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

A LOS ASISTENTES A LOS CURSOS DEL CENTRO DE EDUCACION CONTINUA

Las autoridades de la Facultad de Ingeniería, por conducto del Jefe del Centro de Educación Continua, otorgan una constancia de asistencia a quienes cumplan con los requisitos establecidos para cada curso. Las personas que deseen que aparezca su título profesional precediendo a su nombre en la constancia, deberán entregar copia del mismo o de su cédula a más tardar el SEGUNDO DIA de clases, en las oficinas del Centro con la señorita encargada de inscripciones.

El control de asistencia se llevará a cabo a través de la persona encargada de entregar las notas del curso. Las inasistencias serán computadas por las autoridades del Centro, con el fin de entregarle constancia solamente a los alumnos que tengan un mínimo del 80% de asistencia.

Se recomienda a los asistentes participar activamente con sus ideas y experiencias, pues los cursos que ofrece el Centro están planeados para que los profesores expongan una tesis, pero sobre todo, para que coordinen las opiniones de todos los interesados constituyendo verdaderos seminarios.

Es muy importante que todos los asistentes llenen y entreguen su hoja de inscripción al inicio del curso. Las personas comisionadas por alguna institución deberán pasar a inscribirse en las oficinas del Centro en la misma forma que los demás asistentes, entregando el oficio respectivo.

Con objeto de mejorar los servicios que el Centro de Educación Continua ofrece, al final del curso se hará una evaluación a tráves de un cuestionario diseñado para emitir juicios anónimos por parte de los asistentes.

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
FACULTAD DE INGENIERIA
DIVISION DE ESTUDIOS SUPERIORES
CENTRO DE EDUCACION CONTINUA
DIRECTORIO GENERAL

REGISTRO DE ASISTENTES Y PROFESORES.

NOMBRE DEL CURSO: ______________________ FOLIO [ ][ ][ ][ ][ ] (1–5) CLAVE ASOC [ ][ ] (6 7)

[ ][ ][ ][ ][ ][ ] (8–13) [ ] ... [ ] (14–41) NOMBRE(S) APELLIDO PATERNO APELLIDO MATERNO

REG. FED. CAUS. [ ] (42–51) CED. PROF. [ ] (52–58)

TEL. PARTICULAR [ ] (59–65) TEL. OFICINA [ ] (66–72) EXTENSION [ ] (73–76)

MARQUE CON UNA CRUZ

ASISTENTE [ ] PROFESOR [ ] [ ] 77 [1] 80

[ ] ... [ ] 8 DOMICILIO PARTICULAR (CALLE. NUMERO Y No. INTERIOR) 41

[ ] ... [ ] 42 COLONIA 71 Z.P. [ ][ ] 72 73

______________________ ESTADO [ ][ ] 74 75

______________________ TITULO PROFESIONAL [ ][ ] 76 77 ______________________ ESPECIALIDAD [ ][ ] 78 79 [2] 80

[ ] ... [ ] 8 DOMICILIO DE OFICINA (CALLE, NUMERO Y No. INTERIOR) 41

[ ] ... [ ] 42 COLONIA 71 Z.P. [ ][ ] 72 73

______________________ ESTADO [ ][ ] 74 75 [3] 80

SOCIACIONES A LAS QUE PERTENECE.

______________________ [ ][ ] ______________________ [ ][ ]

______________________ [ ][ ] ______________________ [ ][ ]

FCEC-01

MANEJO DE SISTEMAS DE INFORMACION GEOGRAFICA.

Programa de la segunda semana 21-25 noviembre 1977

Lunes 21

18:00-18:30	Entrega de las prácticas de laboratorio. Explicación de la práctica con el sistema IMGRID para el manejo de información de millas.
19:00-20:00	Demostración del equipo básico de graficación de la firma Tektronix. American Trade Center Liverpool 31 tel. 591-01-55 Dr. Douglas A. Thorson.
20:00-21:00	Codificación de la práctica con IMGRID.

Martes 22

18:00-19:30	Procesamiento de información geográfica. Interpolación. Métodos de SYMAP. Equipos de computación gráfica.
19:30-19:40	Descanso. Cafetería
19:40-21:00	Prácticas. Utilización de barreras con SYMAP. Modelos de IMGRID, utilizando información geográfica.

Miércoles 23

18:00-19:00	Descripción del sistmea de percepción remota del IMAS, UNAM.
19:00-20:00	Prácticas del laboratorio PR. Grupo 1 Prácticas de laboratorio SYMAP, IMGRID. Grupo 2
20:00-21:00	Prácticas de laboratorio PR. Grupo 2 Prácticas SYMAP, IMGRID. Grupo 1. Dr. Armando Jinich, y Dra. Rosa Ma. S to Ciudad Universitaria, IMAS.

Jueves 24

18:00-10:00	Bancos de Datos geográficos. El sistema CETENAL. Interconexión de sistemas de información geográfica con sistemas de percepción remota. Dr. Adolfo Guzmán Arenas.

20:00-20:10	Descanso. Cafetería.
20:10-21:00	Entrega de resultados y corrección de prácticas.

Viernes 25

18:00-19:30	Mesa redonda y conclusiones
19:30-20:00	Entrega de resultados de las prácticas.
20:00-20:30	Entrega de constancias y clausura del curso
20:30-21:30	Convivio.

DIRECTORIO DEL CURSO MENEJO DE SISTEMAS DE INFORMACION GEOGRAFICA EN PLANEACION

ARQ. BRAULIO HORNEDO ROCHA
INVESTIGADOR
D E S F I
SECC. DE PLANEACION
UNAM
TEL.: 550.52.15 E.4492

MAT. MARIO RODRIGUEZ GREEN
INVESTIGADOR
SECC. DE PLANEACION
D E S F I
UNAM
TEL.: 550.52.15 E. 4492

ING. ALBERTO TORFER MARTELL
INGENIERO DE SISTEMAS
I B M DE MEXICO, S.A.
MARIANO ESCOBEDO 595
MEXICO, D.F.
TEL.: 250.90.11

ARQ. ALEJANDRO VILLANUEVA EGAN
INVESTIGADOR
SECCION DE PLANEACION
D E S F I
UNAM
TEL.: 550.52.15 E. 4492

'edcs. 24, XI, 77.

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

MANEJO DE SISTEMAS DE INFORMACION GEOGRAFICA EN PLANEACION

SISTEMAS DE INFORMACION GEOGRAFICA

ING. ALBERTO TORFER MARTELL

NOVIEMBRE, 1977.

SISTEMA DE INFORMACION GEOGRAFICA

VISION GENERAL.-

El propósito de este artículo es el de señalar la finalidad de los sistemas de información geográfica, establecer un marco de referencia para clasificarlos y describir algunos sistemas típicos.

I N D I C E

DEFINICION

Un sistema de información geográfica o sistema de geoprocesamiento es todo aquel sistema de computación donde la información que se maneja tiene el atributo de posición geográfica de la entidad generadora de la información, haciendo uso de ella de forma eficiente, para relacionar la información entre sí.

Algunos ejemplos de las operaciones que se pueden realizar son:

a).- la correlación de información, donde el sistema nos puede proporcionar automáticamente los datos de las entidades vecinas, o de las que están comunicadas entre sí, por carretera, telefono u otro medio.

b).- la generalización, mediante la cual podemos agregar la información para que esta, a un nivel menor de detalle, nos refleje las condiciones de áreas cada vez más extensas.

El objetivo de almacenar la información con su posición geográfica es el de establecerla como marco de referencia, que resulta universal. Otro marco de referencia es el del momento en el que se capta la información. Identificándola con estos dos parámetros tendremos un identificador que no puede ser duplicado, ya que en un momento determinado y en un lugar determinado solo pueda existir una sola entidad.

La identificación precisa del lugar y la fecha permiten establecer un código que puede servir de manera unívoca para la identificación de -

recursos naturales, de infraestructura y humanos. El conocimiento de ellos es esencial en las labores administrativas y de planeación de gran parte de las labores institucionales cuya actividad tenga una dispersión geográfica.

TOMA DE DECISIONES

Las causas fundamentales por las cuales se está recorriendo cada vez en mayor medida a los sistemas de información geográfica son básicamente dos:

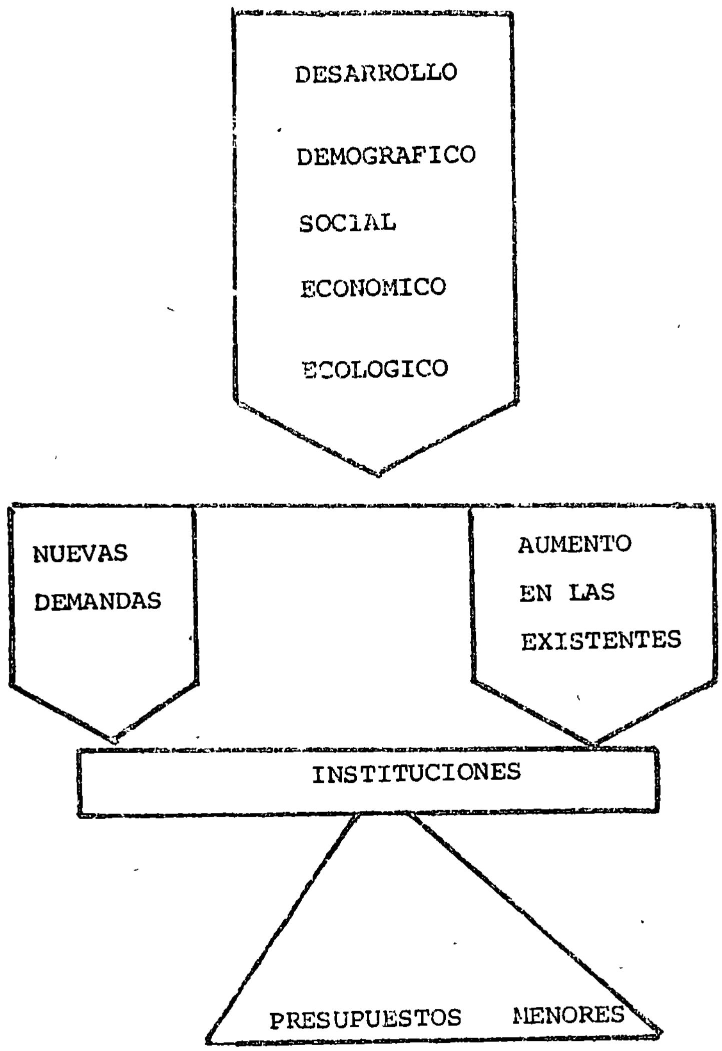
- El crecimiento en la demanda de todo tipo de servicios.
- la cada vez mayor compleja interacción entre los factores que conforman a una sociedad.

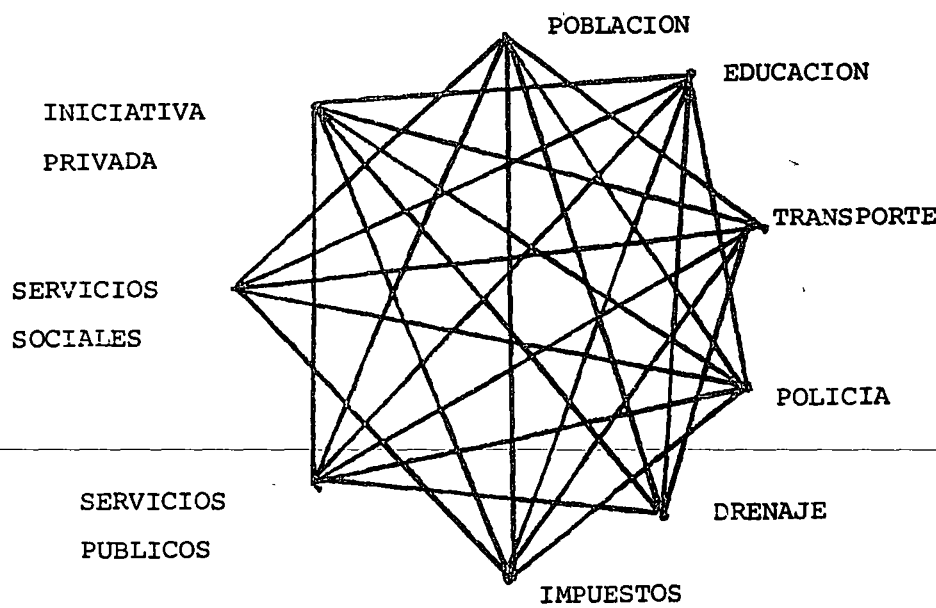
El desarrollo demográfico, económico y social de nuestros tiempos está creando por un lado nuevas necesidades, pero también está aumentando la demanda de las ya atendidas anteriormente.

A esto se suma la reducción proporcional de los presupuestos, pues si bien éstos crecen, no lo hacen en la misma proporción de la demanda.

Es interesante señalar que para el año 2,000 se tendrá que construir un número de viviendas y servicios equivalente al construido hasta la fecha, dado que para el fin de siglo la población será el doble de la población actual.

Por otra parte, todos los factores que conforman a una sociedad interactuan entre sí y las decisiones de cualquier sector y a cualquier ni


DESARROLLO
DEMOGRAFICO
SOCIAL
ECONOMICO
ECOLOGICO
NUEVAS DEMANDAS
AUMENTO EN LAS EXISTENTES
INSTITUCIONES
PRESUPUESTOS MENORES



INTERACCION DE LAS ACTIVIDADES

vel afectan a todos los restantes.

Así por ejemplo, si se construye una industria en una población esta tendera a atraer a la población lo que hará que se requieran servi-- cios públicos, sociales, viviendas, etc.

Un ejemplo reciente de la complejidad de esta interacción es el vivi do recientemente con la construcción de la siderúgica Lázaro Cárde-- nas Las Truchas, dunde al principio no se pudo prever con exactitud ni se pudo coordinar la interacción de las diversas dependencias que proporcionaron los servicios requeridos por la afluencia de habitan- tes, presentándose fenómenos de escacez, carestía y especulación.

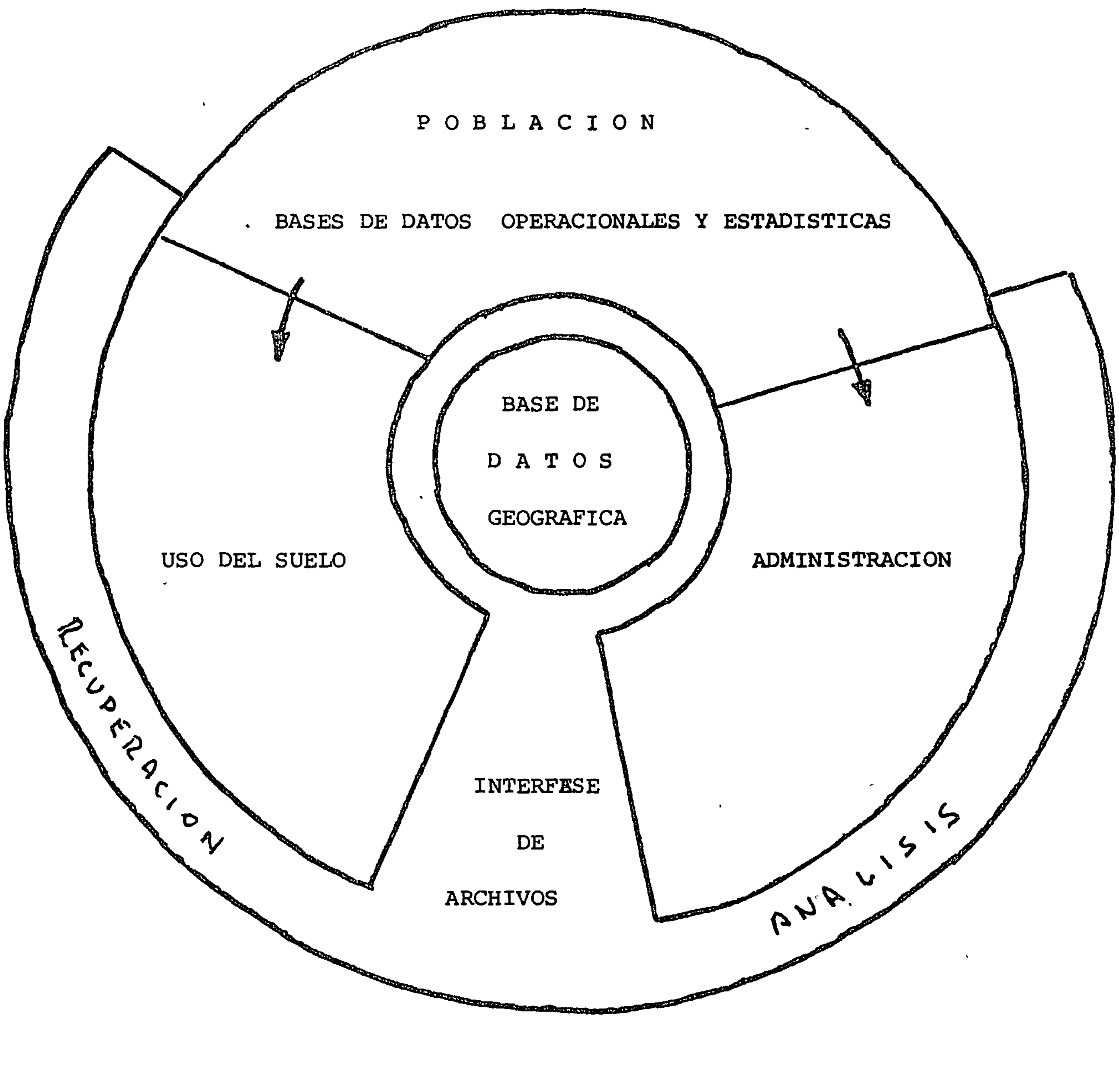
Es por estas causas que se requeria de sistemas que permitan la más efectiva y ágil planeación. Para efectuar una buena planeación se - requiere de información y esta debe integrarse de manera coherente.

INTEGRACION DE INFORMACION

La mejor forma de lograr esto es a través de una base de datos geo- gráfica que permita, a través de un lenguaje de interfase, la recupe ración y el análisis de la información.

Básicamente las decisiones que respalde dicho sistema son con respec to al uso del suelo y en decisiones administrativ as de las institu- ciones que tengan sus actividades dispersas geográficamente.

Un concepto fundamental que debe respetar esta interfase entre las -



INTEGRACION DE INFORMACION CON BASE GEOGRAFICA bases de datos es de que se deben de poder aprovechar las bases de datos operativas de las instituciones con poco o ninguna necesidad de intervención manual.

NIVELES DE INFORMACION

Por otra parte, el sistema de información geográfico debe adaptarse a los diferentes tipos de decisiones que se puedan presentar: Esto - es, decisiones operativas, administrativas, tacticas, estratégicas y políticas. Para esto debe poder agregar la inform ación desde los niveles más detallados a los niveles más generales de manera automá tica.

El objetivo de esto es el de que la información con la que se toman las decisiones políticas sea la misma que con la que se toman las de cisiones técnicas y administrativas.

En ocasiones hay una interrupción en el flujo de información y las de cisiones no son consistentes porque la información no lo es. Un ejem plo de agregación sería en relación a un marco de referencia donde - se tuviera, la división de país en estados, municipios, localidades y predios. Esta serviría para toma de decisiones políticas, estraté gicas, tacticas, administrativas y operativas para todo el país.

Para un área urbana se podrían tener los niveles de áreas metropoli- tana, delegación, sección, predio y facilidades.

No hay que olvidar que las divisiones a nivel nacional afectan a las regiones y que en general no existen a la fecha mecanismos adecuados

N I V E L E S

AGREGACION DE INFORMACION

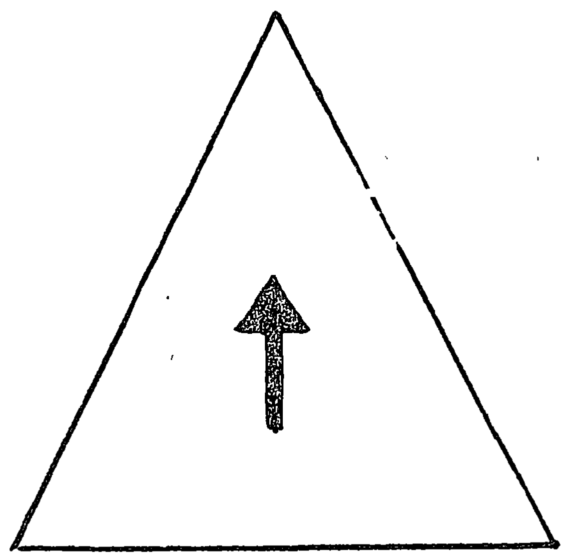
PAIS O AREA METROPOLITANA

ESTADO O DELEGACION

MUNICIPIO O SECCION

LOCALIDAD O MANZANA

PREDIO



TOMA DE DECISIONES

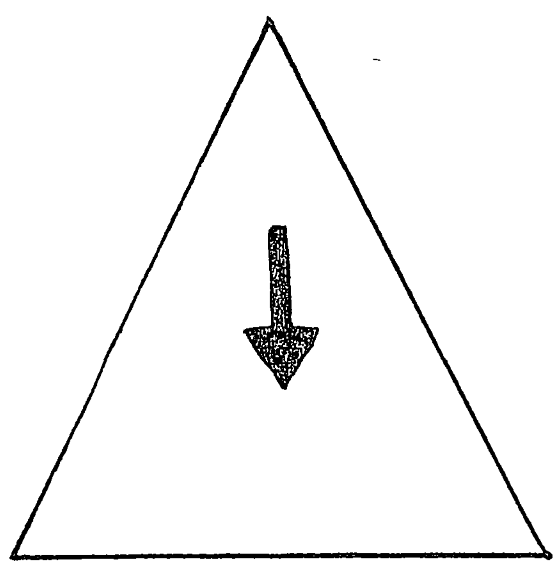
POLITICAS

ESTRATEGICAS

TACTICAS

ADMINISTRATIVAS

OPERATIVAS



para determinar el impacto de una decisión a nivel nacional en el ámbito regional.

Los sistemas geográficos pueden obviar este problema, ya que se pueden tener sistemas de cubrimiento regional o nacional

Las clases de información donde la información requiere de la identificación geográfica son muchas y muy variadas. A efecto ilustrativo mencionaremos las siguientes:

Recursos Físicos	- Todo tipo de recursos naturales - bosques - lagos - tierras de cultivo
Recursos humanos	- población - médicos - maestros
Obras de Infraestructura	- puentes - líneas telefónicas - caminos - distritos de riego - pozos
Producción	- agrícola - ganadera

- forestal
- de servicios

En fin, se puede concluir que toda actividad humana se desarrolla dentro de un contexto geográfico, si bien hay algunos en los cuales este contexto no es significativo.

ETAPAS DEL GEOPROCESAMIENTO

Ahora bien, dentro del proceso de utilización de información geográfica tenemos tres etapas bien definidas que son:

ETAPA I.- Analizar las relaciones físicas entre los factores del medio ambiente.

- uso del suelo
- características físicas
- obras de infraestructura

ETAPA II.- Se complementa con información estadística y socioeconómica de bases de datos institucionales.

- población
- valor de la propiedad

ETAPA III.- Se analiza el efecto de las políticas y cambios propuestas en:

- el uso del suelo
- facilidades
- servicios

Todo con objeto de coadyuvar a la toma de decisiones a diferentes niveles con respecto a la

- asignación de recursos
- ubicación de facilidades
- determinación del uso potencial del suelo
- análisis de la demanda de servicios

DESARROLLO DE UN PROYECTO

De lo anteriormente expuesto pueden verse algunas de las implicaciones del geoprocesamiento que harán que para que éste tenga éxito, se cumplan las premisas que se mencionarán posteriormente.

Las implicaciones son:

a).- el geoprocesamiento cruza límites funcionales y jurisdiccionales.

b).- la información puede provenir de una gran variedad de fuentes.

c).- los usos pueden ser multiples.

d).- puede requerirse de más información de la actualmente disponible.

e).- se requiere de asignación de recursos financieros, materiales y humanos.

De esta resultan los siguientes requisitos para la aplicación exitosa del geoprocesamiento:

a).- se debe formar un grupo representante de las áreas usuarias que determine tanto las necesidades que debe cumplir el sistema como la información que el sistema puede absorver directamente.

b).- apoyo a un alto nivel directivo que coordine a los usuarios y que puede decidir sobre la asignación de recursos al proyecto.

TIPOS DE INFORMACION.-

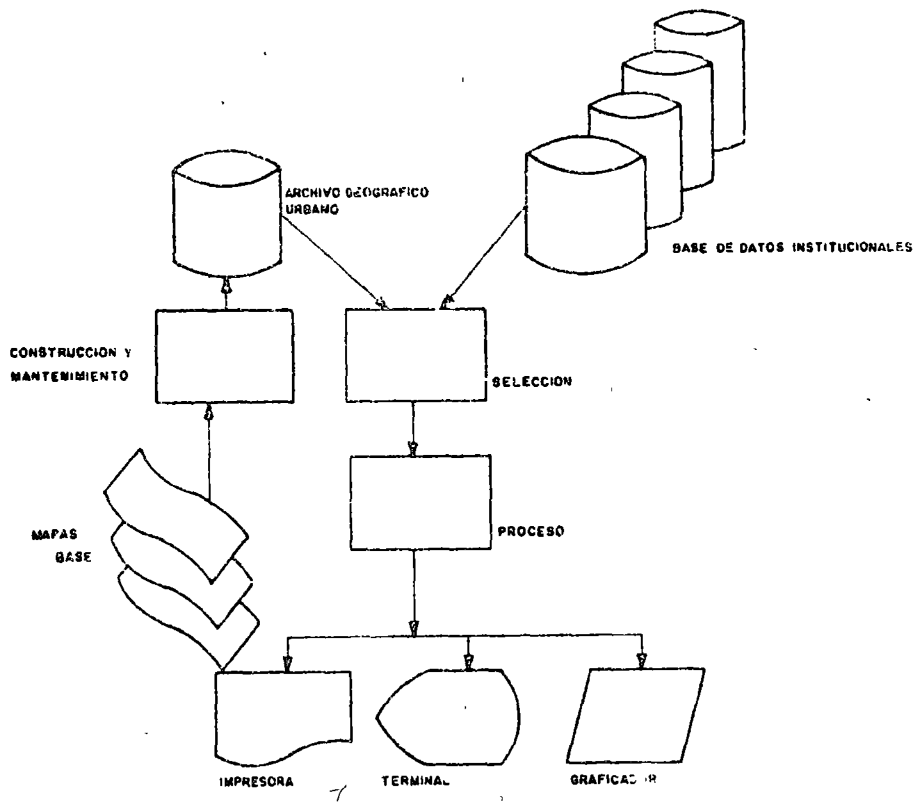
Básicamente, la información que puede utilizar un sistema de geoprocesamiento puede clasificarse como información:

Nominal - que es aquella que esta localizada por nombres geográficos y cuyo ejemplo más relevante es el de la información censal.

Ordinal - que es la información localizada por coordenadas, como la representada en mapas.

De aquí resulta que los procesos del sistema son:

- construcción y mantenimiento de la base de


ARCHIVO GEOGRAFICO URBANO
BASE DE DATOS INSTITUCIONALES
CONSTRUCCION Y MANTENIMIENTO
SELECCION
PROCESO
MAPAS BASE
IMPRESORA
TERMINAL

datos geográfica

- extracción de información, según el problema a resolver y proceso de la información.

Los resultados pueden darse en forma de listados de impresora o mapas realizados por impresora, resultados por pantalla, de caracteres alfanuméricos o de vectores y graficadores.

NIVELES DE DETALLE

De los diferentes niveles de detalle de la información se puede establecer una clasificación de sistemas de geoprocesamiento.

Los sistemas pueden ser a nivel regional o urbano dentro de cada uno se pueden tener los niveles de ingenieria, de límites, de localización o el geoestadístico.

Para entender mejor esta última división hablaremos de los diferentes niveles:

NIVEL DE INGENIERIA URBANO.-

Este nivel esta constituído por la información acerca de la posición de obras de infraestructura para proporcionar los servicios urbanos como por ejemplo, ubicación de líneas eléctricas, de agua potable y de drenaje dentro de una ciudad.

Generalmente las compañías de servicios públicos cuentan con mapas que indican la ubicación de sus facilidades. El objeto de geocodificar esta información puede ser por ejemplo, el de correlacionar la infor-

	INGENIERIA	LIMITES	LOCALIZACION	ESTADISTICO
NIVEL REGIONAL				
NIVEL URBANO				

mación de demanda de servicio con la de facilidades, con objeto de planear mejor el mantenimiento y extensión de las facilidades en función del crecimiento de la demanda.

NIVEL DE LIMITES URBANOS.-

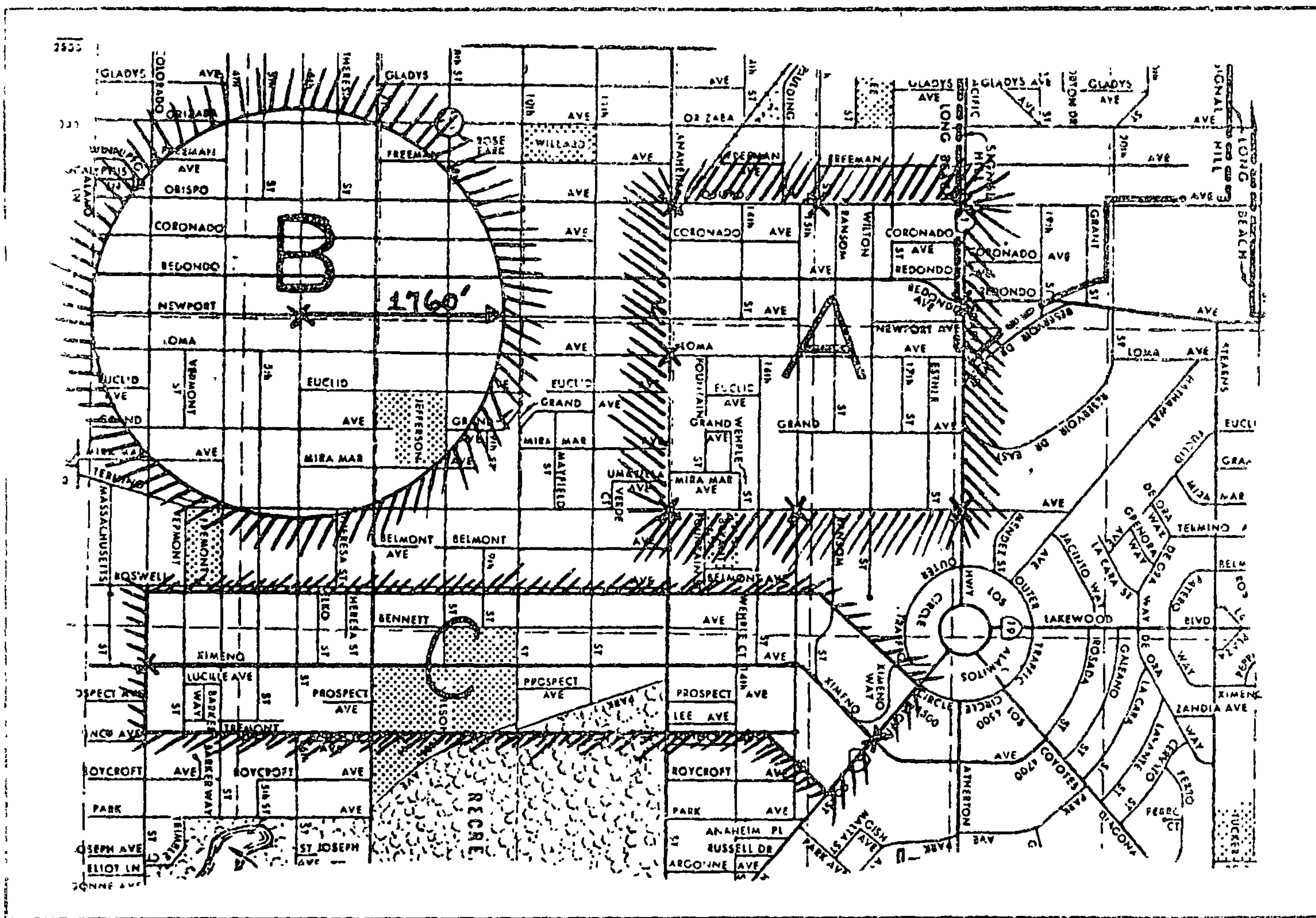
Este nivel es el resultado de un catastro urbano y el objeto de geocodificarlo puede ser por ejemplo, la determinación de los predios, y el valor de los mismos, que se verán afectados por la construcción de una nueva avenida. También puede servir para determinar a cuales -- predios se les puede revaluar en función de el mejor servicio que se les proporciona con esa avenida, etc.

NIVEL DE LOCALIZACION-URBANO.-

Es un nivel donde la información está agregada a nivel de manzana y puede servir para determinar el mercado potencial de un nuevo conjunto de comercios o de agencias gubernamentales. Puede servir para efectuar una regionalización para efectos censales o de padrón de - votantes, etc.

NIVEL GEOESTADISTICO.-

En este nivel ya se tiene la información agregada a nivel de sección o cuartel y su uso puede enfocarse a la comparación relativa de diversas secciones en cuanto a nivel de población y de facilidades y servicios, etc.


B
1760'
A
C
GLADYS
COLORADO
ORIZABA
FREEMAN
OBISPO
CORONADO
REDONDO
NEWPORT
LOMA
EUCLID
GRAND
MIRA MAR
TERMINO
MASSACHUSETTS
VERMONT
ROSWELL
BELMONT
BENNETT
XIMENO
LUCILLE AVE
PROSPECT
TREMONT
ROYCROFT
PARK
ST JOSEPH
ELIOT LN
JEFFERSON
WILLARD
ROSE PARK
MAYFIELD
UMATILLA
VERDE CT
WEHRLE
RANSOM
WILTON
ESTHER
GRANT
LONG BEACH
SIGNAL HILL
NEWPORT AVE
RESERVOIR DR EAST
LOS ALAMITOS
TRAFFIC CIRCLE
LAKEWOOD
BLVD
ROSADA
GALEANO
LA CARA
DE ORA WAY
ATHERTON
COYOTES
LEE AVE
ANAHEIM PL
RUSSELL DR
ARGONNE
ZANDIA AVE
RECRE

A nivel regional se pueden tener los mismos niveles de detalle de información: el nivel de ingeniería, el de límites, de localización y el estadístico.

NIVEL DE INGENIERIA RURAL.-

Este nivel puede considerarse constituído por modelos digitales del terreno. Se pueden tener MDT que cubren áreas muy extensas con poca precisión, lo que haría que no se pudiesen clasificar en esta categoría.

En algunos países se tienen bases de datos cartográficos a escalas 1:2000, pèro son realmente escasas.

NIVEL LIMITE-REGIONAL.-

En este sistema el nivel de información es el proporcionado por un catastro rural, que puede efectuarse a una escala entre 1:10 000 y 1:20 000. Aquí se tiene la delimitación de todos los predios rurales y su utilización puede ser semejante al correspondiente urbano, se puede utilizar para la determinación de los predios afectados por una presa o una carretera. Eventualmente, a través de la información recopilada a este nivel se podrían tener mapas de uso del suelo.

NIVEL LOCALIZACION-REGIONAL.-

En la actualidad, se están desarrollando bancos de datos de recursos que están a la escala 1:25 000 a 1: 250 000 elaborados a través de -

fotogrametría y fotointerpretación o por medio de clasificación automática de sátelites.

Este tipo de Banco de Datos pueden servir para la planeación del uso del suelo, la determinación del área cultivada, etc.

NIVEL GEOESTADISTICO-REGIONAL.-

En este nivel, se empieza a usar una división política para agregar la información, tanto de recursos como socioeconómico. Las actividades para las que puede usarse este tipo de Banco de Datos son genéricas. Es decir, podemos seleccionar entidades que cumplan una serie de requisitos, por ejemplo, todos aquellos cuya población sea mayor de 100000 habitantes y que sólo se tengan 10 hospitales, etc.

En general, el costo del desarrollo de los sistemas varía geometricamente según el nivel en el que se trabaje. Así, si un sistema geoestadístico para todo el país cuesta 1, el nivel siguiente costará 2, el siguiente 4 y el siguiente 8.

En general los geostadísticos manejan solo información nominal y los otros manejan información ordinal con cada vez mayor grado de detalle.

Las ventajas de los niveles más detallados es que siempre podemos agregar información para alimentar el siguiente nivel, mientras que lo contrario es imposible.

Así, si se desarrolló un sistema de límites rural, con costo de 4 - podremos tener también sistemas de localiaación y estadísticos, que si se hubieran realizqdo independientemente nos costarían 3 unidades sin disponer del nivel de límites rural.

Es por lo tanto muy importante para un país determinar la convenien cia de empezar un sistema a nivel límites-rural o uno de localización, decisión nada fácil de tomar y que hasta ahora no se había planteado dado que el desarrollo de los sistemas geográficos de información que permiten la generalización automática se han desarrollado muy recientemente y no hay conciencia clara de esta alternativa ni se han estudiado cuidadosamente sus factores e implicaciones.

CLASIFICACION

Los sistemas también pueden clasificarse en subsistemas de construcción y mantenimiento, de extracción y proceso.

A continuación se describen algunos sistemas típicos de algnos niveles:

a).- Banco de Datos Geográfico de CETENAL

La información que contiene este banco es la de los mapas elaborados por DETENAL sobre topografía, geología, edafología, uso del suelo y uso potencial.

El formato de la información almacenada permite tener una precisión variable de acuerdo a las necesidades de detalle de la zona geográfica - correspondiente. La actualización de estas propiedades se facilita por el diseño del formato y el usuario puede elaborar cualquier tipo de --

pregunta que se puedan descomponer en una combinación booleana de preguntas o predicados elementales.- Su nivel es de localización regional.

Como ejemplo de las preguntas que el sistema puede responder tenemos:

¿ qué pueblos tienen más de mil habitantes y no tienen medios de comunicación?.

¿ cuáles son los mejores lugares para pasar una carretera, ¿ qué pueblos esten a menos de diez kilométros de una aeropista, ¿ cuántos kilométros cuadrados de pastizal inducido hay sobre terrenos de aluvión en una cierta zona?, etc.

La República Mexicana se dividió en 2400 zonas cada una de 1000 km2., cada zona se dividió en 25 cuadros y éstos a su vez se dividieron en 12 subcuadros de 1' x 1', equivalente a 1.8 km., aproximadamente.

Para cada cuadro y de cada materia se codifican los siguientes tipos de propiedades:

Superficiales.- Son todas aquellas definidas por un área, como el tipo del suelo, la clasificación geológica o el uso del suelo.

Lineales.- Definidas por una línea que puede ser recta, curva o irregular como carreteras, líneas de telefonos, etc.

Particulares.- Son aquellas definidas por un punto, como por ejemplo, pueblos, depositos de agua, minas, etc.

Para la consulta se tieen funciones relacionadas y rutinas, además de funciones lógicas. Algunas de las funciones relacionadas son:

MAYORA.- Determina zonas, cuadros o subcuadros donde el porcentaje de una propiedad sea mayor que una cierta entidad.

MENORQ.- Que se amenor

ENTRE.- Que el valor de la propiedad se encuentre entre 2 valores que se indican a continuación.

DEFIRE.- Que el valor sea diferente.

IGUALQ.- Que el valor sea igual.

BUSCA.- La función BUSCA nos indica la función de búsqueda y el nivel el cual se hace la consulta.

Algunas de las funciones lógicas son:

PRO (A1,A2,A3).- Búsqueda de propiedades.

donde:

A1.- Número de la propiedad buscada.

A2.- Función relacional.

A3.- Valor dado.

EVALUA.- Sirve para hacer una evaluación de las propiedades de un lugar.

CERCA (A1, A2,A3,A4).- Permite relacionar propiedades de diferentes lugares y consta de 4 argumentos

A1.- Primer arreglo.

A2.- Segundo arreglo.

A3.- Distancia deseada.

A4.- Nivel de búsqueda.

PUEBLO.- (A1) Sirve para encontrar los pueblos por medio de un número previamente asignado.

HAYVIA (A1,A2).- Determina que tipo de caminos o comunicaciones que llegan a un publo.

A1.- Es el tipo de comunicación.

A2.- Es el número del pueblo.

SERPOB.- (NUM, a, b, c, d, e, f, g) Indica los servicios con que cuenta una población, de acuerdo a la clasificación de la carta de Uso del Suelo de DETENAL.

Todas estas funciones se pueden combinar por medio de parámetros booleanos como el AND, OR, y NOT. Pudiendose construir preguntas complejas.

b).- Sistema Geomunicipal de Información.-

Este sistema es del tipo geoestadístico-regional y la unidad básica de información es el municipio, por lo tanto, cualquier información que se desee alimentar a la base debe cumplir con la restricción de estar levantada a nivel municipal (1970).

Por medio de este sistema se pueden conocer la información de municipios o estados, construir conjuntos de municipios que cumplan con ciertas condiciones especificadas por el usuario, obtener reportes con información y formatos seleccionados, crear formulas e índices y aplicarlas a los municipios. También se puede analizar los municipios con respecto a sus vecinos inmediatos.

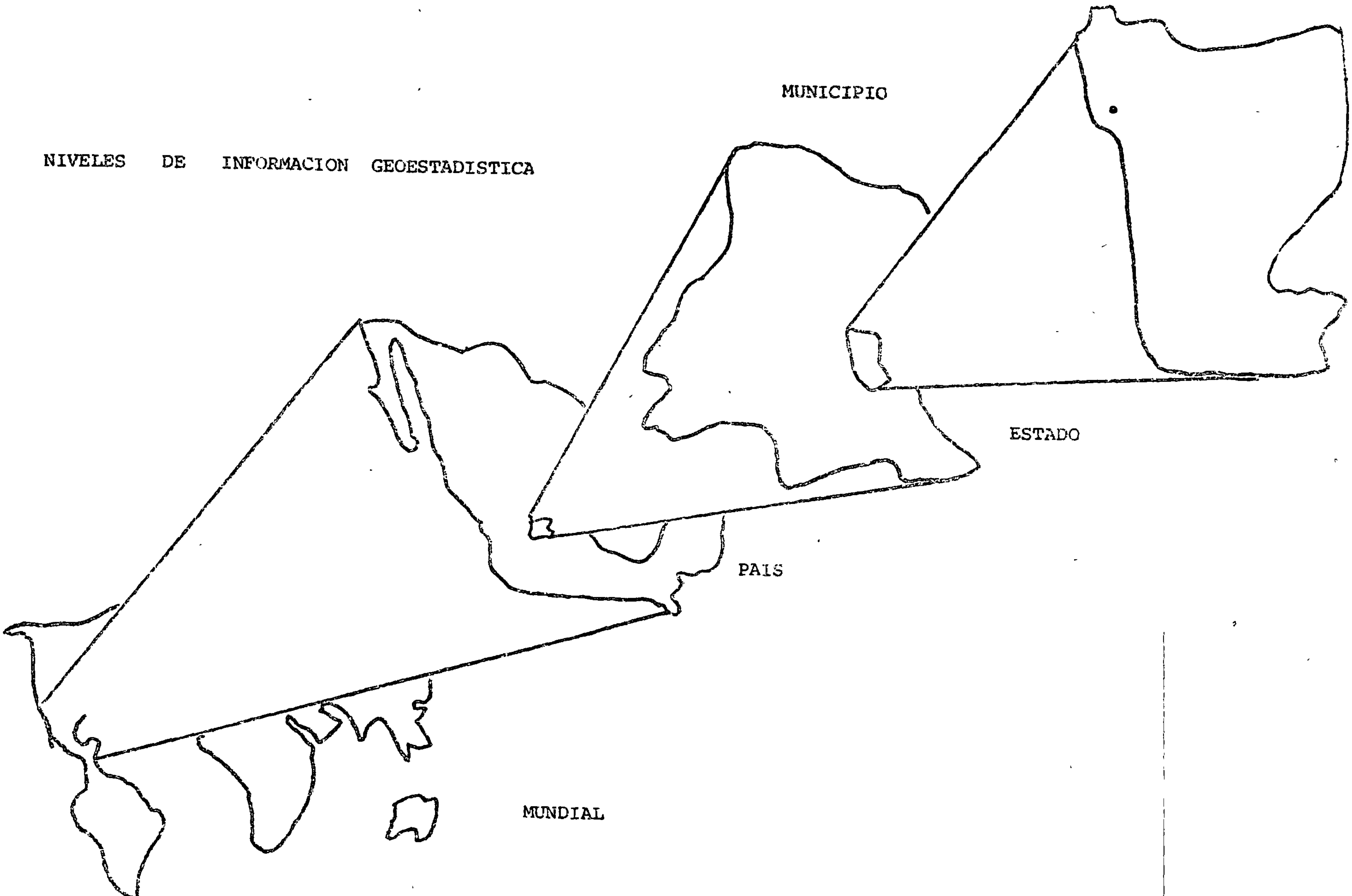
El sistema tiene las siguientes funciones:

CONSULTA.- Se tacha el número del municipio y a continuación aparecen los datos del mismo.

También se puede consultar por Estado.

VECINOS.- Nos da el nombre y clave de municipios vecinos.

NIVELES DE INFORMACION GEOESTADISTICA



CONJUNTO 1.- Agrega municipios que cumplan con una condición dada.

MUESTRA.- Nos da las claves de los municipios incluidos en un conjunto.

CONJUNTO 2.- Sirve para unir, intersectar u obtener el complemento de dos conjuntos existentes.

CONJUNTO MANUAL.- Permite definir un conjunto arbitrario.

DEFINE.- Permite construir expresiones algebraicas utilizando cualquier tipo de índice o variable definida anteriormente.

DESPLIEGA.- Muestra los conjuntos, entidades, formatos, campos o índices que están en los directorios.

FORMATO.- Para definir el formato de impresión y las variables que se imprimirán.

IMPRIME.- Imprimir un conjunto con un formato definido.

CARACTERIZA.- Se puede obtener para una variable o índice el valor total, máximo, mínimo y promedio.

ANALIZA.- Sirve para obtener un diagrama que indica la situación del municipio en relación a sus servicios con respecto a un o más variables.

ESTRATIFICA.- Separa en diversos estractos un conjunto utilizando algunas variables o índice como criterio de estratificación.

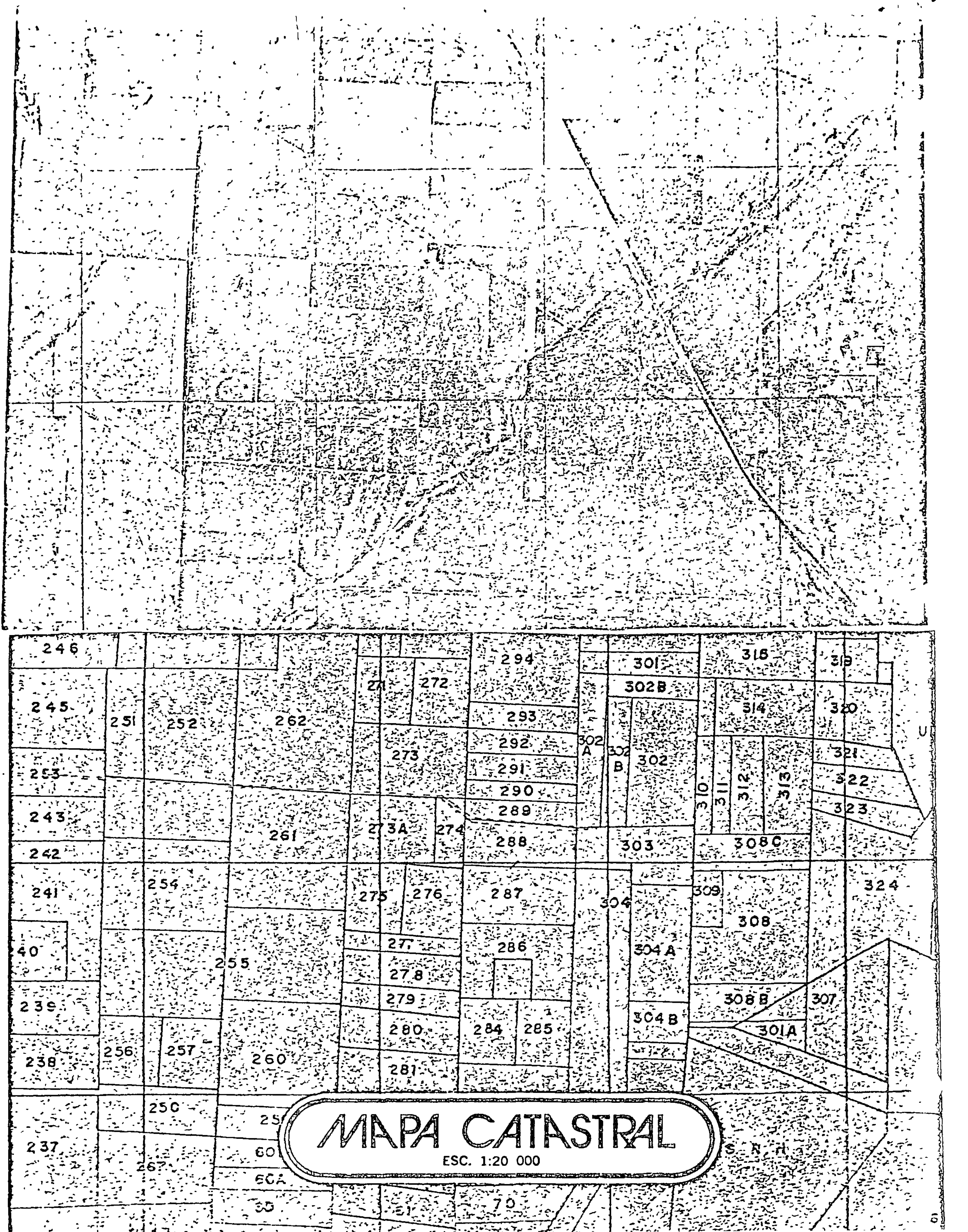
El sistema actualmente (el Sistema Geomunicipal) tiene información de los censos de población, agrícolas y ganaderos y de planeación educativa. Se le ha estado agregando otro tipo de información como por ejemplo la del Banco de Crédito Rural, etc.

<u>Sistema de Información de uso del suelo y registro de la propiedad de las provincias marítimas de Canadá.</u>-

Las provincias marítimas de Canadá se unieron para lograr los siguientes objetivos:

a).- Realizar la geodesia y mapeo integrado como base para toda la información requerida para todo el desarrollo ligado al suelo.

b).- Reemplazar sus sistemas de registros manuales de la propiedad con un sistema en línea.


MAPA CATASTRAL
ESC. 1:20 000
246
245
251
252
262
272
294
293
292
291
290
289
288
273
273A
274
261
301
302B
302 A
302 B
302
303
318
319
314
320
321
322
323
310
311
312
313
308C
253
243
242
241
254
275
276
287
304
309
308
324
240
255
277
278
279
280
281
286
284
285
304 A
304 B
308 B
307
301A
239
238
256
257
260
250
237
267
60A
70

c).- Examinar la viabilidad de un Banco de Datos del uso del suelo.

Para lograr lo anterior LRIS (Land Registration and Information Service) se estableció en 4 etapas :

ETAPA 1.- Establecer los puntos de centros secundarios con mayor densificación que la proporcionada por la Red Geodésica Nacional.

ETAPA 2.- Producir los mapas planimétricos, topográficos y de propiedades a las escalas requeridas para cumplir con las demandas generales por la administración de recursos, desarrollo urbano e identificación de propiedades.

ETAPA 3.- Reemplazar el sistema de registro de la propiedad con un nuevo sistema que aproveche la tecnología moderna y la información de las fases anteriores.

ETAPA 4.- Establecer un banco de datos para las estadísticas geográficas, usando la información de los predios, obtenidas en las fases anteriores.

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

SISTEMA DE INFORMACIUN GEOGRAFICA URBANA

VISION GENERAL

Ing. Alberto Torfer Martell

Ana Elena Ferrer Ramírez

RESUMEN

Se presenta un sistema de información urbana con geocodificación que permite la integración de información de diversas fuentes, facilitando su agregación y presentación, especialmente en función de la distribución geográfica. Realizar este tipo de operaciones por medios manuales o sin un marco de referencia adecuado, consumiría cientos de horas hombre.

El sistema puede ser una herramienta muy útil tanto en la toma de decisiones administrativas como en la planeación.

C O N T E N I D O

CAPITULO I.-

INTRODUCCION

El medio ambiente ecológico, social y económico actual, está planteando demandas nuevas e imprecedentes, en las posibilidades administrativas y de planeación.

La explosión demográfica está provocando esfuerzos en nuestro medio ambiente. Durante los últimos 50 años se han producido más bienes, que durante todos los años anteriores registrados en la historia. Los vehículos están saturando las vías de comunicación. El aumento de fraccionamientos está inundando los sistemas de alcantarillado. El gran consumo de energéticos está contaminando el medio ambiente.

La efectividad de los gobiernos locales, se deteriora constantemente mientras las demandas siguen en aumento. Por ejemplo, la cantidad de basura por persona está aumentando constantemente. Al mismo tiempo, el tamaño y densidad del área metropolitana aumenta el problema de los basureros. En algunos casos se debe transportar la basura largas distancias para llegar a las áreas apropiadas. A esto se une el aumento de crímenes, necesidades de más escuelas, agua potable, alcantarillado, etc.

Todos estos problemas recaen en los administradores. Tienen que resolver los problemas duales del deterioramiento de la productividad y el incremento de demandas de nuevos servicios.- todo esto con presupuestos reducidos.

De acuerdo con estos problemas, se plantean interrogantes cada vez más complejas. Muchas de estas preguntas requieren datos de múltiples áreas geográficas y funcionales. Por ejemplo:

> ¿Qué impacto tendrá un nuevo fraccionamiento, sobre las escuelas, transportes, protección policiaca, drenaje, agua y alcantarillado?.
>
> ¿Cuál sería el impacto probable de una orden de reordenamiento urbano?.
>
> ¿Cómo se puede incrementar la productividad de los sistemas de protección contra incendios, con objeto de satisfacer las nuevas demandas?.

¿Cómo se deben desplegar los recursos actuales de patrullas policiacas para cubrir los cambiantes patrones sociales y criminales, sin sacrificar la seguridad pública?

Estos complejos problemas a nivel político no se pueden decidir con sistemas de información que únicamente proporcionan una orientación fucional. Se deben recolectar datos a través de fronteras funcionales y geográficas. Si no se tienen las respuestas adecuadas, es imposible la administración y planeación del medio ambiente ecológio, social y económico, actualmente complejo.

CAPITULO II.-

GEOCODIFICACION URBANA

La Geocodificación Urbana es una técnica reciente de procesamiento de datos, que se ha desarrollado para satisfacer los requerimientos de información de las agencias gubernamentales estatales y locales. Básicamente, relaciona la localización geográfica con las bases de datos existentes (por ejemplo, la localización geográfica se puede relacionar con los archivos de incidencias de policía, archivos de población escolar, etc.); la geografía es el común denominador que puede ligar virtualmente todos los datos. Los archivos del departamento de policía contienen información acerca de dónde y cuándo ocurren los casos. Los archivos de propiedad, así como los de permisos de construcción, contienen datos acerca de dónde se localizan las cosas. Los archivos de salud contienen información acerca de - quién y dónde, etc., y se podría continuar la lista indefinidamente.

El problema es cómo relacionar estos datos salvando el obstáculo que representa la diversidad de límites administrativos. Todas las agencias tienen sus propios métodos particulares para registrar su información. La solución sería relacionar de alguna manera esta información, con un común denominador (esto es, una localización geográfica). De esta manera, se podría recuperar toda la información perteneciente a esa localización, sin importar en qué base de datos se tenga. Esto es, en esencia, la geocodificación. La geocodificación urbana puede integrar y correlacionar los datos a través de los límites - administrativos de diversas dependencias, añadiendo de esta manera, toda una nueva dimensión al procesamiento de datos. Con la geocodificación, muchas dependencias pueden estudiar las bases de datos extremadamente diversificadas de otras, relacionando los datos en una base de datos geográficos común obteniendo respuestas a preguntas específicas. Anteriormente, bajo condiciones manuales, los complejos proyectos relacionados con el intercambio de datos entre diversas dependencias, hubiera sido prohibitivamente costoso en términos de hora-hombre; inclusive imposibles.

GEOCODIFICACION URBANA.-

La Geocodificación urbana puede ser vista como un proceso en tres pasos:

1) Utilizando mapas se analizan las relaciones físicas o geográficas entre las diferentes características relativamente estables del medio ambiente. ¿Cuáles son los usos básicos del suelo (por ejemplo, granjas, minería, residencial)? ¿Cuáles son las características físicas del área (Terreno, cuerpos de agua, tipo de suelo)? ¿Cuáles son algunos de los rasgos de infraestructura (presas, caminos)?

2) De nuevo, usando mapas, junto con datos estadísticos de varias bases de datos, se determinan los patrones de actividad generados por los diferentes usos del suelo. Cuáles son las rutas de tránsito colectivo, las poblaciones escolares, los valores de propiedad individual?

3) Se analizan los cambios propuestos para uso del suelo, facilidades o servicios que afectarán a estos patrones de actividad. Por ejemplo, ¿Cómo se pueden redistribuir las estaciones de bomberos para dar protección a nuevos edificios actualmente en construcción? ¿Cuál sería el efecto de un nuevo puente o carretera?.

Estos tres pasos podrán consumir incontables horas hombre si se hacen manualmente; se tendría que investigar en archivos enteros de diferentes departamentos. Con el geoprocesamiento de la información, una computadora podría hacer toda la búsqueda y correlación de los datos.

VENTAJAS DE LA GEOCODIFICACION

La ventaja más grande de la geocodificación es su potencial habilidad para incrementar la productividad; supera la suma de la utilidad de las aplicaciones individuales. Puede coadyuvar a mejorar la productividad operacional y de planeación de dependencias funcionalmente independientes, y a la eficiencia de todo el gobierno. A través de la mejor información, se podrá aumentar la eficiencia e inclusive, se podrán reasignar recursos. Esta posibilidad de reasignación es una de las ventajas principales del uso de estos sistemas.

Ya que la Geocodificación Urbana utiliza las bases de datos existentes, se podría perfeccionar la calidad total de las bases de datos individuales, de las siguientes maneras:

- Se podría eliminar la duplicación de una gran cantidad de datos, porque todos los identificadores geográficos serían mantenidos en la base de datos geográficos y no en los archivos de datos operacionales.

- Se podrían estandarizar los identificadores geográficos usados por todos los archivos de datos.

- Se podría checar la calidad e integridad de los datos almacenados.

- Se podría reducir el trabajo tedioso de codificación de identificadores particulares en las transacciones.

Cuando los funcionarios puedan evaluar exactamente las facilidades y servicios públicos requeridos por el desarrollo y puedan comparar estas demandas con los servicios y facilidades disponibles, entonces existirá una base válida para tomar decisiones. Se podrían consolidar fácilmente los datos de diferentes dependencias para formar una base general para toma de decisiones. Por ejemplo, la mayoría de los estados requieren de un informe del impacto del medio ambiente para un mejor desarrollo. Esto puede ser un requisito complejo que involucra datos de muchas dependencias independientes entre sí. ¿Cómo se afectará el medio ambiente? ¿Las facilidades existentes (escuelas, agua, etc.) serán capaces de soportar la carga adicional, o se tendrán que aumentar?. Las respuestas a estas preguntas requieren de la interrelación de una gran variedad de datos locales y ésta es otra ventaja del procesamiento geográfico - la de ser capaz de integrar datos a través de límites jurisdiccionales y funcionales.

La integración y correlación de la información a través de líneas funcionales, puede mejorar las siguientes áreas de política gubernamental y toma de decisiones:

- Asignación de recursos
- Ubicación de facilidades
- Uso del suelo
- Análisis de mercados
- Trabajos de ingeniería relacionados con los puntos anteriores.

Todas estas actividades son esenciales para la planeación de un mejor desarrollo. Todas las fases del desarrollo podrían ser cuidadosamente estudiadas por los planificadores, utilizando un sistema de procesamiento geográfico. Por tanto, se podrían trazar vastas redes de caminos, cañerías y líneas de energía para las necesidades óptimas presentes y futuras. Se podrían localizar eficientemente las estaciones de policía y bomberos, así como otras oficinas municipales. Se podrían localizar hidrantes para casos de incendio con objeto de dar la mejor protección posible a casas y oficinas. Se localizarían convenientemente escuelas y tiendas. Las líneas de transporte colectivo prodrían ser construidas donde más se necesitaran. Si se hiciera manualmente, cada proyecto representaría un enorme trabajo en términos de tiempo, dinero y recursos humanos. Sin embargo, un sistema de información urbano, podría correlacionar todos los datos necesarios para apoyar decisiones y políticas de importancia.

CAPITULO III.-

PRESENTACION GENERAL DEL SISTEMA

Un sistema de información urbana con geocodificación puede pensarse como una herramienta cuyo valor depende completamente de la habilidad y experiencia del usuario. El sistema capacita al usuario para agregar y desplegar rápidamente datos geográficos. Proporciona información para ayudar a los directivos y administradores a tomar decisiones efectivas. El sistema no toma las decisiones por sí mismo.

Los conceptos básicos de la geocodificación pueden ilustrarse mejor a través de un ejemplo: Evaluar el impacto de aplazar la construcción de un sistema de alcantarillado, en un área donde el crecimiento excede a los servicios; en otras palabras, las facilidades de alcantarillado - actuales no pueden satisfacer un crecimiento adicional. Un retraso - sería una de las posibles opciones para enfrentarse al problema, hasta que se pudieran construir facilidades adicionales que abarataran el costo total del proyecto.

La geografía del área se representa como una serie de mapas sobrepuestos.

1.- El primero (Figura 1) muestra la geografía física relacionada de un área hipotética de planeación; el río, los arroyos, y sus cuencas de drenaje asociadas. Las líneas punteadas representan los límites de las cuencas. El significado de las cuencas de drenaje es que cada una está servida por una sola red de alcantarillado, con una capacidad - dada.

2.- La segunda sobrepuesta (Figura 2) muestra las facilidades físicas o de infraestructura. En este ejemplo, se muestra la red de carreteras, Otras sobrepuestas podrían mostrar escuelas, tiendas, áreas recreativas, etc.

3.- La tercera sobrepuesta (Figura 3) muestra los límites póliticos del área. En este ejemplo, zonificación de distritos (comercial rodeada por residencias, rodeada por agricultura). Otras sobrepuestas podrían involucrar red de caminos, límites de parcelas, etc.

4.- La siguiente sobrepuesta (Figura 4) muestra los límites administrativos. En este ejemplo, se muestran las zonas escolares, pero también se podrían usar distritos de bomberos, distritos electorales, etc. En esta etapa, el mapa muestra la fisonomía física, facilidades de - infraestructura, límites políticos y límites administrativos. Con esta descripción básica de las relaciones físicas y geográficas, se podría hacer manualmente un análisis elemental.

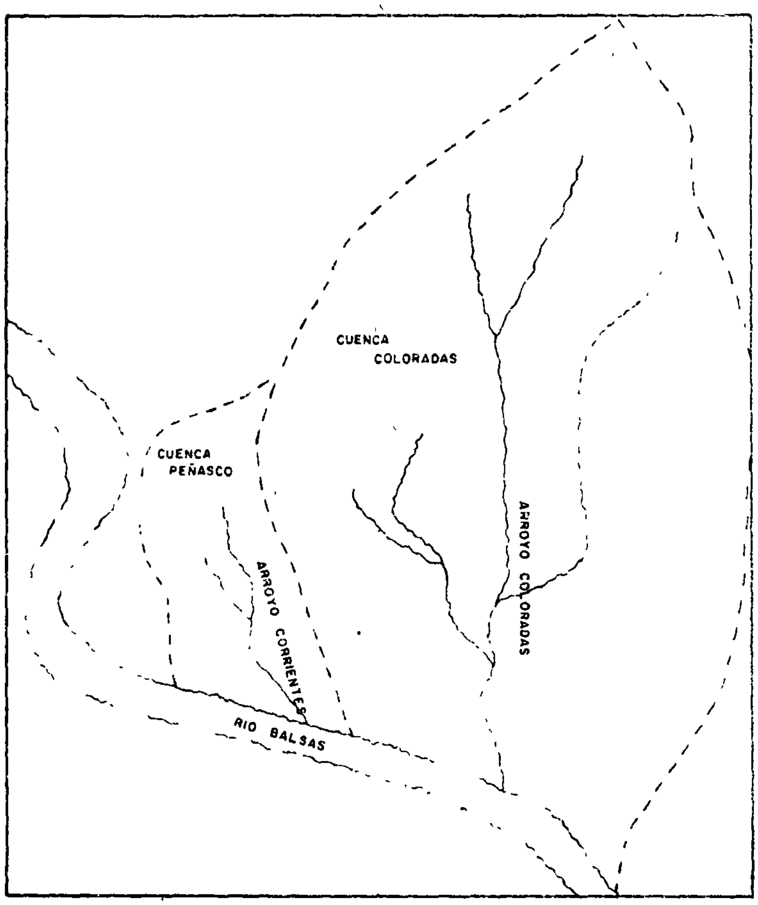


Figura 1. EJEMPLO DE UN MAPA TOPOGRAFICO SIMPLIFICADO

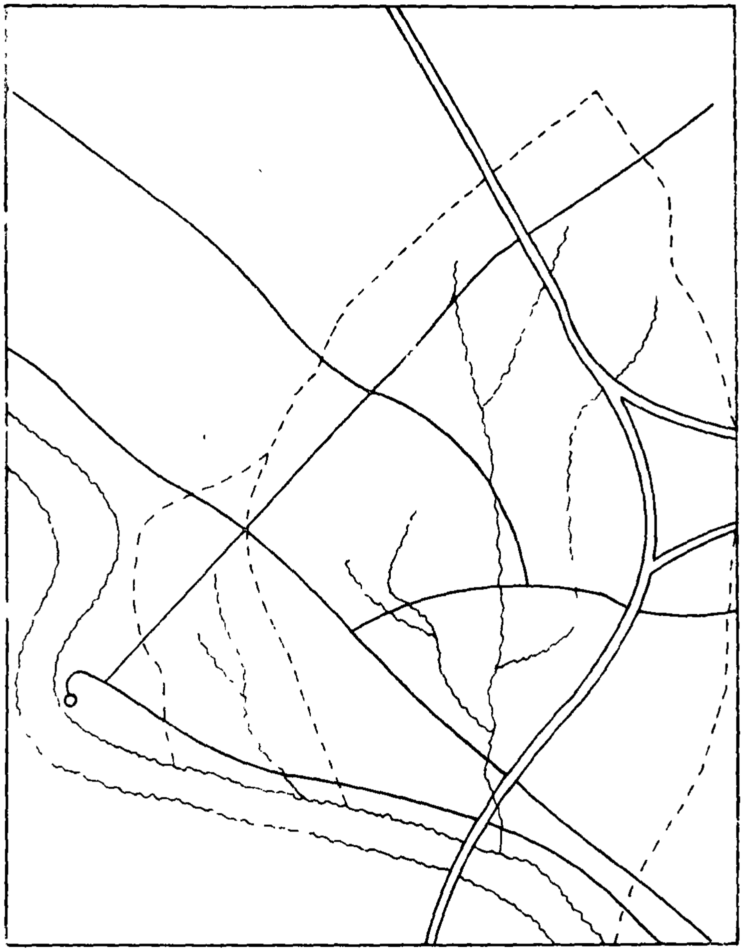


Figura 2 SOBREPUESTA DE OBRAS DE INFRAESTRUCTURA

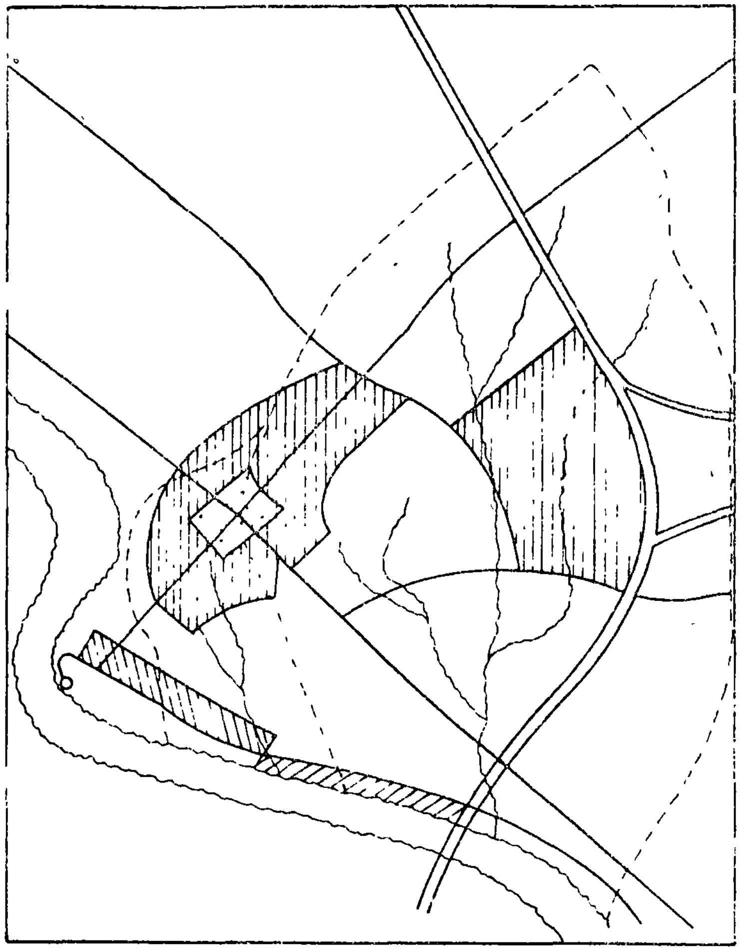


Figura 3 SOBREPUESTA DE LIMITES POLITICOS

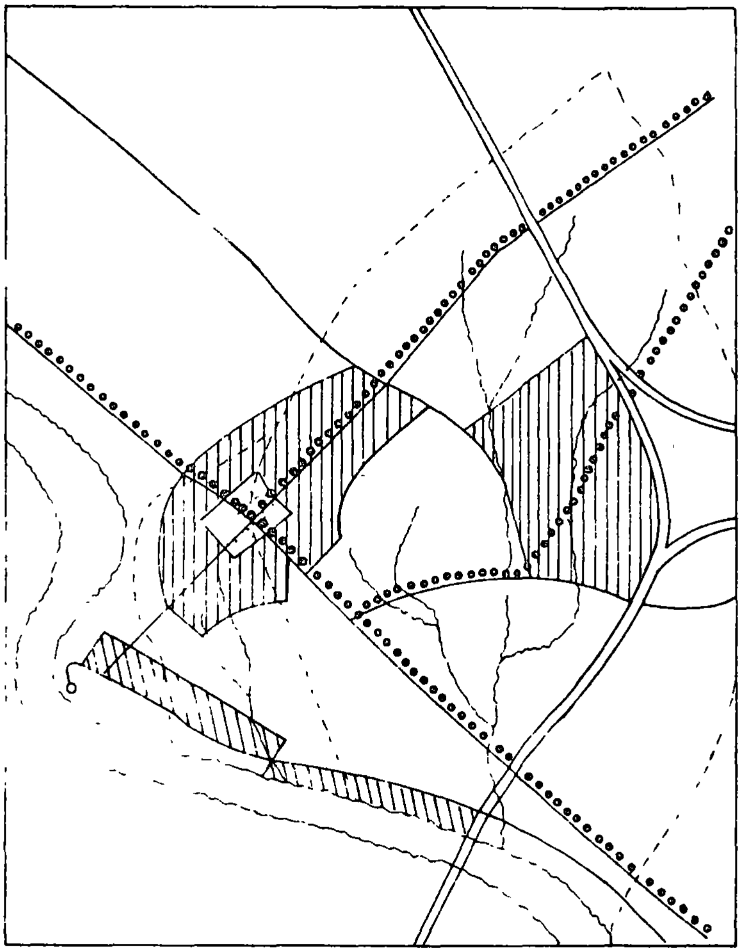


Figura 4 SOBREPUESTA DE LOS LIMITES ADMINISTRATIVOS

Llevando el problema un paso más adelante, esta geografía puede relacionarse con otros archivos de datos, como archivos de permisos de construcción, archivos de casos de incendio, policiacos, etc., para lograr un análisis más preciso de la situación. Para este problema en particular, se saca del archivo de predios, una lista de predios vacantes. Los predios vacantes se relacionan con las cuencas de drenaje y límites de zona y se hace un proyecto del potencial de construcción. El número y tipo de construcciones propuestas llevan a un proyecto preciso de la capacidad potencial de alcantarillado requerida para cada cuenca de drenaje. El impacto del retraso del alcantarillado se podrá calcular con precisión.

(Un análisis diferente podría ser el impacto sobre la inscripción escolar proyectada, utilizando límites de zonas escolares en lugar de los límites de cuencas de drenaje).

Manualmente esta aplicación, podría consumir varios cientos de horas-hombre. Muchas dependiencias hacen análisis de este tipo, únicamente para problemas de rutina, y no están relacionados con proyectos especiales. Este procedimiento completo, y muchos otros semejantes, pueden automatizarse utilizando un sistema de información geográfica urbana.

La figura 5 muestra un modelo simple de un sistema de procesamiento geográfico. Los mapas alimentados al sistema son los sobrepuestos ya mencionados. Las características físicas, obras de infraestructura, límites políticos y áreas administrativas. El procesador geográfico combina estas entradas para producir reportes o mapas que apoyen al proceso de toma de decisiones.

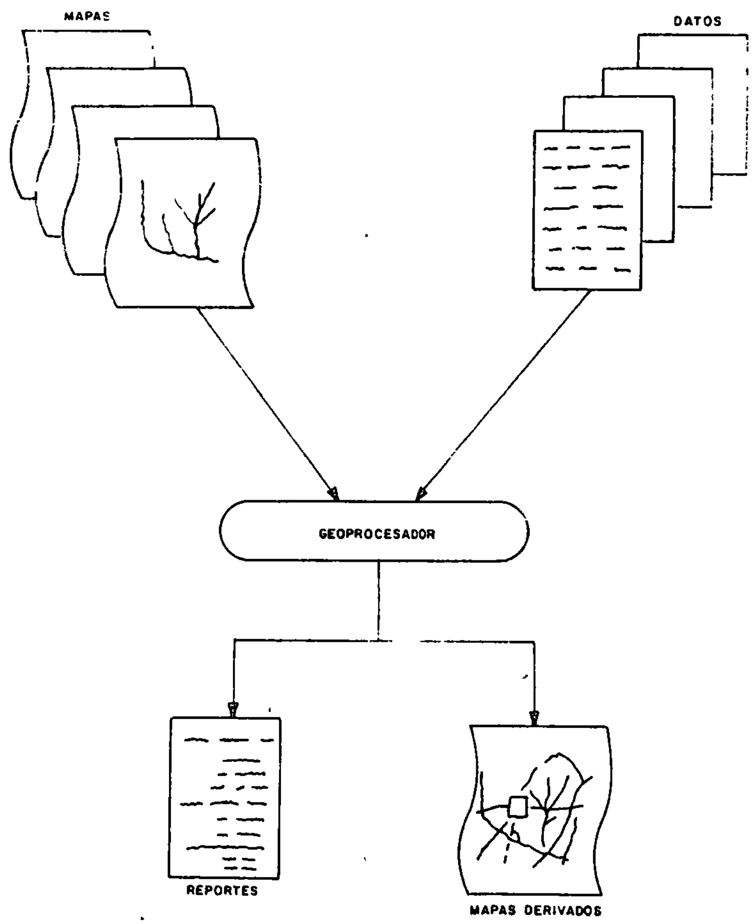


Figura 5. MODELO SIMPLE DE UN SISTEMA DE INFORMACION URBANA

CAPITULO IV.-

DESCRIPCION DEL SISTEMA

Un sistema de geoprocesamiento se puede dividir en tres fases distintas: Construcción/Mantenimiento, Selección y Proceso (ver figura 6).

CONSTRUCCION/MANTENIMIENTO.-

La construcción/mantenimiento, tiene la función de construir, editar y actualizar el archivo geográfico urbano(AGU). Este AGU es una descripción digitalizada de la geografía de un área específica, junto con sus índices asociados (incluyendo el número de calle y/o parcela). El AGU es un modelo de computación de un área geográfica, casi del mismo modo que un mapa es un modelo gráfico.

DIGITALIZAR un mapa es poner los datos en una forma legible de máquina, con lo cual se crea un archivo de procesamiento de datos del mapa. Esto se logra asignando coordenadas X, Y, a los puntos predeterminados en el mapa. Por ejemplo, se les pondrían coordenadas X, Y a las intersecciones de calles, así como a otros rasgos, como localizaciones de hidrantes para incendios, cementerios, puentes y carreteras. Entonces se le da un nombre a cada uno de estos puntos. Esta lista de nombres forma el índice de calles que se asocia a el AGU y se puede considerar como un "diccionario de datos". Cada nombre muestra la información - disponible para cada punto y dónde se puede encontrar esa información en la base de datos.

SELECCION, la Selección busca datos apropiados en el archivo geográfico urbano, y los correlaciona con los archivos de datos existentes (departamento de policía, consejo de planeación, asesoría, etc.) Por ejemplo, en el AGU se puede correlacionar un inventario de carreteras, un archivo de escuelas primarias, y una base de datos de parcelas, para analizar los sitios adecuados para construir una nueva escuela. La correlación se realiza haciendo coincidir los identificadores de localización que aparecen en el AGU y en las bases de datos operacionales.

PROCESO, el Proceso organiza los datos en el formato deseado y despliega el resultado. La produccción puede ser gráfica (mapas), alfanumérica (tabulación estadística o de terminal interactiva). Esta producción se puede hacer en impresoras, graficadores, terminales remotas, o terminales gráficas de vectores. El modelo se puede manejar en tiempo real.

Los mapas se pueden producir en una impresora de líneas, un graficador o un tubo de rayos catódicos (pantalla de video).

Para el público se pueden hacer mapas con graficadores de alta calidad, mientras que para otros mapas que no necesitan ser tan refinados se - puede usar impresoras electrostáticas.

El proceso puede contener un modelo para preguntas: de "qué pasaría si". Por ejemplo:

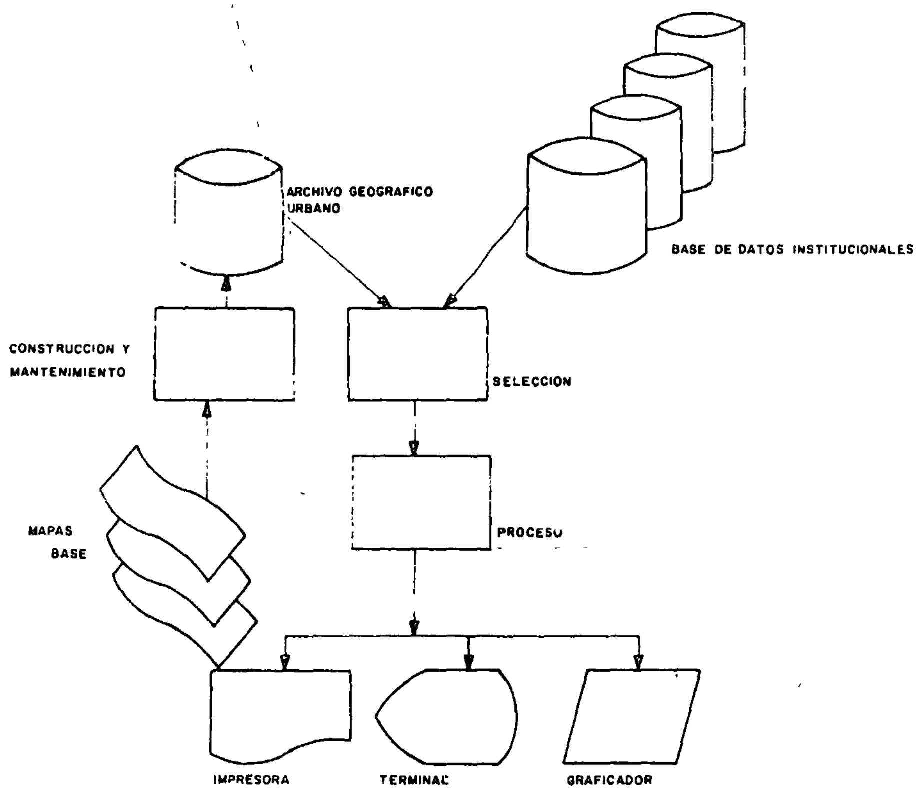


Figura 6 UN SISTEMA DE GEOPROCESAMIENTO

- El modelo de análisis de tráfico evalúa el nivel adicional de tráfico resultante de un nuevo conjunto comercial en el lugar. Se usa un modelo interactivo por medio de una terminal, donde se pueden fácilmente añadir, suprimir o cambiar datos. (¿Qué pasaría si se construyeran 13 tiendas en lugar de 8? ¿Cuál sería el impacto?).

- El modelo de contaminación del aire computa las concentraciones de contaminantes gaseosos (hidrocarbonos, monóxidos de carbono, etc.), que resultan de los equipos móviles y estacionarios que contaminan el ambiente. (¿Qué pasaría si hubiera 400 automóviles en lugar de 50? ¿Cuáles serían los niveles de contaminación?).

- El modelo de ordenamiento territorial mide el impacto de nuevos fraccionamientos sobre los servicios públicos, escuelas, facilidades recreativas, de salubridad y facilidades de proyección contra incendios, etc. (¿Qué pasaría si se constuyeran habitaciones para 15 familias adicionales? ¿Cuál sería el impacto?).

Estos tres modelos utilizan un archivo geográfico urbano que se correlacionan con las bases de datos para un sistema interactivo que posibilita a los administradores y planificadores resolver los problemas actuales y asegurar las necesidades futuras. Se puede añadir, cambiar o suprimir cualquiera de las partes de la información por medio de una terminal, y el programa se puede volver a correr para las determinaciones del último minuto.

CAPITULO V.-

MANEJO DE LA BASE DE DATOS

El Sistema de Información Geográfica Urbana correlaciona los datos de las bases institucionales, con el archivo geográfico urbano (figura 7).

Los archivos geográficos urbanos se pueden construir en cuatro niveles de complejidad; explícito, estadístico, de límites, y de ingeniería. A medida que se procede desde el básico hasta el más complejo, se pueden manejar una variedad mayor de aplicaciones; también se incrementa el costo y los requerimientos de mantenimiento del sistema.

El nivel de complejidad explícito, es el acceso básico al procesamiento geográfico (figura 8). En este nivel, cada registro de datos contiene el geocódigo. Por ejemplo, el distrito de bomberos está contenido en cada reporte de incidencia de incendios; o la cuenca de drenaje está contenida en cada registro de predio. Con este nivel, se puede hacer un procesamiento geográfico simple, sin el auxilio de un sistema de computación especial. Empero, existen los siguientes inconvenientes:

1.- No es posible cruzar la información de diferentes límites funcionales, a menos que todas las bases de datos estén geocodificadas.

2.- Es virtualmente imposible actualizar los archivos geográficos cuando se cambian los límites administrativos por una nueva - zonificación, división, etc.

3.- Las preguntas de "qué pasaría si", no pueden ser económicas porque las fronteras geográficas propuestas no están en el archivo. Por estas razones, el nivel de complejidad explícito no se utiliza tan ampliamente en los sistemas de procesamiento geográfico.

ENFOQUE MULTIDISCIPLINARIO

A las tres distintas fases de un sistema determinado de información urbana (Construcción/Mantenimiento, Selección y Proceso) se les puede denominar el enfoque multidisciplinario. Existen tres niveles de complejidad, (en realidad existen cuatro, pero no se incluye el nivel explícito por su simplicidad) por localización (direcciones y geocodificación simple), por límites (administrativos, políticos o debidos a la malla urbana) y el de facilidades, que defieren entre sí, principalmente en el nivel de detalles geográficos en los archivos geográficos urbanos.

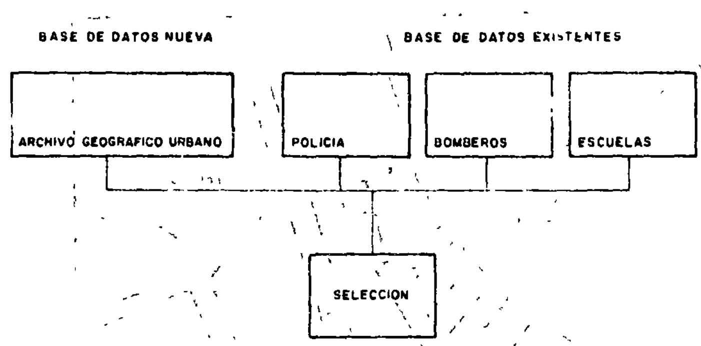


Figura 7 CORRELACION DE INFORMACION FUNCIONAL

NUMERO DE PREDIO	ZONA ESCOLAR	DELG POLICIA	DIST PLANEA	Z ALCANTA	P INUNDA	DATOS PREDIO
10 04 - 1 - 24 1	321	512	5	4	0	
10 05 - 1 - 24 2	321	512	5	4	1	
17 65 - 3 - 14 21	625	327	7	3	1	
17 65 - 3 - 14 22	626	32[illegible]	7	3	[illegible]	
—	—	—				
—	—					
—	—					

Figura 8 EJEMPLO DE UN METODO DE GEOPROCESAMIENTO

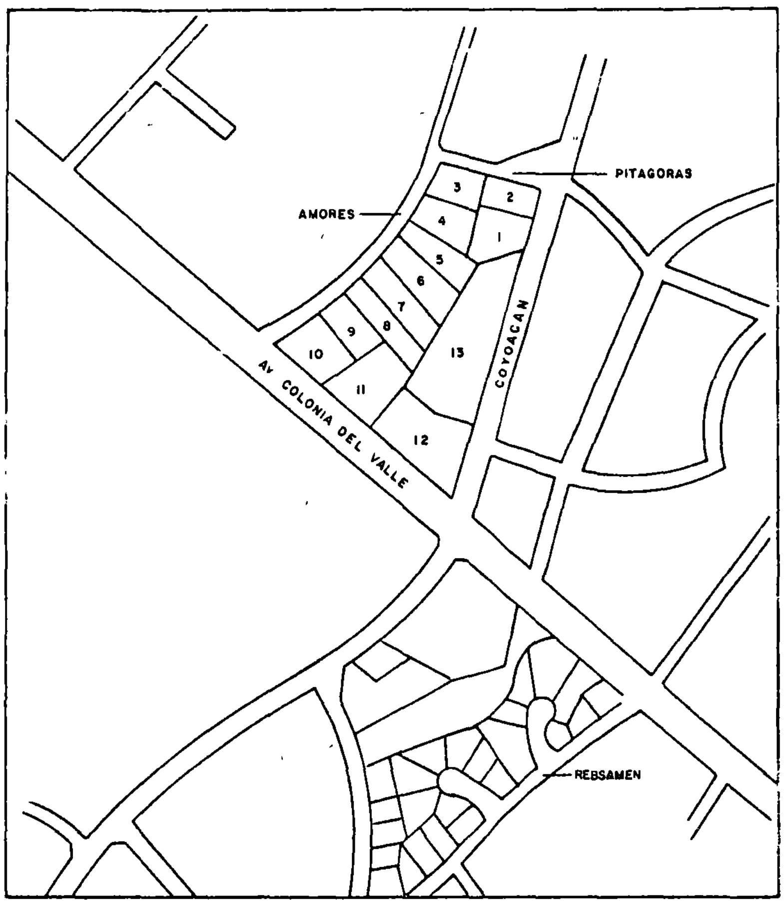


Figura 9. EJEMPLO DE UN MAPA PREDIAL

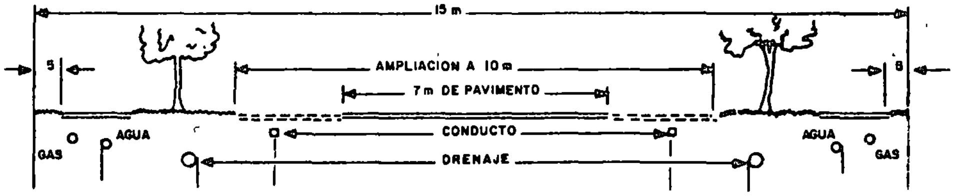


Figura 10. EJEMPLO DE UN DISEÑO DE INGENIERIA DETALLADA

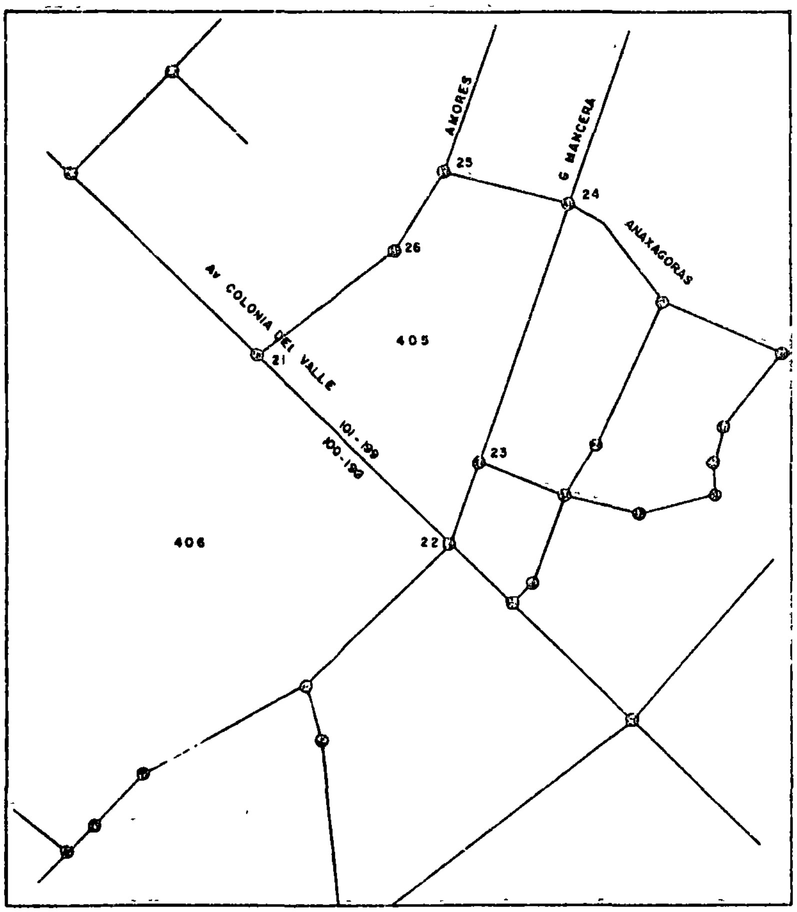


Figura 11 EJEMPLO DE UN MAPA AGU/CDMU

1.- El nivel de localización simple resuelve la ubicación de acontecimientos en áreas geográficas (por ejemplo, en el distrito de bomberos 14 ocurrió un incendio en Ave. Taxqueña 181, o en el área de planeación 5 ocurrió una subdivisión del predio 10-1-24). Responde a la cuestión básica: "¿En qué áreas geográficas ocurrieron los hechos?. Tiene la posibilidad de producir reportes estadísticos que sirven para el análisis o respuestas interactivas para una acción inmediata. Este nivel también puede apoyar la producción de mapas de incidencia o de densidad.

2.- El de límites resuelve la geometría de los acontecimientos, en otras palabras, las intersecciones actuales de los polígonos. (Por ejemplo, en la figura 9. ¿Qué porción de qué predio cruzará exactamente el trazo propuesto para una nueva calle o carretera?.) El nivel estadístico listaría solamente los predios afectados. El sistema de límites da una respuesta geométrica más que un sumario estadístico. Este nivel sería útil para el valuador catastral; por ejemplo, podría pedirse contribución especial a todos los predios a 50 metros de una nueva avenida?. Se podrían preparar planos catastrales, para aquellos predios que estén incluídos, aunque sea parcialmente.

3.- El nivel de facilidades (o Ingeniería) resuelve los detalles a nivel de ingeniería (por ejemplo, porporciona trazos de ingeniería sin requerir de un levantamiento preliminar). Ver figura 10.

Todos estos niveles se pueden usar en el Enfoque Multidisciplinario, y son compatibles en nivel ascendente. Cada nivel más alto permite una mejor visión, pero es más complejo, y por consiguiente más costoso en su construcción y mantenimiento. Generalmente es más práctico empezar con el nivel de direcciones (localización) y más tarde extenderse hacia arriba de acuerdo con las demandas que se presenten. El Enfoque Multidisciplinario apoya las aplicaciones en tiempo real, como los despachos de emergencia, así como las aplicaciones de planeación.

ARCHIVO GEOGRAFICO URBANO/CODIFICACION DUAL INDEPENDIENTE PARA MAPAS URBANOS.-

Ya son varios los países que han implementado un sistema de información para áreas urbanas cuyas características principales son:

Una retícula de canevá en proyección cónica de Lambert para el caso de Suecia, sirve de base a la incorporación de información en la base de datos (FRIS).

En Canadá la Unidad Básica de Información está constituida por elementos lineales representado un costo del bloque de manzana (G.R.D.S.R.) y en el caso de los países escandinavos (NIMS) E.U.A. (DIME) y Francia (R.G.U.) la Unidad Básica es el segmento vial limitado por dos nodos o cruceros.

La Oficina de Levantamiento Censal de E.U.A. ha creado un archivo geográfico urbano para las áreas metropolitanas más grandes del país. Estos archivos se utilizarán para el Censo de 1980 y algunos del Censo de 1970, ya están disponibles. Ver figura 11. Sin embargo, algo más importante es que estos archivos (AGU/CDIMU) tienen los datos básicos para apoyar un sistema de procesamiento geográfico en el nivel de localización. La existencia de un archivo AGU/CDIMU para un área en particular, puede reducir bastante el tiempo y gastos necesarios para construir un sistema de procesamiento geográfico.

Los usuarios de estos sistemas serán incontables y ayudarán a enriquecer las bases de información de datos urbanos multidisciplinarios que existan.

CAPITULO VI.-

DESARROLLO DEL SISTEMA

El desarrollo de un sistema de información urbana es una empresa compleja. Es absolutamente esencial un apoyo a nivel directivo. Ya que el procesamiento geográfico cruza los límites funcionales, debe existir un apoyo a alto nivel, para cuando las decisiones crucen estos límites.

Se debe organizar un grupo con representantes de las dependencias que vayan a usar el sistema. Estos representantes participarían en la planeación de todo el sistema, con objeto de asegurar la satisfacción expedita de las necesidades y requerimientos de todas ellas. Se debe consultar a los usuarios potenciales durante el diseño del sistema, con el fin de conocer si aceptarán el producto terminado.

Este grupo de representantes de cada área debe llevar a cabo un estudio de lo siguiente:

- Aplicaciones específicas deseadas para cada dependencia (incluyendo los beneficios potenciales).

- Bases de datos disponibles y las que se tendrán que construir para apoyar las aplicaciones de cada unidad.

- La investigación y la transferencia de códigos o técnicas de potencial aplicación al sistema.

- Los recursos disponibles tanto locales, estatales o federales.

Entonces se procede a la elaboración de un plan de implementación basado en las aplicaciones requeridas por cada área y los recursos disponibles. Se podrían añadir al sistema las capacidades y aplicaciones, de manera modular, estableciendo un programa de implementación.

CAPITULO VII.-

JUSTIFICACION ECONOMICA

El valor real de un sistema de información urbana, se debe medir en términos del aumento de productividad en todas sus aplicaciones.

Se puede analizar la productividad desde tres puntos de vista:

1.- Reducción de costos.- Eliminación de algún desembolso de capital en el futuro. (Por ejemplo, evitar la construcción de una nueva estación de bomberos, porque el equipo actual y los recursos humanos se aprovechen más eficientemente).

2.- Redistribución de costos, recursos humanos, o equipo.- A través de una asignación más eficiente.

3.- Valor agregado.- Resultado de las posibilidades adicionales que antes eran muy costosas o muy difíciles de realizar. La posibilidad de establecer una regionalización nueva cae en esta categoría; frecuentemente, los requerimientos federales y estatales demandan informes sobre el impacto en el medio ambiente para nuevos proyectos, y éstos pueden ser muy complejos. Las horas-hombre empleadas en la investigación manual, podría ser astronómicas.

APLICACIONES POTENCIALES DEL PROCESAMIENTO GEOGRAFICO

En toda administración municipal existen aplicaciones del procesamiento geográfico. Continuamente se hacen algunas aplicaciones (como despachos de emergencia de bomberos o policía). Otras, como la regionalización se hacen sólo ocasionalmente como una función de la planeación. Todavía otras, como las inspecciones y permisos de construcción, se hacen rutinariamente. A continuación se presenta una lista de las aplicaciones potenciales del sistema de información urbana para poblaciones de 100,000 habitantes en adelante. No pretende ser completa, únicamente indicativa.

1.- Despacho de policía.- Despacho de recursos policiacos basados en la asignación geográfica y la historia local.

2.- Despacho de bomberos.- Despacho de recursos de protección contra incendios, basado en la localización geográfica, condiciones locales, y la historia del caso.

3.- Zonificación escolar.- Planeación de la utilización óptima de las facilidades educacionales, basada en la geografía de la población escolar.

4.- Transporte escolar.- Planeación de las rutas de autobuses más eficientes, basada en la población escolar y la localización.

5.- Planeación urbana.- Análisis del impacto de las decisiones urbanas sobre todas las dependencias del gobierno local.

6.- Planeación de transporte.- Planeación del transporte colectivo más eficiente, modificando las rutas y destinos de los transporres alimentadores.

7.- Recreación.- Localización de las mejores áreas recreativas, basada en la localización geográfica de los sitios y usuarios.

8.- Análisis de delincuencia.- Asistencia en la delimitación de áreas de alta incidencia de delitos para optimizar la protección policiaca.

9.- Cuidado de parques públicos.- Planeación del cuidado de zonas - verdes dentro de la ciudad, basada en la localización de estas zonas.

10.- Colecta de basura.- Planeación de rutas, basada en la geografía del uso del suelo.

11.- Impacto en el medio ambiente.- Análisis del impacto de las urbanizaciones propuestas, sobre la ecología y las dependencias municipales.

12.- Listas de empadronamiento.- Producción de listas de electores por área geográfica.

13.- Localización de servicios públicos.- Planeación de reparaciones y nuevas construcciones de líneas eléctricas, telefónicas, de agua potable, alcantarillado, etc.

14.- Uso del Suelo.- Mantenimiento de archivos exactos de uso del suelo.

15.- Monitores del crecimiento.- Análisis contínuo de los cambios de localidad, basado en los análisis geográficos de licencias y permisos de construcción.

16.- Inspecciones de construcción.- Planeación de programas de inspección, basada en la distribución geográfica de permisos y licencias.

17.- Mapas base para Catastro Urbano.- Determinación de los impuestos prediales.

18.- Valuaciones especiales.- Asistencia en el trazo de áreas candidatas a impuestos especiales.

19.- Mapa base (Ingeniería).- Mantenimiento de mapas de ingeniería exactos para eliminar la necesidad de estudios preliminares.

No todas las aplicaciones se usarán en todas las organizaciones; de cualquier modo cada aplicación sirve para, por lo menos, dos dependencias. Cada aplicación tiene un cierto potencial de productividad. La suma total de la productividad justificará un sistema de información geográfica urbana como el que se plantea.

CAPITULO VIII.-

CONCLUSION

La implementación de un sistema de información geográfica urbana puede dar muy buenos resultados para los gobiernos estatales y locales. Como con la implementación de cualquier sistema computarizado sofisticado, se necesita un apoyo completo y una gran confianza por parte de los altos niveles gubernamentales. La geocodificación integra información a través de límites funcionales (dependencias) y jurisdiccionales (políticos); el procesamiento geográfico no puede justificarse en base a una sola dependencia. Por tanto, la construcción, mantenimiento y uso del sistema, requieren de la completa cooperación y compromiso de todos los funcionarios, además del coordinador del proyecto. Sólo entonces se pueden obtener los recursos necesarios para tener éxito, facilidades, dinero y personal. Se debe consultar al funcionario de nivel más alto durante todas las fases de la planeación e instalación de la base de datos del sistema geográfico urbano.

MANEJO DE SISTEMAS DE INFORMACION GEOGRAFICA EN PLANEACION

TEMA: PROGRAMAS DISPONIBLES EN EL GEOGRAPHY PROGRAM
EXCHANGE MICHIGAN STATE UNIVERSITY.

PROF. ARQ. ALEJANDRO VILLANUEVA EGAN.

NOV. 14-XI, 1977.

PROGRAM ORDERS
(revised June, 1977)

The Geography Program Exchange makes available all of its computer program holdings according to any of the following options for program orders:

(1) OPTION 1: Listing of Program Documentation and Listing of Program Source Code (Prices for option 1 orders reflect charges for paper stock, computer time, postage, and handling).

Items included in an option 1 order:

(a) Listing of Program Documentation
--Statement of purpose
--Input requirements and deck make-up
--Description of output
--References
--Program source

(b) Listing of Source Code
--Main program
--Internal functions and subroutines
--Test data set (if available)

(2) OPTION 2: Listing of Program Documentation, Listing of Program Source Code on Punched Cards (Prices for option 2 reflect charges for paper and card stock, computer time, postage, and handling).

Items included in a option 2 order:

(a) Listing of Program Documentation (includes the same features as those listed under item 1-a).

(b) Listing of Source Code (includes the same features as those listed under item 1-b).

(c) Program Source Code on Punched Cards.*
--Main program
--Internal functions and subroutines
--Test data set (if available)

(d) Sample Program Output--if sample output is available for a given program, an asterisk (*) will follow the option 2 price for that program on the price list.

(3) OPTION 3: Program Documentation and Program Source Code on Magnetic Tape (Prices for option 3 orders reflect the charges for computer time and handling.** Charges for paper stock and computer time are added if sample output is available).

(a) Program Documentation (including the same features as those listed under item 1-a) recorded in card-image form on magnetic tape.

(b) Program Source Code (including the same features as those listed under item 1-b) recorded in card-image form on magnetic tape.*

(c) Sample Program Output--If sample output is available for a given program, an asterisk (*) will follow the option 3 price for that program on the price list.

*All program source code is made available by the GPE in either BCD or EBCDIC coding. BCD coding contains the basic character set of the IBM-026 keypunch, while EBCDIC contains the expanded character set of the IBM-029 keypunch. All orders should specify which type of coding is desired. If none is specified, the order will be filled using BCD coding.

In addition, the reader will notice in the following "Index of Programs Currently Available" that the GPE usually maintains programs in either CDC or IBM FORTRAN. Several of the programs, however, are made available in both versions of FORTRAN. It is important, therefore, that all orders indicate which version of FORTRAN is desired for each program requested. In the case where the type of FORTRAN is not specified for a program available in either version, the CDC FORTRAN version will default.

**A base tape charge of $20.00 is incurred with each order of one or more programs recorded on magnetic tape (order Option 3). This price includes the cost of a new magnetic tape (2400 feet at $15.50), insurance and postage ($3.00), and handling ($1.50). All magnetic tapes are insured for $50.00 and mailed by parcel post.

The facilities of the Computer Laboratory at Michigan State University permit the GPE to create either seven-track or nine-track magnetic tapes at densities of 200, 556 (7 track only), 800, or 1600 (9 track only) bits per inch.

PAYMENT POLICY

When ordering program materials from the GPE, please send a check, money order, or purchase order for the total amount of the order. Checks and money orders should be made payable to MICHIGAN STATE UNIVERSITY and forwarded to:

Geography Program Exchange
Computer Institute for Social Science Research
510 Computer Center
Michigan State University
East Lansing, MI 48824

IMPORTANT: Payment must accompany an order totalling less than $20.00. The policy of prepayment is also in effect for all orders, regardless of the total dollar value of the order, originating outside the U.S.A. In addition, all foreign orders must be paid in either U.S. currency (example; bank draft on U.S. bank) or by international money order. All prices noted on the price list of programs currently available from the GPE are subject to change without further notice.

SHIPMENT POLICY

DOMESTIC ORDERS (U.S.A. and Canada): All printed matter will be forwarded to the customer's address through third class mail. All parcels (i.e., magnetic tapes and packages of punched cards) will be shipped via parcel post. If shipment by air express is desired, the GPE will provide a quotation for the air express charge upon special request.

INTERNATIONAL ORDERS: All parcels will be forwarded to the shipping address via international parcel post. If a less time-consuming mode of shipment is desired, the GPE will provide a quotation for that shipment charge (e.g., via air express) upon special request. In the case of international orders, any special shipment charge must be prepaid.

Program Prices and Option Availability

Acronym	IBM Option 1	IBM Option 2	IBM Option 3	CDC Option 1	CDC Option 2	CDC Option 3
ACCESS	N.A.	N.A.	N.A.	$12.40	$22.70	$11.10
ADJUST	N.A.	N.A.	N.A.	9.40	16.80*	10.00*
ALLOC	$ 8.80 #	$13.70 #	$ 7.30 #	N.A.	N.A.	N.A.
ALTERN	8.20 #	12.50 #	6.80 #	N.A.	N.A.	N.A.
AZMAP	8.50	12.30	6.80	8.80	15.30*	8.90*
BASICS	N.A.	N.A.	N.A.	8.70	12.80	6.80
BIVAR (PDP)	8.50	15.70*	9.00*	N.A.	N.A.	N.A.
BLCK	8.20	11.80	6.80	8.50	16.30*	9.30*
CANON	8.70	13.80	6.80	N.A.	N.A.	N.A.
CANTRN (UVAC)	9.50	15.40	8.00	N.A.	N.A.	N.A.
CART1	10.30	17.80	8.50	N.A.	N.A.	N.A.
CART2	9.60	15.20	8.00	N.A.	N.A.	N.A.
CENDA	8.10	11.00	6.80	N.A.	N.A.	N.A.
CENTRO	8.80	13.80	7.30	9.00	19.20*	11.10*
CHIINT	N.A.	N.A.	N.A.	7.80	10.60	6.80
CHOROS	9.80	15.30	8.00	N.A.	N.A.	N.A.
CLCOUNT	8.10	11.00	6.80	8.10	11.10	6.80
CLIMAT (GE)	5.50 #	9.60 #	5.50 #	N.A.	N.A.	N.A.
CLUSTER (UVAC)	8.20	11.40	6.80	N.A.	N.A.	N.A.
CLUSTR	N.A.	N.A.	N.A.	8.50	13.50*	7.60*
CMAP	8.00	10.70	6.80	N.A.	N.A.	N.A.
CNGRP	8.90	13.70	7.30	9.10	24.20*	14.60*
CNTOUR	9.10	13.30	7.30	N.A.	N.A.	N.A.
COBMAP	9.10	13.90	7.30	N.A.	N.A.	N.A.
COLMOG	N.A.	N.A.	N.A.	8.10	11.10	6.80
CONDIST	N.A.	N.A.	N.A.	8.70	12.90	7.30
CONGRP	8.90	14.10	7.30	9.00	13.90	7.30
CONRAT	N.A.	N.A.	N.A.	8.20	11.20	6.80
CONTOUR	N.A.	N.A.	N.A.	8.80	13.30	7.30
CONTR	8.30	12.00	6.80	N.A.	N.A.	N.A.
CONTUR	9.80	16.80	8.50	N.A.	N.A.	N.A.
CONWGT (UVAC)	8.90	13.70	7.30	N.A.	N.A.	N.A.
COORD	8.70	13.30	7.30	N.A.	N.A.	N.A.
CORD (UVAC)	8.70	13.30	7.30	N.A.	N.A.	N.A.
COSINE	8.20	11.10	6.80	N.A.	N.A.	N.A.
COVAR	9.50	16.00	8.00	9.50	19.10*	10.00*
CPLETH	8.70	12.30	6.80	N.A.	N.A.	N.A.
CURVES	N.A.	N.A.	N.A.	8.70	12.50	6.80
DENDRO (UVAC)	9.00	13.70	7.30	N.A.	N.A.	N.A.
DISAGG	8.10	11.00	6.80	8.10	11.00	6.80
DISGRP	9.40	13.50	7.30	N.A.	N.A.	N.A.
DISITR	8.80	13.50	7.30	N.A.	N.A.	N.A.
DISTORT	8.20	11.40	6.80	N.A.	N.A.	N.A.
DON	9.50	15.30	8.00	N.A.	N.A.	N.A.
DOUBLE	8.90	12.60	7.30	N.A.	N.A.	N.A.

Acronym	IBM Option 1	IBM Option 2	IBM Option 3	CDC Option 1	CDC Option 2	CDC Option 3
DSTAZ	$ 8.20	$11.80	$ 6.80	$ 8.20	$11.60	$ 6.80
DYAD	8.70	13.10	6.80	N.A.	N.A.	N.A.
ELIPS	8.20	11.60	6.80	N.A.	N.A.	N.A.
ENTROPY	8.10 #	14.70*#	8.70*#	N.A.	N.A.	N.A.
EQUAL	8.50	12.30	6.80	8.50	14.30	8.20
EUCLID	8.10	11.20	6.80	8.10	12.10*	7.40*
EXTRAP	8.10	11.20	6.80	8.10	10.00	6.80
FLTREC (UVAC)	10.60	13.70	9.20	N.A.	N.A.	N.A.
FORFIT (UVAC)	8.80	12.90	7.30	N.A.	N.A.	N.A.
GENEB	8.50	12.60	6.80	8.50	20.80*	12.60*
GEODIS	8.00	11.10	6.80	N.A.	N.A.	N.A.
GEOFIT	9.50	16.20	8.00	N.A.	N.A.	N.A.
GEOPAK	9.50#	16.00#	8.00#	N.A.	N.A.	N.A.
GIPSY	9.10	13.90	7.30	9.40	17.90	10.70
GRAVITY	N.A.	N.A.	N.A.	8.10	12.60 *	7.60*
GRAVTY (UVAC)	8.30	12.00	6.80	N.A.	N.A.	N.A.
GRID	8.30	12.10	6.80	8.30	13.30*	7.40*
GVAR	8.20	11.80	6.80	N.A.	N.A.	N.A.
HAAAG	N.A.	N.A.	N.A.	8.70	12.50	6.80
HAGMEVA	N.A.	N.A.	N.A.	7.80	10.60	6.80
HAGPLOT	N.A.	N.A.	N.A.	7.80	10.70	6.80
HENTRO	6.80	N.A.	N.A.	N.A.	N.A.	N.A.
HEXAGON	8.50	12.90	6.80	N.A.	N.A.	N.A.
HVAR	8.20	11.60	6.80	N.A.	N.A.	N.A.
INDIFF	8.90	13.50	7.30	N.A.	N.A.	N.A.
INPOUT (UVAC)	8.20	11.20	6.80	N.A.	N.A.	N.A.
INTCYL (UVAC)	9.60	15.80	8.00	N.A.	N.A.	N.A.
INTPMED	8.30 #	13.10 #	6.80 #	N.A.	N.A.	N.A.
INTRMAP	7.80	10.60	6.80	N.A.	N.A.	N.A.
INTRPOL	8.20	11.80	6.80	N.A.	N.A.	N.A.
INVERSE	8.70	13.50	6.80	N.A.	N.A.	N.A.
IOWAP	N.A.	N.A.	N.A.	8.00	10.70	6.80
IPCIDA	8.20	11.10	7.00	N.A.	N.A.	N.A.
ISODEN	N.A.	N.A.	N.A.	6.80	N.A.	N.A.
ISOMET	N.A.	N.A.	N.A.	13.50	24.50	12.30
ITERIM	N.A.	N.A.	N.A.	13.50	21.60	12.30
KCOLOR	N.A.	N.A.	N.A.	8.30	11.80	6.80
KOPPEN	8.80	12.60	7.30	N.A.	N.A.	N.A.
LANDUSE	10.10	16.00	8.50	10.00	16.00	8.50
LAP	9.10 #	16.20 #	7.30 #	N.A.	N.A.	N.A.
LATMAP	N.A.	N.A.	N.A.	9.60	23.10*	13.70*
LINEAR (UVAC)	8.20	11.40	6.80	N.A.	N.A.	N.A.
LISTER	N.A.	N.A.	N.A.	8.30	11.60	6.80
LOCATE	8.50	12.50	6.80	8.50	12.60	6.80
LOCPOT	9.40	18.70*	10.40*	N.A.	N.A.	N.A.
LUAM	3.20	N.A.	N.A.	N.A.	N.A.	N.A.

Acronym	IBM Option 1	IBM Option 2	IBM Option 3	CDC Option 1	CDC Option 2	CDC Option 3
MAPDATA	N.A.	N.A.	N.A.	$ 8.10	$11.00	$ 6.80
MAPIT	N.A.	N.A.	N.A.	10.30	17.50*	9.60*
MAP3D	N.A.	N.A.	N.A.	8.20	11.80	6.80
MAPLOT	$ 8.80	$12.60	$ 7.30	N.A.	N.A.	N.A.
MAPTRAN	8.80 #	14.60 #	7.30 #	N.A.	N.A.	N.A.
MARKOV1	8.70	13.30	6.80	8.50	12.60	6.80
MARKOV2	8.20	12.00	6.80	8.20	11.80	6.80
MAVE	8.00	11.00	6.80	N.A.	N.A.	N.A.
MDISC	9.50	15.30	8.50	N.A.	N.A.	N.A.
MINIMAP	8.80 #	12.10 #	7.30 #	N.A.	N.A.	N.A.
MINPATH	N.A.	N.A.	N.A.	8.30	12.00	6.80
MLNEGBI	N.A.	N.A.	N.A.	8.10	11.10	6.80
MOCON	N.A.	N.A.	N.A.	7.80	10.40	6.80
MOMENS	3.20	N.A.	N.A.	N.A.	N.A.	N.A.
MONA	8.30	11.80	6.80	N.A.	N.A.	N.A.
MPLSQ	N.A.	N.A.	N.A.	13.60	22.20	12.30
MULTI	9.40 #	16.80 #	8.00 #	N.A.	N.A.	N.A.
MUNKRE	N.A.	N.A.	N.A.	8.30	12.80*	7.10*
NAYBOR	8.10	11.20	6.80	N.A.	N.A.	N.A.
NEARNBR	8.00	11.00	6.80	N.A.	N.A.	N.A.
NEGBIN	N.A.	N.A.	N.A.	7.80	10.60	6.80
NODAC	N.A.	N.A.	N.A.	8.70	12.60	6.80
NOLIDA	8.70	12.50	6.80	N.A.	N.A.	N.A.
NONCEL	8.70	12.50	6.80	8.80	12.50	7.30
NONCELO	N.A.	N.A.	N.A.	9.70	25.30*	15.30*
NONPLOT	N.A.	N.A.	N.A.	7.80 #	11.20 #	6.80 #
NORLOC (UVAC)	8.80	12.90	7.30	N.A.	N.A.	N.A.
NORM	8.30	12.00	6.80	8.50	15.30*	8.50*
NSCAT	8.30	12.10	6.80	N.A.	N.A.	N.A.
OPTREG	15.60	21.70	14.10	N.A.	N.A.	N.A.
PCPA	11.30 #	23.60*#	9.80*#	N.A.	N.A.	N.A.
PELTO	8.10	11.40	6.80	7.80	11.40	6.80
PERS	9.00	13.10	7.30	N.A.	N.A.	N.A.
PLOTMAP	10.20	17.00	8.50	N.A.	N.A.	N.A.
PLOTMP (UVAC)	8.20	12.00	6.80	N.A.	N.A.	N.A.
PLOTS	8.00	12.30	7.40	N.A.	N.A.	N.A.
POINT	N.A.	N.A.	N.A.	8.10	12.30*	7.40*
POINTS	N.A.	N.A.	N.A.	8.70	12.00	6.80
POISSN	N.A.	N.A.	N.A.	7.80	10.40	6.80
POLYFIT	10.30	19.10	8.50	N.A.	N.A.	N.A.
POPMAP	8.30	10.60	6.80	N.A.	N.A.	N.A.
POPPYR	7.80	10.60	6.80	N.A.	N.A.	N.A.
PREDEN (UVAC)	8.90	13.80	7.30	N.A.	N.A.	N.A.
PRESLOC	8.10 #	12.00 #	6.80 #	N.A.	N.A.	N.A.
PROLO	N.A.	N.A.	N.A.	11.30	20.60	9.80
PYRAMID	8.10 #	12.00 #	6.80 #	N.A.	N.A.	N.A.
RANKD (UVAC)	8.00	10.70	6.80	N.A.	N.A.	N.A.

Acronym	IBM Option 1	IBM Option 2	IBM Option 3	CDC Option 1	CDC Option 2	CDC Option 3
REGLAR (UVAC)	$ 8.10	$11.00	$ 6.80	N.A.	N.A.	N.A.
REGRESS	N.A.	N.A.	N.A.	$ 8.30	$11.80	$ 6.80
REVPREF	10.30 #	20.10 #	8.50 #	11.70	23.60*	13.40*
RGRID	9.00	14.30	7.30	9.00	16.30*	9.00*
RGRID2	9.10	14.70	7.30	N.A.	N.A.	N.A.
RIDGERE	3.40	N.A.	N.A.	N.A.	N.A.	N.A.
RSTAR	8.80	12.60	7.30	N.A.	N.A.	N.A.
SAMPLE (UVAC)	8.20	12.60	6.80	N.A.	N.A.	N.A.
SANDM	N.A.	N.A.	N.A.	8.00 #	11.60 #	6.80 #
SCALE	N.A.	N.A.	N.A.	10.30	16.00	8.50
SDIS	8.00	11.00	6.80	N.A.	N.A.	N.A.
SINGLE	8.80	12.80	7.30	N.A.	N.A.	N.A.
SIRES	9.40	20.60	11.40	9.00	22.50*	12.70*
SLOPE	8.10	11.10	6.80	8.30	13.30*	7.60*
SMOOTH	N.A.	N.A.	N.A.	8.00	10.70	6.80
SMOTHR	7.80	10.60	6.80	N.A.	N.A.	N.A.
SMRATE (UVAC)	8.10	11.10	6.80	N.A.	N.A.	N.A.
SNORT (UVAC)	9.00	14.10	7.30	N.A.	N.A.	N.A.
SOLUP	8.80	12.90	7.30	8.90	14.10*	8.30*
SPA	9.40 #	16.40 #	8.00 #	N.A.	N.A.	N.A.
SPACE	3.50	N.A.	N.A.	N.A.	N.A.	N.A.
SPAN	N.A.	N.A.	N.A.	8.90	20.60*	12.20*
SPECTR	8.80	12.80	7.30	N.A.	N.A.	N.A.
SPHERE	N.A.	N.A.	N.A.	8.30	13.10	7.30
STDROP (HNYWL)	11.50	21.10	9.80	N.A.	N.A.	N.A.
SURGE	8.80	13.10	7.30	8.70	17.00*	9.80*
SYMBOLS	N.A.	N.A.	N.A.	8.20	11.10	6.80
TAXON (UVAC)	8.20	11.20	6.80	N.A.	N.A.	N.A.
THIESEN	8.70	12.50	6.80	N.A.	N.A.	N.A.
TORG	8.80	14.10	7.30	8.90	14.30	7.30
TORG2	8.80	14.10	7.30	N.A.	N.A.	N.A.
TORN	8.20 #	12.50 #	6.80 #	N.A.	N.A.	N.A.
TRANS	8.10	11.20	6.80	8.10	17.70*	11.40*
TREND	11.30	22.00	9.80	11.50	26.80*	11.10*
TRENDC	10.60	20.20	9.20	N.A.	N.A.	N.A.
TRID	N.A.	N.A.	N.A.	9.00	13.90	7.30
TWAIN	8.20 #	12.60 #	6.80 #	N.A.	N.A.	N.A.
USDATA (UVAC)	9.40	15.30	8.00	N.A.	N.A.	N.A.
VALRAT	N.A.	N.A.	N.A.	8.00	10.70	6.80
VALRATI	8.50	12.30	6.80	8.70	13.80*	7.60*
WARELOC	8.30 #	13.10 #	6.80 #	N.A.	N.A.	N.A.
WEBER (UVAC)	8.10	11.00	6.80	N.A.	N.A.	N.A.
WEBERI	8.10 #	12.00 #	6.80 #	N.A.	N.A.	N.A.
XCUT	N.A.	N.A.	N.A.	8.80	13.80	7.30
XMAPP	8.00	11.10	7.30	8.00	11.10	7.30
ZSCORE (UVAC)	8.00	11.00	6.80	N.A.	N.A.	N.A.

*	Indicates test output is included
#	Documentation not available from GPE
N.A.	Not Available
UVAC	UNIVAC
HNYWL	HONEYWELL

GEOGRAPHY PROGRAM EXCHANGE

The Geography Program Exchange (GPE) has been established to assist universities and other non-profit organizations with the interchange of computer software which relates to problems of a geographic nature. The GPE operates on a self-sustaining basis. Its services are made available to users anywhere in the world.

The GPE maintains a central file of specialized computer programs with associated documentation and test data sets, and makes copies of all or part of this material available to users at cost. The programs held by the GPE represent specialized materials developed for use in geographic research and teaching; no attempt is made to duplicate standard statistical programs. Distribution of available programs is in the form of listings, punched cards, or magnetic tapes.

Program holdings consist of items from the specialized program libraries of a large number of geography departments in North America as well as programs submitted by geographers from other countries. The GPE actively solicits programs from all sources. If the program submitted appears to be of enough general interest, it is added to those available for distribution. All software handled by the GPE is obtained and distributed on a non-royalty basis.

Although the GPE operates with limited resources, it does attempt to provide additional assistance to those departments and individuals who are just starting to make use of the digital computer as a research and teaching tool.

Overall management of the GPE is provided by a board consisting of Professors Duane F. Marble (Suny/Buffalo), Waldo R. Tobler (University of California-Santa Barbara), and Robert I. Wittick (Michigan State University). Dr. Wittick is in charge of operations and all inquiries relating to activities of the GPE should be directed to him at the following address:

Dr. Robert I. Wittick
Computer Institute for Social Science Research
Michigan State University
East Lansing, Michigan 48824
U.S.A.

Telephone: 517/353-2040

A European outlet for the GPE programs has been established in West Germany. For further information on the availability of programs from West Germany, write to:

Dr. Andre Kilchenmann
Geographisches Institut
Universitat Karlsruhe (TH)
75 Karlsruhe 1, Kaiserstrasse 12
Germany, BRD

INDEX OF PROGRAMS CURRENTLY AVAILABLE

June 1977

ACCESS: Tape Library Utility Program

A utility program which is used to create and edit a tape library of computer programs. Programs on tape may be listed, punched, stored on another tape, or compiled.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: December, 1971
Contributor: Department of Geography
Northwestern University

ADJUST: Matrix Adjustment

A program designed to perform iterative multiplicative adjustment of matrices to desired marginal totals. The program can also be used in the analysis of transaction flows.

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: July, 1972
Similar Programs: INDIFF
Contributor: Department of Geography
Michigan State University

ALLOC: Heuristic Solutions to Multi-Facility Location Problems on a graph

Program to compute optimal locations for facilities based on minimized aggregate distance using Marazana, and Tietz and Bart Algorithms.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: January, 1974
Contributor: Department of Geography
University of Iowa

ALTERN:	Multi-Source Location-Allocation Algorithm (Heuristic)
	This program alternately locates sources at the minimum point of their assigned sinks and allocates sinks to their closest sources, to approximate the optimal location for M sources to serve N sinks in a Euclidean plane.
Computer:	IBM 360/65
Compiler:	FORTRAN G
Date Received:	January, 1974
Similar Programs:	LAP, MATRAN, TORN
Contributor:	Department of Geography University of Iowa

AZMAP:	Azimuthal Map Transformation
	The program calculates and produces azimuthal map transformations on the pen-and-ink plotter, as a means of displaying geographic data.
Computer:	IBM 360/50
Compiler:	FORTRAN G
Date Received:	May, 1971
Contributor:	Department of Geography University of Rhode Island

Other Versions: (AZMAP)

Computer:	CDC 6500
Compiler:	6000 FORTRAN EXTENDED
Date Received	January, 1972
Contributor:	Computer Institute for Social Science Research Michigan State University

BASICS:	Calculation of Basic Statistical Series
	Program to calculate and plot (on pen-and-ink plotter) one, or all of the following, which are commonly used in basic analyses of statistical series: (1) a histogram, (2) a data array, and (3) a cumulative-frequency graph.
Computer:	CDC 6600
Compiler:	6000 FORTRAN EXTENDED
Date Received:	March, 1972
Contributor:	Department of Geography University of Minnesota

BIVAR: Bivariate Means and Medians

Computes bivariate means and approximates bivariate medians by iterative means to a specified accuracy or a prescribed number of iterations. Also computes geometric coordinates, and total travel to the means and medians.

Computer: PDP-10
Compiler: FORTRAN IV
Date Received: March, 1972
Contributor: Department of Geography
Western Michigan University

BLCK: Block Diagram Plotting

The program is used for rapid pen-and-ink plotting of isometric profiles from a rectangular matrix of data

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: August, 1970
Similar Programs: ISOMET; MAP 3D; PERS; SURGE; TRID
Contributor: Department of Geography
University of Michigan

Other Versions: (BLCK)

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: December, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

CANON: Canonical Correlation

The program calculates the interrelationships between two sets of measurements made on the same observations.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: May, 1970
Contributor: Department of Geography
University of Iowa

CANTRN: Canonical Trend-Surface Analysis

This program determines geographic patterns of several variates simultaneously, thereby producing a multivariate trend surface. The purpose is to determine linear combinations for both geographic coordinates and the variates, Z (J), so as to maximize correlation between the two linear expressions.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Similar Programs: DYAD, MPLSQ, POLYFIT, TREND, TRENDC
Contributor: Department of Geography
State University of New York, Albany

CART1: Irregular Cartograms

Computes cartograms from areas defined by irregular polygons.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: October, 1971
Similar Programs: CART2
Contributor: Department of Geography
University of Michigan

CART2: Regular Cartograms

Computes regular cartograms from areas defined by a geographical matrix

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: October, 1971
Similar Programs: CART1
Contributor: Department of Geography
University of Michigan

CENDA: Distance-to-Centroid Discriminant Analysis

Classifies subjects on the basis of distance to group centroids.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: November, 1972
Similar Programs: COSINE, DISITR, ITERIM, MDISC, NEARNBR, NOLIDA
Contributor: Department of Geography
University of Waterloo

CENTRO: Centrographic Measures

The program computes descriptive measures of spatial distributions from coordinate data. Weighted or non-weighted point sets are allowable. Output includes mean center, standard radius, coefficient of circularity, and angle of orientation of the distribution.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: May, 1970
Similar Programs: ELIPS
Contributor: Department of Geography
University of Iowa

Other Versions: (CENTRO)

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: August, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

CHIINT: Interactions for Chi-Square

To compute interactions for chi-square

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Contributor: Department of Geography
Northwestern University

CHOROS: Application of Linear Neighborhood Operators to Choropleath Maps

Program to apply linear neighborhood operators to choropleath maps by using the reaggregation method devised by Tobler.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: October, 1972
Contributor: Department of Geography
University of Michigan

CLCOUNT: Point Counting Within Rectangular Cells of a Rectangular Region

To count the number of points within rectangular cells of a rectangular region.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Contributor: Department of Geography
Northwestern University

Other Versions: (CLCOUNT)

Computer: IBM 360/91
Compiler: FORTRAN G
Date Received: December, 1971
Contributor: Department of Geography
University of California, Los Angeles

CLIMAT: Recognition and Classification of Climate Types

A set of CAI programs for analyzing, cataloguing, and classifying raw climate data using a logical, uniform, and open ended system which is easily capable of being plotted and mapped.

Computer: GE-635
Compiler: GE BASIC
Date Received: March, 1972
Contributor: Department of Geography
Dartmouth College

CLUSTER: Clustered Pattern Recognition

The program establishes the location and areal extent of point clusters in a spatial distribution.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Similar Programs: LINEAR, REGLAR
Contributor: Department of Geography
State University of New York, Albany

CLUSTR: Cluster Analysis

Provides a means of grouping variables on a set of observations using cluster analysis. The clustering of variables begins with the pair of variables having the highest correlation and then proceeds by adding the variables having the highest correlation with the elements of the cluster.

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: January, 1972
Similar Programs: DON
Contributor: Computer Institute for Social Science Research
Michigan State University

CMAP: Choropleth Mapping

Designed to produce choropleth maps utilizing small computers.

Computer: IBM 360/40
Compiler: FORTRAN E
Date Received: December, 1971
Similar Programs: CPLETH, XMAPP, COBMAP, INTRMAP
Contributor: Department of Geography
University of Idaho

COBMAP: Choropleth Mapping in COBOL

This program is a COBOL version of CMAP. The program will produce one map for each submission.

Computer: IBM 360/65
Compiler: ANS COBOL
Date Received: March 1976
Similar Programs: CMAP, CPLETH, XMAPP, INTRMAP
Contributor: Department of Geography
University of Tennessee

CNGRP: Contiguity Grouping

The program groups observations using the criterion of euclidean proximity in a p-dimensional vector space if the data are given in the form of geographical matrices, the program can automatically impose a geographical contiguity constraint on the grouping.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: August, 1970
Similar Programs: CONGRP, DISGRP
Contributor: Department of Geography
University of Michigan

Other Versions: (CNGRP)

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: August, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

CNTOUR: Contouring a Grid

A package of FORTRAN subroutines to produce a contour map (pen-and-ink plot) of a set of data points which form a grid. The grid does not necessarily have to be regular.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: January, 1972
Similar Programs: CONTR, CONTUR
Contributor: Computing Centre
University of British Columbia

COLMOG: D-Statistic for Kolmogorov-Smirnov Test

The program computes the d-statistic for the Kolmogorov-Smirnov Test for equivalent rows or columns of the two matrices output from CONDIST.

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: June, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

CONDIST: Relationship of a Population to Distributed Facilities

The program computes the spatial relationship between a sample of points and a distribution of different sized facilities. The program computes the number of points within specified sizes. Also computed are the distances separating successive percentile groups of the points from their nearest facility larger than a specified size.

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: June, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

CONGRP: Contiguity Grouping

Stepwise multivariate grouping of observations using one of several distance criteria. Contiguity restraints may be imposed upon the groupings.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Similar Programs: CNGRP, DISGRP
Contributor: Department of Geography
Northwestern University

Other Versions: (CONGRP)

Computer: IBM 360/65
Compiler: FORTRAN G
OS/360 Assembler-F
Date Received: May, 1970
Contributor: Department of Geography
University of Iowa

CONRAT: Contiguity Ratios

Computation of statistics for contiguity measures to provide an evaluation of the hypothesis of randomness in the areal arrangement of the values of a specified property. The program is used for a two-category nominal, or binary, scaling of values.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Similar Programs: KCOLOR, VALRAT, VALRAT1
Contributor: Department of Geography
Northwestern University

CONTOUR: Contour Mapping

Construction of a contour map on the line printer from a series of data points whose locations are defined in terms of grid squares.

Computer: CDC 6400
Compiler: 6000 RUN-FORTRAN
Date Received: June, 1970
Similar Programs: LATMAP, RGRID
Contributor: Department of Geography
Northwestern University

CONTR: Automatic Contour Mapping

The program produces contour maps either on a pen-and-ink plotter, or on a cathode ray tube. The program will accept either a rectangular or an irregular shaped (not spaced) data set within a rectangular area.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: May, 1970
Similar Programs: CNTOUR, CONTUR
Contributor: Department of Geography
University of Iowa

CONTUR: Contour Mapping

The program uses a pen-and-ink plotter to draw contour maps from data given in the form of geographical matrices. Stereograms and perspective contours can also be obtained. An option allows conversion of the contours to a map projection before plotting.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: August, 1970
Similar Programs: CNTOUR, CONTR
Contributor: Department of Geography
University of Michigan

CONWGT: Contiguity-Structured Class Limits

The program weighs data values at each enumeration area by considering the spatial arrangement and similarity of the data units.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Contributor: Department of Geography
State University of New York, Albany

COORD: Coordinate Conversion

Program to convert section, township, and range notation to Cartesian coordinates.

Computer: IBM 360/91
Compiler: FORTRAN G
Date Received: December, 1971
Contributor: Department of Geography
University of California, Los Angeles

CORD: Generation of Coordinate Information for Mapping Routines

Coordinate systems are made readily available for use in the SYMAP, GIPSY, and Trend Surface mapping routines.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Contributor: Department of Geography
State University of New York, Albany

COSINE: Correlation Discriminant Analysis

Discriminant analysis based on angular proximity to group centroids.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: November, 1972
Similar Programs: CENDS, DISITR, ITERIM, MDISC, NEARNBR, NOLIDA
Contributor: Department of Geography
University of Waterloo

COVAR: Covariance Analysis

Performs a covariance analysis on a set of multiple regressions. An unequal N size is allowed for the regressions. Includes tests for homoscedasticity, equality of regressions, and equality of origins.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: May, 1970
Contributor: Department of Geography
University of Iowa

Other Versions: (COVAR)

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: August, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

CPLETH: Choropleth Mapping

Produces maps on the computer line printer using varying shades of darkness to indicate the intensity of phenomena in different geographical areas.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: October, 1971
Similar Programs: CMAP, XMAPP, COBMAP, INTRMAP
Contributor: Department of Geography
University of Michigan

CURVES: Plotting Series of Superimposed Curves

Program for plotting a series of up to four superimposed curves. The pen-and-ink plot is intended to permit visual analysis of the covariation or progression of several different statistical series.

Computer: CDC 6600
Compiler: 6000 FORTRAN EXTENDED
Date Received: March, 1972
Contributor: Department of Geography
University of Minnesota

DENDRO: Dendrograph Plotter Program

The program produces a two dimensional plotter diagram for depiction of the mutual relationships among a group of objects whose pair-wise similarities are given. Displays both within-group and between-group similarities measured as correlation coefficients or distance functions.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Contributor: Department of Geography
State University of New York, Albany

DISAGG: Geographical Matrix Disaggregation

To prepare data from pairs of geographical matrices for input to a multiple regression program.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: August, 1970
Contributor: Department of Geography
University of Michigan

Other Versions: (DISAGG)

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: December, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

DISGRP: Distance Grouping

Performs a cluster analysis based on Ward's technique of minimum group distance (within group sum of squares) increase at each step.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: November, 1972
Similar Programs: CNGRP, CONGRP
Contributor: Department of Geography
University of Waterloo

DISITR: Discriminant Iterations

Performs an iterative multiple discriminant analysis using orthogonal data.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: May, 1970
Similar Programs: CENDA, COSINE, ITERIM, MDISC, NEARNBR, NOLIDA
Contributor: Department of Geography
University of Iowa

DISTORT: Distortion of Geographical Maps

Calculates the distortion on geographical maps via the theorem of Tissot. The program is for distortion on projections given as a five degree table and uses a finite difference approximation to the derivatives (The program can read punched output from CART2).

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: October, 1971
Contributor: Department of Geography
University of Michigan

DON: Optimal Clustering Routine

Partitions objects into optimally homogeneous groups on the basis of empirical measures of similarity among those objects. Two methods are used to form clusters.

Computer: IBM 360/91
Compiler: FORTRAN G
Date Received: December, 1971
Similar Programs: CLUSTR
Contributor: Department of Geography
University of California, Los Angeles

DOUBLE: Double Fourier Series

Computes a double fourier series for irregularly spaced data. The series has been used as an alternative model to the polynomial for trend surface analysis.

Computer: IBM 360/91
Compiler: FORTRAN G
Date Received: December, 1971
Similar Programs: SINGLE
Contributor: Department of Geography
University of California, Los Angeles

DSTAZ: Computation of Distances on a Sphere

Rapid computation of the great circle distance and azimuth between any two points on either the lunar or terrestrial sphere.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Similar Programs: GEODIS, SDIS, SPHERE
Contributor: Department of Geography
Northwestern University

Other Versions: (DSTAZ)

Computer IBM 360/91
Compiler FORTRAN G
Date Received December, 1971
Contributor Department of Geography
University of California, Los Angeles

DYAD: Trend Surfaces by Eigenvector Dyads

The program produces a series of trend surfaces each consecutive member of which better approximates an original input distribution.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: August, 1970
Similar Programs: MPLSQ, POLYFIT, TREND, TRENDC
Contributor: Department of Geography
University of Michigan

ELIPS: Plotting of Bivariate Standard Deviations

The program may be used to provide some summary measures of geographical distributions. Includes pen-and-ink plot output.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: August, 1970
Similar Programs: CENTRO
Contributor: Department of Geography
University of Michigan

ENTROPY: Entropy Maximizing Techniques

A set of five programs written as subroutines in FORTRAN corresponding to a family of gravity models using entropy maximizing techniques. These spatial interaction models correspond in order to: (1) unconstrained, (2) origin constrained, and (3) total costs constrained flows.

Computer: not stated
Compiler: FORTRAN
Date Received: July, 1972
Contributor: Centre for Land Use and Built Form Studies
University of Cambridge

EQUAL: Lines of Equilibrium

The program uses the pen-and-ink plotter to depict lines of equilibrium about multi-centers of attraction. Such centers could include cities, individual firms, public institutions, etc.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: May, 1970
Similar Programs: GENEB, NAYBOR
Contributor: Department of Geography
University of Iowa

Other Versions: (EQUAL)

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: December, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

EUCLID: Geographical Distributions Correspondence In Euclidean Space

Computes the spatial correlation and linear regression between two associated distributions of N points each. The dependent set is in (U,V) space and the independent set is in (X,Y) space.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: October, 1971
Contributor: Department of Geography
University of Michigan

Other Versions: (EUCLID)

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: January, 1972
Contributor: Computer Institute for Social Science Research
Michigan State University

EXTRAP: Univariate Geographical Forecasting

Extrapolation of geographical matrices in the time domain using a positionally invariant, time varying, linear, local operator.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: August, 1970
Contributor: Department of Geography
University of Michigan

Other Versions: (EXTRAP)

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: December, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

FLTREC: Flow-Linkage Trend Recognition

Points or areas are linked to similar neighbors within a search region and these linkages are used to indicate similarities or dependencies. Resulting planar graphs may be useful in indicating the spatial pattern of a process.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Contributor: Department of Geography
State University of New York, Albany

FORFIT: A double Fourier Surface-Fitting Program for Irregularly Spaced Data Points

The double fourier series of trigonometric functions can be used as an alternative model to power-series polynomial trend-surface analysis. It provides a means for interpolation and limited extrapolation of data that are suspected to be periodic.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Similar Programs: DOUBLE, SINGLE
Contributor: Department of Geography
State University of New York, Albany

GENEB: Geographical Neighbors

Given NC points identified by rectangular coordinates X(I), Y(I); I = 1,...NC, the program produces the adjacency matrix of neighbors of order NN less than or equal to 9. These can then be used to find the neighbors to points of a regular lattice.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: August, 1970
Similar Programs: EQUAL, NAYBOR
Contributor: Department of Geography
University of Michigan

Other Versions: (GENEB)

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: August, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

GEODIS: Distances from Coordinates

The program reads a vector of geographical coordinates, and then punches the scalar product matrix of all pairs of great circle distances, in degrees.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: March, 1972
Similar Programs: DSTAZ, SDIS, SPHERE
Contributor: Department of Geography
University of Michigan

GEOFIT: Determination of Geographical Origins

The program estimates sets of source coordinates from empirical geographical distributions.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: August, 1970
Contributor: Department of Geography
University of Michigan

GEOPAK: A package of FORTRAN IV programs for the spatial filtering of digital geological maps.

Computer: IBM 360
Compiler: FORTRAN G
Date Received: June, 1974
Similar Programs: SMOOTH, MAVE
Contributor: Geosystems division of Lea Associates Limited

GIPSY: Geographical Incremental Plotting System

A computer mapping program for producing point and line symbols on the digital increment plotter.

Computer: IBM 360/50
Compiler: FORTRAN G
Date Received: June, 1971
Similar Programs: MAPIT, POPMAP, SYMBOLS
Contributor: Department of Geography
University of Rhode Island

Other Versions: (GIPSY)

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: July, 1972
Contributor: Computer Institute for Social Science Research
Michigan State University

GRAVITY: Iterative Fitting of Gravity Model

Used to fit the social gravity model to data describing spatial interaction (either inflow or outflow) between a set of areas and a user selected focal area.

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: July, 1972
Contributor: Department of Geography
Michigan State University

GRAVTY: Gravity Model Fitting Program

Program involves a data matrix that measures the inflows or outflows from one point to other points (observations). For each observation, distance, mass and index of interaction are calculated. Additional information is provided by subjecting a data matrix to simple regression using index of interaction as dependent variable and mass and distance as independent variables.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Similar Programs: GRAVITY
Contributor: Department of Geography
State University of New York, Albany

GRID: Geographical Interpolation

Interpolation to a square lattice from measures given at scattered geographical (X,Y) positions.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: August, 1970
Similar Programs: INTRPOL
Contributor: Department of Geography
University of Michigan

Other Versions: (GRID)

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: August, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

GVAR: Scale-Variance Detector with Even Hierarchy

The program uses a modified fixed effects analysis of variance procedure to detect variation at different nested scale levels from rectangularly gridded data.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: March, 1972
Similar Programs: HVAR
Contributor: Department of Geography
University of Michigan

HAAAG: Diffusion on a Regular Lattice

This program is based on the spatial diffusion model of Hagerstrand and simulates the location and number of knowers after each generation for each simulation.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Similar Programs: NONCEL, NONCELO
Contributor: Department of Geography
Northwestern University

HAGMEVA: Spatial Series Summary of HAAAG Output

To provide summary information on the spatial series generated by the tape output version of HAAAG.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Similar Programs: SANDM
Contributor: Department of Geography
Northwestern University

HAGPLOT: Temporal Series Summary of HAAAG Output

To provide summary information on the temporal series generated by the tape output version of HAAAG. Includes pen-and-ink plot output.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Similar Programs: NONPLOT
Contributor: Department of Geography
Northwestern University

HENTRO. Information Theory Entropy Measures

The program computes measures of information theory entropy for bivariate distributions. The following measures are calculated: information content, row entropy, column entropy, expected joint information, information gain, maximum entropies, and certain measures of redundancy.

Computer. IBM 370/135
Compiler: DOS Full FORTRAN
Date Received: October, 1974
Contributor: Department of Geography
University of Ibadan, Nigeria

HEXAGON: Hexagon Plotting and Transformation

Program to draw (using the pen-and-ink plotter) and transform hexagons.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: March, 1972
Contributor: Department of Geography
University of Michigan

HVAR: Scale-Variance Detector with Uneven Hierarchy

The program uses a modified fixed effects analysis of variance to detect scale variation in an uneven fully nested hierarchy.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: March, 1972
Similar Programs: GVAR
Contributor: Department of Geography
University of Michigan

INDIFF: Transaction Flows Analysis

A generalized program for the analysis of transaction flows.

Computer: IBM 360/91
Compiler: FORTRAN G
Date Received: December, 1971
Similar Programs: ADJUST
Contributor: Department of Geography
University of California, Los Angeles

INPOUT: Simplified Input-Output Model for Interindustry and Interregional Transaction Flows

For each run, the program outputs final demand and production vectors for a two or three region interregional flow model based on user specified input-output coefficients and final demands for each sector.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Contributor: Department of Geography
State University of New York, Albany

INTCYL: Interrupted Cylindrical Map Projection

Program to print out the earth's graticule according to the cylindrical projection specified.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Similar Programs: LATMAP, PLOTMAP
Contributor: Department of Geography
State University of New York, Albany

INTPMED: Evaluation of Student Solutions to Location Problems on a Network

The program uses the spatial assignment section of program ALLOC to partition the demand points among the centers chosen by the students. The optimum solution is instructor input and the student's solution is compared to it, giving a ranking of the coefficient of efficiency and a relative frequency distribution.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: January, 1974
Contributor: Department of Geography
University of Iowa

INTRMAP: Interactive Choropleth Mapping

This program is an interactive version of CMAP. It is designed to be used on standard character cathode ray tubes with 78 character lines or less.

Computer: IBM 360
Compiler: FORTRAN G
Date Received: November, 1975
Similar Programs: CMAP, COBMAP, CPLETH, XMAPP
Contributor: Department of Geography
Virginia Commonwealth University

INTRPOL: Grid Interpolation

Irregularly spaced data are rendered regular by local interpolation onto a cartesian grid.

Computer:	IBM 360/65
Compiler:	FORTRAN G
Date Received:	May, 1970
Similar Programs:	GRID
Contributor:	Department of Geography University of Iowa

INVERSE: Inverse Bivariate Interpolation

A program to perform inverse bivariate interpolation on a given set of tables of latitude and longitude coordinates for a map projection. A drawing of the resultant inverse tables is made via a pen-and-ink plotter.

Computer:	IBM 360/67
Compiler:	FORTRAN G
Date Received:	October, 1971
Contributor:	Department of Geography University of Michigan

IOWAP:

The program graphs the cumulative percentile distances on a pen-and-ink plotter using as input the matrix described in output on CONDIST.

Computer:	CDC 3600
Compiler:	3600 FORTRAN
Date Received:	June, 1971
Contributor:	Computer Institute for Social Science Research Michigan State University

IPCIDA: Summarized Areal Classifications

This program is designed to summarize areal classifications based on point samples. It is an aid to characterizing polygon-shaped areas, defined in a larger realm, on the basis of the number of point-occurences of some discrete phenomenon found to be contained in each of these areas. The total weighted values of the data points found in each area are summed and output.

Computer: IBM 370
Compiler: FORTRAN G
Date Received: May 1976
Contributor: Department of Geography
University of Denver

ISODEN: Digital Isodensitometry

The program uses the isodensitracer technique for the automatic mapping of isopleths.

Computer: CDC 6000 series (model not specified)
Compiler: FORTRAN
Date Received: January, 1976
Contributor: Department of Geography
Tel-Aviv University

ISOMET: Isometric Diagram Plotting

The program acts as a driver program that can access any one or all of the (13D) three dimensional isometric drafting subroutines. These subroutines provide an efficient and fast method of displaying a surface--a single valued function or two variables--in isometric or similar pictorial representation.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: December, 1971
Similar Programs: BLCK, MAP3D, PERS, SURGE, TRID
Contributor: Department of Geography
Northwestern University

ITERIM: Iterative Improvements Program for Evaluation and Improvement of Classifications

The program is designed primarily to assess and improve classifications, although it can be used also for principal components analysis, discriminant analysis, and one-way multivariate analysis of variance. Three criteria, pooled within groups sums of squares, Wilks' Lambas, and the sum of the eigenvalues associated with discriminant functions, are computed to assess and compare classifications.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: December, 1971
Similar Programs: CENDA, COSINE, DISITR, MDISC, NEARNBR, NOLIDA
Contributor: Department of Geography
Northwestern University

KCOLOR: Contiguity Measures for K-Color Maps

Computation of statistics for contiguity measures to provide an evaluation of the hypothesis of randomness in the areal arrangement of the values of a specified property.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Similar Programs: CONRAT, VALRAT, VALRATI
Contributor: Department of Geography
Northwestern University

KOPPEN: Koppen Climatic Classification

Classified climatic stations into major climatic types according to standard Koppen criteria.

Computer: IBM 360/91
Compiler: FORTRAN G
Date Received: December, 1971
Similar Programs: CLIMAT
Contributor: Department of Geography
University of California, Los Angeles

LANDUSE: Land Use and Market Area Model

Development of land use and market area patterns according to the basic Thunen model.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: December, 1971
Similar Programs: SOLUP, SIRES
Contributor: Department of Geography
Northwestern University

Other Versions: (LANDUSE)

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: October, 1972
Contributor: Department Of Geography
University of Michigan

LAP: Location Allocation Package

The program solves constrained and unconstrained continuous space location-allocation problems by variations of the alternation heuristic (algorithm). Transportation and Weber problems can be solved by requesting zero iterations

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: January, 1974
Similar Programs: ALTERN, MAPTRANS, TORN, WEBER, WEBERI
Contributor: Department of Geography
University of Iowa

LATMAP: Mapping Latitude/Longitude Coordinates on Polyconic Projection

Converts latitude/longitude points to Cartesian coordinates based on a polyconic projection, and as options: (1) interpolates the coordinates to a lattice, and (2) constructs an isarithmic line printer map of the data.

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: December, 1971
Similar Programs: GRID, INTRPOL, RGRID
Contributor: Computer Institute for Social Science Research
Michigan State University

LINEAR: Linear Pattern Recognition

Local lineations within a scatter of points are detected based on the elongation ratio for a standard deviation ellipse rotated so that its major axis trends along the lineation.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Similar Programs: CLUSTER, REGLAR
Contributor: Department of Geography
State University of New York, Albany

LISTER: Lister/Reproducing Punch

Emulates a lister/reproducing punch combination with greater flexibility and much less user effort. The available options include: (1) reordering of the sequence of variables, (2) integer to decimal and decimal to integer conversion, (3) serial numbering of card output, and (4) multiple copies of listed and punched output.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: December, 1971
Contributor: Department of Geography
Northwestern University

LOCATE: Ring and Sector Counting

Rapid counting on a ring and sector basis, about one or more base points, of a series of points whose locations are defined in terms of Cartesian Coordinates.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Contributor: Department of Geography
Northwestern University

Other Versions: (LOCATE)

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: May, 1970
Contributor: Department of Geography
University of Iowa

LOCPOT:

An input-output gravity model simulating industrial attraction

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: March, 1975
Contributor: Department of Geography
University of Georgia

LUAM: Land Use Allocation Model (Version 1)

Development of land use allocations according to the Walsh and Grava Land Use allocation model.

Computer: IBM 7040
Compiler: FORTRAN IV
Date Received: June, 1973
Contributor: Department of Town and Country Planning
University of Sydney

MAPDATA: Descriptive Statistics of Map Data

The input is a rectangular array of data, and the program computes the frequency distribution, the mean, variance, and crude moments. The program then aggregates the array to two by two blocks. Properties of these aggregated data may then be computed.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Contributor: Department of Geography
Northwestern University

MAPIT: Map Drawing on a Pen-and-Ink Plotter

The program performs map-drawing on a pen-and-ink plotter. Input is in the form of a two-dimensional X-, Y-coordinate system.

Computer: CDC 3600
Compiler: 3600 FORTRAN
Date Received: June, 1971
Similar Programs: GIPSY, POPMAP, SYMBOLS
Contributor: Computer Institute for Social Science Research
Michigan State University

MAP3D: Three Dimensional (3-D) Block Diagram

A FORTRAN program for plotting three-dimensional surfaces (topographic as well as statistical) as block diagrams.

Computer: CDC 6600
Compiler: 6000 FORTRAN EXTENDED
Date Received: March, 1972
Similar Programs: BLCK, ISOMET, PERS, SURGE, TRID
Contributor: Department of Geography
University of Minnesota

MAPLOT: A Map Plotting Program

To draw lines (streets, boundaries etc.) and to label them and to plot distinguishing marks at coordinate locations.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: October, 1971
Contributor: Department of Geography
University of Michigan

MAPTRAN: Multiple Facility Location for a Continuous Demand Surface

Map transformation algorithm for locating a set of points in a non-homogeneous population density area such that approximately equal populations are in the dirichlet regions of all points.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: January, 1974
Similar Programs: ALTERN, LAP, TORN
Contributor: Department of Geography
University of Iowa

MARKOV1: Regular Markov Chains

Analysis of regular markov chains.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June 1970
Similar Programs: MARKOV2
Contributor: Department of Geography
Northwestern University

Other Versions: (MARKOV1)

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: May, 1970
Contributor: Department of Geography
University of Iowa

MARKOV2: Simple Absorbing Markov Chains

Analysis of simple absorbing markov chains.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Similar Programs: MARKOV1
Contributor: Department of Geography
Northwestern University

Other Versions: (MARKOV2)

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: May, 1970
Contributor: Department of Geography
University of Iowa

MAVE: Binomially Weighted Smoothing

The program removes high frequency spatial components from a matrix of geographical data by use of a nine point binomially weighted local smoothing operation.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: August, 1970
Similar Programs: SMOOTH, SMOTHR
Contributor: Department of Geography
University of Michigan

MDISC: Linear Discriminant Analysis with Classification

A linear discriminant analysis with classification of subjects, based on likelihood ratios, assuming normal distribution and equal variances-covariances.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: November, 1972
Similar Programs: CENDA, DOSINE, DISITR, ITERIM, NEARNBR, NOLIDA
Contributor: Department of Geography
University of Waterloo

MINIMAP: Choropleth Mapping

Designed to produce choropleth maps from map and data file input utilizing small computers.

Computer: IBM 360
Compiler: FORTRAN IV
Date Received: June, 1973
Similar Programs: CMAP, CPLETH
Contributor: Department of Geography and Environmental Planning
Towson State College

MINPATH: Minimum Path Networks

Computation of minimum path networks from each of up to 1000 nodes to all other nodes in a given transportation system. The maximum number of edges is 2000, but multiple jobs may be processed.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Similar Programs: SPA, SPAN
Contributor: Department of Geography
Northwestern University

MLNEGBI: Maximum Likelihood Parameters of Negative Binomial Distribution

This program calculates the maximum likelihood parameters P and K of the negative binomial distribution, from the observed frequency distribution. Using P and K it then fits the negative binomial to the data and computes the chi-square values.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Similar Programs: NEGBIN
Contributor: Department of Geography
Northwestern University

MOCON: Central Moments, Skewness and Kurtosis from Crude Moments

To compute central moments, skewness, and kurtosis from curde moments.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Contributor: Department of Geography
Northwestern University

MOMENS: Computation and Mapping of Areal Moment Surfaces

The program calculates first, second or inverse first areal moments about population masses and optionally produces a line printer contour plot of resulting surface.

Computer: IBM 7040
Compiler: FORTRAN IV
Date Received: June, 1973
Contributor: Department of Town and Country Planning
University of Sydney

MONA: Mapping of Nominally Classified Activities

Produces choropleth-type maps of activity patterns from nominally classified activities.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: October, 1971
Contributor: Department of Geography
University of Iowa

MPLSQ: Surface-Fitting Program for Areally-Distributed Data from the Earth Sciences and Remote Sensing

Provides a method of reducing data recorded for a large number of areally-distributed sample sites to the form of simple isopleth maps using various techniques of trend surface analysis.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: December, 1971
Similar Programs: DYAD, POLYFIT, TREND, TRENDC
Contributor: Department of Geography
Northwestern University

MULTI: **Multi-Source Location-Allocation Algorithm (Exact)**

This program uses a minimum-storage (Backtrack) branch and bound technique to optimally locate M sources to serve N destinations in a Euclidean plane.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: January, 1974
Similar Programs: TWAIN, WEBER, WEBERI
Contributor: Department of Geography
University of Iowa

MUNKRE: **Transportation Problem**

Uses Munkree's Algorithm to solve the transportation problem of linear programming.

Computer: CDC 6500
Compiler: CDC FORTRAN EXTENDED
Date Received: November, 1971
Similar Programs: TRANS
Contributor: Computer Institute for Social Science Research
Michigan State University

NAYBOR: **Nearest Neighbor Statistic**

Calculation of the nearest neighbor statistic for a set of points distributed in a defined area on a plane. The program may also be used to identify nearest neighbors of each point.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: March, 1972
Similar Programs: EQUAL, GENEB
Contributor: Department of Geography
University of Michigan

NEARNBR: **Nearest-Neighbor-Discriminant Analysis**

Program to classify subjects on the basis of distance to nearest-neighbor individual.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: November, 1972
Similar Programs: CENDA, COSINE, DISITR, INTERIM, MDISC, NOLIDA
Contributor: Department of Geography
University of Waterloo

NEGBIN: Negative Binomial Probability Law

The program computes the individual terms of the negative binomial probability law.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Similar Programs: MLNEGBIN
Contributor: Department of Geography
Northwestern University

NODAC: Node Accessibility Indices

Computation of node accessibility indices (as defined by Shimbel and Katz) for transport networks.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Contributor: Department of Geography
Northwestern University

NOLIDA: Generalized Non-linear Discriminant Analysis with Classification

Computes non-linear discriminant functions which can then be used to classify unknown individuals or to reclassify those individuals from which the discriminant functions were originally derived.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: November, 1972
Similar Programs: CENDA, COSINE, DISITR, ITERIM, MDISC, NEARNBR
Contributor: Department of Geography
University of Waterloo

NONCEL: Simulation of Diffusion through Area of Non-Gridded Units

A program for simulation of diffusion through an area of non-grid units, such as political areas and/or other minor civil divisions.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: December, 1971
Similar Programs: HAAAG, NONCELO
Contributor: Department of Geography
Northwestern University

Other Versions: (NONCEL)

Computer: IBM 360/91
Compiler: FORTRAN G
Date Received: December, 1971
Contributor: Department of Geography
University of California, Los Angeles

NONCELO: Simulation of Diffusion through an Area on Non-Grid Units

The overlay of programs permits the simulation of diffusion according to standard Hagerstrand rules over an irregular lattice (non-grid units). The over-lay (1) simulates the location and number of knowers after each generation for each simulation (2) provides summary information on the spatial series generated and (3) provides summary information on the temporal series generated (which includes pen-and-ink plot output).

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: July, 1972
Similar Programs: HAAAG, NONCEL
Contributor: Department of Geography
Michigan State University

NONPLOT: Temporal Series Summary of NONCEL Output

To provide summary information of the temporal series generated on tape output by NONCEL. Includes pen-and-ink plot output.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: December, 1971
Similar Programs: HAGPLOT
Contributor: Department of Geography
Northwestern University

NORLOC: Multiple Location Analysis Program

The program solves problem of locating a number of facilities to serve a dispersed population. The routine finds 1) coordinates for positions of facilities, 2) service area of each facility and 3) capacity of each facility from basic input information.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Similar Programs: STDROP
Contributor: Department of Geography
State University of New York, Albany

NORM: Normality Check

The program tests data sets for normality using the Kolmogorov-Smirnov test. If the data are not normal in their original form, the following transformations can be applied: 1) Log, 2) Log-Log, 3) Square Root. The level of significance is chosen by the user.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: May, 1970
Contributor: Department of Geography
University of Iowa

Other Versions: (NORM)

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: August, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

NSCAT: Rapid Data Plotting

Rapid data screening. The program produces the N*(N-1)/2 printer plots, linear regressions, and simple correlations for all pairwise combinations of N variables.

Computer: IBM 360/67
Compiler: FORTRAN G
(Requires RUMPLOT subroutines from SHARE library)
Date Received: August, 1970
Similar Programs: REGRESS
Contributor: Department of Geography
University of Michigan

OPTREG: Optimal Regression Analysis

An optimal regression program to calculate regression equations on a dependent variable such that the residual sum of squares is a minimum.

Computer: IBM 360/91
Compiler: FORTRAN G
Date Received: December, 1971
Contributor: Department of Geography
University of California, Los Angeles

PCPA: Analysis of Paired Comparison Preference Data

A program to perform paired comparisons analysis on preferential data.

Computer: IBM 360/75
Compiler: FORTRAN G
Date Received: March, 1972
Similar Programs: REVPREF
Contributor: Department of Geography
McGill University

PELTO: Pelto D-Function and Relative Entropy

The program calculates two measures of the degree of mixing in multicomponent systems.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: August, 1970
Contributor: Department of Geography
University of Michigan

Other Versions: (PELTO)

Computer: CDC 6500
Compiler: 6400 FORTRAN EXTENDED
Date Received: December, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

PERS: Plotting Perspective Views of Surfaces

A set of FORTRAN subroutines that will plot a perspective view of a surface which is represented by a regular grid. Portions of the grid which are hidden from the viewer are not plotted.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: January, 1972
Similar Programs: BLCK, ISOMET, MAP3D, MRGE, TRID
Contributor: Computing Centre
University of British Columbia

PLOTMAP: Map Plotting from Tape

To draw outline maps on various projections to scale from data stored on magnetic tape. Will draw cartograms using double bivariate interpolation.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: October, 1971
Contributor: Department of Geography
University of Michigan

PLOTMP: Printer Grid Mapping Program

The program plots numerical valves and draws an outline around the area of interest.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Contributor: Department of Geography
State University of New York, Albany

PLOTS: Special Subroutine for Curve Plotting

A subroutine for time-series plotting on a computer line printer.

Computer: IBM 360/91
Compiler: FORTRAN G
Date Received: December, 1971
Contributor: Department of Geography
University of California, Los Angeles

POINT: Point-in-Polygon Testing

Used for testing whether a point, defined by Cartesian grid coordinates, is located inside, outside, or on an edge of a polygon defined by a set of coordinates indicating the polygon vertices.

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: December, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

POINTS: Point Set Mapped on a Plane or Torus

Computes descriptive measures for a point set mapped onto either a plane or torus. The method of order distance is used to summarize the point pattern.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: November, 1971
Contributor: Department of Geography
Northwestern University

POISSN: Poisson Probability Law

The program computes individual terms of Poisson Probability Law and, as an option, individual terms of the dispersed Poisson Probability Law.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Contributor: Department of Geography
Northwestern University

POLYFIT: Trend Surface Analysis for Degrees One Through Five

Fits algebraic polynomials of degree one through five to geographic data.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: August, 1970
Similar Programs: DYAD, MPLSQ, TREND, TRENDC
Contributor: Department of Geography
University of Michigan

POPMAP: Population Maps

The program reads rectangular coordinates (X,Y) and populations, and then draws population maps on the pen-and-ink plotter, with cities represented as circles.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: August, 1970
Similar Programs: GIPSY, MAPIT, SYMBOLS
Contributor: Department of Geography
University of Michigan

POPPYR: Population Pyramids

Program to plot population pyramids and compute basic demographic measures.

Computer: IBM 370/158
Compiler: FORTRAN G
Date Received: July, 1975
Contributor: Department of Geography
Virginia Polytechnic Institute

PREDEN: Dendrograph Input Generation

The program converts raw data into one of a variety of similarity measurements and offers option of punching these as a lower triangular matrix for direct input to the dendrograph program (DENDRO).

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Contributor: Department of Geography
State University of New York, Albany

PRESLOC: Evaluation of Total Costs of a System of Present Warehouses

The program calculates the costs of the present Warehouse locational configuration for any configuration input by the user. The program may be used in conjunction with program WARELOC.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: January, 1974
Contributor: Department of Geography
University of Iowa

PROLO: Probit and Logit Analysis

Permits probit or logit analysis (or both) to be carried out on a data set containing a maximum of 55 variables. The program is basically designed to handle cross-sectional data on which several years of observations may be handled for each member of the cross-section.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: December, 1971
Contributor: Department of Geography
Northwestern University

PYRAMID: Population Pyramid Subroutine

A program designed to produce age/sex population pyramids on the computer line printer.

Computer: IBM 360/67
Compiler: FORTRAN WATFIV
Date Received: April, 1972
Contributor: Department of Geography
Pennsylvania State University

RANKD: Ranking Data and Inspecting for Natural Breaks in the Distribution

The program rank orders original data and calculates differences between the ranked numbers and outputs them on a scaled bar graph.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Contributor: Department of Geography
State University of New York, Albany

REGLAR: Regular Pattern Recognition

The program determines the location and areal extent of regions in a point array exhibiting regular, rather than random, distributions. Regularity is assessed in terms of a nearest neighbor test.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Similar Programs: CLUSTER, LINEAR
Contributor: Department of Geography
State University of New York, Albany

REGRESS: Linear Regression Calculation and Plot

A FORTRAN program that calculates and plots (using pen-and-ink plotter) for any two data series, a scatter diagram, the line of regression, and delimits the zone for the standard error of estimate on the scatter diagram.

Computer: CDC 6600
Compiler: 6000 FORTRAN EXTENDED
Date Received: March, 1972
Similar Programs: NSCAT
Contributor: Department of Geography
University of Minnesota

REVPREF: Paired Comparisons Analysis from Revealed Spatial Preference Data

Applies the method of paired comparisons to data describing spatial choices of individuals. It develops a proximity matrix which describes the aggregate perceived similarity between locational types.

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: June, 1971
Similar Programs: PCPA
Contributor: Computer Institute for Social Science Research
Michigan State University

Other Versions: (REVPREF)

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: October, 1971
Contributor: Department of Geography
University of Iowa

RGRID: Map Plotting and Contouring

The program produces isamithmic maps on the line printer from scattered observations by first interpolating to a lattice and then contouring these regular values. A location map of the observations, a list of the values at the grid intersections, and the countoured map are returned.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: August, 1970
Similar Programs: LATMAP, RGRID2
Contributor: Department of Geography
University of Michigan

Other Versions: (RGRID)

Computer: CDC 6500
Compiler: FORTRAN EXTENDED
Date Received: June, 1971
Contributor Computer Institute for Social Science Research
Michigan State University

RGRID2: Map Plotting and Contouring

Similar to RGRID; however, this program allows up to eight variables to be mapped.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: March, 1972
Similar Programs: LATMAP, GRID
Contributor: Department of Geography
University of Michigan

RIDGREG; Ridge Regression

The program performs ridge regression analysis for a dependent variable and up to 100 independent variables; the data input must be normalized.

Computer: IBM 360/75, 370/165
Compiler: FORTRAN G
Date Received: March, 1973
Contributor: Department of Geography
University of North Carolina

RSTAR: Measure Degree of Association Between Spatially Distinct Observations

This program is designed to measure either the degree to which one retail function (attribute) is spatially associated with any other or the degree to which one shopping center (observation) is like any other.

Computer: IBM 370
Compiler: FORTRAN G
Date Received: May, 1976
Contributor: Department of Geography
University of Denver

SAMPLE: Simulation of Random Sampl

Simulation of random sampling with binomial distribution and normal distribution probability tests.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Contributor: Department of Geography
State University of New York, Albany

SANDM: Spatial Series Summary of NONCEL Output

This program reads the output tape produced by the spatial simulation program NONCEL and produces cell means and variances for each generation of each simulation.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: December, 1971
Similar Programs: HAGMEVA
Contributor: Department of Geography
Northwestern University

SCALE: Solution of the Law of Categorical Judgement

Transforms original data from the psychological scaling technique known as the method of successive categories to interval scale values.

Computer: [illegible]
Compiler: 6000 RUN FORTRAN
Date Received: December, 1971
Contributor: Department of Geography
Northwestern University

SDIS: Spherical Distances

Computes a matrix of distances from a set of points defined by their latitude/longitude coordinates.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: August, 1970
Similar Programs: DSTAZ, GEODIS, SPHERE
Contributor: Department of Geography
University of Michigan

SINGLE: Single Fourier Series

Analyzes data that may be represented by curves. The following expression is used for the single fourier infinite;

Y = (A(0)/2) + SIGMA(N) [A(N) COS(N*PI*X)/L + B(N) SIN(N*PI*X)/L]

where N = 1,2,...,infinity.

Computer: IBM 360/91
Compiler: FORTRAN G
Date Received: December, 1971
Similar Programs: DOUBLE
Contributor: Department of Geography
University of California, Los Angeles

SIRES: Simulation of Urban Residential Segregation

This program uses Monte Carlo simulation procedures to explore the phenomenom of urban residential segregation. The program is designed as a teaching device in that the student can vary a number of parameters such as initial segregation pattern, size and number of neighborhoods to be analyzed, white and black discrimination criteria for moving into or out of a neighborhood, etc. The CDC version is strictly a batch oriented program, but the IBM version is designed for batch or interactive use.

Computer: CDC 6500
Compiler: FORTRAN EXTENDED
Date Received: July, 1975
Similar Programs: LANDUSE, SOLUP
Contributor: Department of Geography
Michigan State University

Other Versions: (SIRES)

Computer: IBM 370/155
Compiler: FORTRAN G
Date Received: January, 1976
Contributor: Department of Geography
Virginia Polytechnic Institute and State University

SLOPE: Spatial Derivative Program

The program calculates the absolute value of the gradient (slope) for data given as a geographical matrix.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: August, 1970
Contributor: Department of Geography
University of Michigan

Other Versions: (SLOPE)

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: December, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

SMOOTH: Binomially Weighted Smoothing

Removes high frequency spatial components from a matrix of geographical data by use of a nine point binomially weighted local smoothing operation.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: December, 1971
Similar Programs: MAVE, SMOTHR
Contributor: Department of Geography
Northwestern University

SMOTHR: Matrix Smoothing

This assembly language subroutine smooths an array of values (assumed to be spatial data) using the method described by Holloway. It is designed to be called as a subroutine in a FORTRAN program.

Computer: IBM 360/370
Compiler: IBM Assembler
Date Received: February, 1977
Similar Programs: MAVE, SMOOTH
Contributor: Department of Geography
University of California, Santa Barbara

SMRATE: Standard Mortality Rates

The program computes standardized mortality or morbidity ratio and rate for areal-administrative units. The computational method employed adjusts crude death rates for variations in age distribution. SMRATE produces relative mortality ratio as a percentage of the national rate. Standard mortality rate is scaled by user.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Contributor: Department of Geography
State University of New York, Albany

SNORT: Sample Normality Testing

The program tests for normality or lognormality through a comparison of the sample frequency distribution and the normal frequency distribution of the same mean and standard deviation as the sample.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Similar Programs: NORM
Contributor: Department of Geography
State University of New York, Albany

SOLUP: Simulation of Land Use Patterns

To construct a map on the printer on which the land use activity which would earn the highest economic rent is computed and identified for each point location. Economic rent is computed from the formula:

R = E(P-A) - E*F*K.

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: June, 1971
Similar Programs: LANDUSE, SIRES
Contributor: Computer Institute for Social Science Research
Michigan State University

Other Versions: (SOLUP)

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: October, 1971
Contributor: Department of Geography
University of Iowa

SPA: Shortest Path Algorithm

A program to determine the shortest paths through a network (1) between all pairs of nodes, or (2) between a specified subset of nodes and all other nodes for specified route nodes.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: January, 1972
Similar Programs: MINPATH, SPAN
Contributor: Department of Geography
University of Iowa

SPACE:

Computes descriptive statistics for an areal distribution with optional normal probability and/or normal correlation surface tests.

Computer: IBM 7040
Compiler: FORTRAN IV
Date Received: June, 1973
Contributor: Department of Town and Country Planning
University of Sydney

SPAN: Shortest Path Analysis of Networks

Given a network of points interconnected by lines of specified lengths, the program finds the shortest paths between specified pairs of points.

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
6000 COMPASS
Date Received: June, 1971
Similar Programs: MINPATH, SPA
Contributor: Computer Institute for Social Science Research
Michigan State University

SPECTR: Computation of Two-dimensional Power Spectra

Computes autocorrelation functions and two-dimensional power spectra for the investigation of the frequency components or wave forms present in an undulating surface.

Computer: IBM 360/91
Compiler: FORTRAN G
Date Received: December, 1971
Contributor: Department of Geography
University of California, Los Angeles

SPHERE: Spherical Distances

Reads a vector of latitude/longitude coordinates and produces the complete distance matrix for all points.

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: August, 1971
Similar Programs: DSTAZ, GEODIS, SDIS
Contributor: Computer Institute for Social Science Research
Michigan State University

STDROP: Algorithm for Dropping Retail Establishments Failing to meet threshold criterion.

Program allocates population to centers, calculates the sales of centers, and discovers whether or not the sales meet the given threshold criterion. If not, the center is either dropped or transformed to a lower central class.

Computer: Honeywell
Compiler: FORTRAN V
Date Received: August, 1972
Contributor: Center for Urban and Regional Studies
Helsinki University of Technology

SURGE: Surfaces from Geographic Grid Data

The program plots three-dimensional surfaces from geographic grid data using the pen-and-ink plotter.

Computer: IBM 360/50
Compiler: FORTRAN G
Date Received: June, 1971
Similar Programs: BLCK, ISOMET, MAP3D, PERS, TRID
Contributor: Department of Geography
University of Rhode Island

Other Versions: (SURGE)

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: December, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

SYMBOLS: Plot of Graduated Symbols Map

A FORTRAN program for pen-and-ink plotting of graduated symbols maps where the area of each symbol is proportional to the data value it represents.

Computer: CDC 6600
Compiler: 6000 FORTRAN EXTENDED
Date Received: March, 1972
Similar Programs: GIPSY, MAPIT, POPMAP
Contributor: Department of Geography
University of Minnesota

TAXON: Groups Pairs of Similar Operational Taxonomic Units (OTU) to Form Clusters or Nested Hierarchies

Value lies in determining the distances, in N-Dimensional space, between one (OTU) and its next most similar (OTU). Option to group only those (OTU)'s which are spatially contiguous.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Similar Programs: CNGRP, CONGRP, DISGRP
Contributor: Department of Geography
State University of New York, Albany

THIESEN: Thiessen Polygons

Calculates vertices of Thiessen polygons for a set of points in two-space and then plots the polygons using a pen-and-ink plotter.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: November, 1972
Contributor: Department of Geography
University of Michigan

TORG: Coordinates from Distances

The program reads a matrix of distances, and then produces a vector of plane coordinates.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: August, 1970
Similar Programs: TORG2
Contributor: Department of Geography
University of Michigan

Other Versions: (TORG)

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: August, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

TORG2: Coordinates from Distances

The program reads a matrix of distances, and then produces a vector of plane coordinates assuming the distances relate N points in a two dimensional Euclidean space.

Computer: IBM 360/67
Compiler: FORTRAN G
Date Received: March, 1972
Similar Programs: TORG
Contributor: Department of Geography
University of Michigan

TORN: Tornqvist Multiple Location Algorithm

The program determines the optimum locations for a number of centers C1, C2, C3,...,Cm with respect to the transportation cost of serving N destinations.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: January, 1974
Similar Programs: ALTERN, LAP, MAPTRAN
Contributor: Department of Geography
University of Iowa

TRANS: Solution of the Transportation Problem

The transportation problem is used to determine the minimum total distance (or cost) and the individual flows involved in moving commodities between M number of supply points and N number of demand points given the distance between all pairs of supply and demand points.

Computer: IBM 360/65
Compiler: FORTRAN G
360/MPS
Date Received: May, 1970
Similar Programs: MUNKRE
Contributor: Department of Geography
University of Iowa

Other Versions: (TRANS)

Computer: CDC 6500
Compiler: CDC FORTRAN EXTENDED
CDC APEX
Date Received: December, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

TREND: Trend Surfaces for Degrees One through Six

This program determines any or all trend surfaces from first through sixth degree. Line printer maps may be produced depicting the original data, the contoured surfaces, and the residuals from the surfaces.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: May, 1970
Similar Programs: DYAD, MPLSQ, POLYFIT, TRENDC
Contributor: Department of Geography
University of Iowa

Other Versions: (TREND)

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: August, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

TRENDC: Identification of Spatial Patterns of Distance-Decay Type

A program designed to isolate successive distance-decay patterns (derived in a stepwise fashion) around central points and the residuals from these patterns.

Computer: IBM 360/91
Compiler: FORTRAN G
Date Received: December, 1971
Similar Programs: DYAD, MPLSQ, POLYFIT, TREND
Contributor: Department of Geography
University of California, Los Angeles

TRID: Three-Dimensional Isometric Diagrams

A program for pen-and-ink plotting of three-dimensional surfaces as block diagrams. Portions of the grid hidden from the viewer are not plotted.

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: July, 1972
Similar Programs: BLCK, ISOMET, MAP3D, PERS, SURGE
Contributor: Computer Institute for Social Science Research
Michigan State University

TWAIN: Two-Source Location-Allocation Algorithm

This program generates all geometrically possible two-group partitions of N points in a plane. It optimally solves the two-source location-allocation problem.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: January, 1974
Similar Programs: MULTI, WEBER, WEBER1
Contributor: Department of Geography
University of Iowa

USDATA: Procedure for Generating a Data File

36 Socio-economic variables have been tabulated by state and stored internally by data statements. The user can call variables, make computational alternations and add up to 50 newly generated variables. Option is available to have 50 original variables punched or printed out.

Computer: U IVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Contributor: Department of Geography
State University of New York, Albany

VALRAT: Contiguity Measures

Computation of statistics for contiguity measures to provide an evaluation of the hypothesis of randomness in the areal arrangement of the values of a specified property.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: June, 1970
Similar Programs: CONRAT, KCOLOR, VALRAT1
Contriburor: Department of Geography
Northwestern University

VALRAT1: Contiguity Measures

Similar to VALRAT; however this version includes a test for normality.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: October, 1971
Similar Programs: CONRAT, KCOLOR, VALRAT
Contributor: Department of Geography
University of Iowa

ther Versions: (VALRAT1)

Computer: CDC 6500
Compiler: 6000 FORTRAN EXTENDED
Date Received: December, 1971
Contributor: Computer Institute for Social Science Research
Michigan State University

WARELOC: A Heuristic Program for Locating Warehouses With Facility Costs

A heuristic algorithm to solve for the 'optimal' number, sizes, and locations of warehouses at prespecified nodes in a network, where there are fixed costs associated with the facilities. The program has its own shortest path routine (shimbel method).

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: January, 1974
Contributor: Department of Geography
University of Iowa

WEBER: Weberian Weight Table

The program attempts to find the optimum location, for a facility so as to maximize accessibility to or from a number of points, according to a vector equilibrium model that assumes a uniform transport surface.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Similar Programs LAP, MULTI, TWAIN, WEBERI
Contributor Department of Geography
State University of New York, Albany

WEBERI: One-Source Location Algorithm

This program uses a second order iterative technique to solve for the point of minimum aggregate travel of a set of weighted points in a plane.

Computer: IBM 360/65
Compiler: FORTRAN G
Date Received: January, 1974
Similar Programs: LAP, MULTI, TWAIN, WEBER
Contributor: Department of Geography
University of Iowa

XCUT:

To provide rapid, approximate solutions to the multi-source Weber problem.

Computer: CDC 6400
Compiler: 6000 RUN FORTRAN
Date Received: October, 1974
Contributor: Department of Geography
State University of New York, Buffalo

XMAPP: Choropleth Mapping Using the Line Printer

The program generates character symbol choropleth maps for terminal or line printer output. Data files may be entered by the user, reside in storage, or be system generated when testing map outlines. Several maps of one or more region may be produced in a single run.

Computer: CDC Cyber 72
Compiler: FORTRAN EXTENDED
Date Received: February, 1977
Similar Programs: CMAP, CPLETH, COBMAP, INTRMAP
Contributor: Department of Geography
Eastern Illinois University

ZSCORE: Data Standardization

The program first computes means and standard deviations of variables and then determines Z-scores. Output includes a punched deck of the Z-score matrix.

Computer: UNIVAC 1108
Compiler: FORTRAN IV
Date Received: March, 1973
Similar Programs: NORM
Contributor: Department of Geography
State University of New York, Albany

A Self Instructional Package:

How to Digitize a Topographic Map

by

David M. Mark

INTRODUCTION

Digitization can be defined as the process by which "analog measures" such as length or location on a map, are converted into "digital computer-usable form" (Peucker, 1972, p.72), in other words, into numbers. When a topographic surface is digitized using surface-specific points, knowledge of the form of the surface being sampled (usually obtained by a visual inspection of a contour map or the land surface itself) is used to select points or lines which contain a maximum amount of "information." The digitization process involves 3 relatively independent phases:

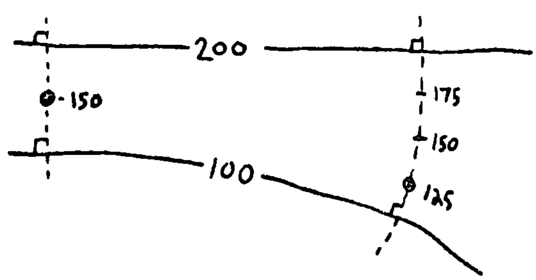
1) The selection of the points which are to be used to represent the surface;

2) the determination of the elevations (Z - co-ordinates) of each of the selected points; and

3) the determination of the planimetric locations (X and Y co-ordinates) of each of the selected points.

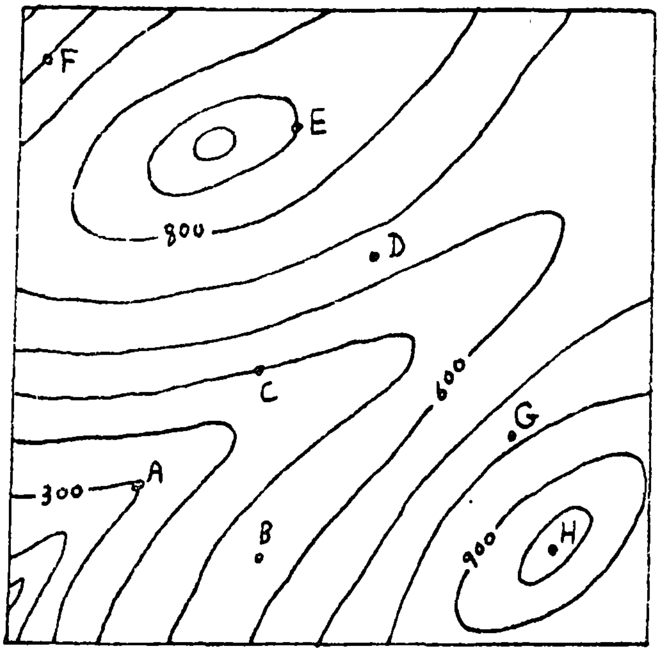
The present package is designed to instruct the reader in how to perform particularly phase 1) and to some extent phase 2) of this process.

A: Reading Contours and Interpolating: When reading a contour map, one should first determine the contour interval, the difference in elevations between adjacent contours. This will usually be the same throughout a map, and is often printed in the margin of the map. If not, it can be determined, simply: Find two nearby contours on a slope which both have their elevations indicated; count the number of spaces between these two labelled contours and divide this number into the elevation difference between the two contours. If one wishes to know the elevation of a point not on a contour, one must use interpolation, generally linear interpolation. First, imagine a straight or curved line perpendicular to both of the neighbouring contours and passing through the point. Next imagine that this line is divided into a number (2, 3, 4, 5, ...) of equal divisions in such a way that one of the divisions falls on the point. Each of the n equal division represents (contour interval / n) units of elevation, and thus the required elevation can be determined. See diagram below.



(1)

Determine the elevations of the eight points on the contour map below.



②

What is the contour interval of the map?

If you got any point wrong,* compare the correct answers with the map and if necessary, re-read section A. If you still do not understand, check with the instructor.

B: PEAKS AND PITS

Perhaps the most importand and most easily recognizable surface-specific points are peaks and pits, which are local maxima and minima respectively on the surface. A peak is a point which is higher than all the immediately surrounding points; to put it another way, the land surface slopes downhill away from a peak in all directions. A peak is shown on a contour map as a

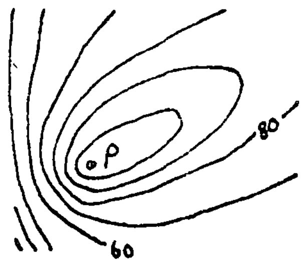
* A = 300   B = 550   C = 500   D = 650   E = 900   F = 600   G = 775 (approx)
H = 1050 (approx)   contour interval = 100 units

closed contour loop which is higher than the surrounding territory. When there is no evidence to the contrary, the peak should be located at the centre of the closed contour (see diagram 3).

P = peak

③

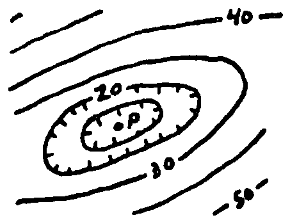
In some cases, however, the surrounding slopes can be used to determine more precisely where the highest point (peak) is located.



④

Here in diagram 4, the slopes clearly indicate that the peak is most likely to be toward the left end of the closed contour. A peak need not be the highest point on a map, and indeed most topographic maps show many peaks; a peak need only be higher than the immediately surrounding land.

Pits are the exact inverse of peaks; they are points which are lower than the surrounding land and are also shown by a closed contour, in this case <u>lower</u> than the surrounding territory. <u>Usually</u> (but not always) contour maps distinguish pits from peaks by placing small "hachures" along the closed depression contours. These lines point downhill and turn into the pit (see diagram 5).

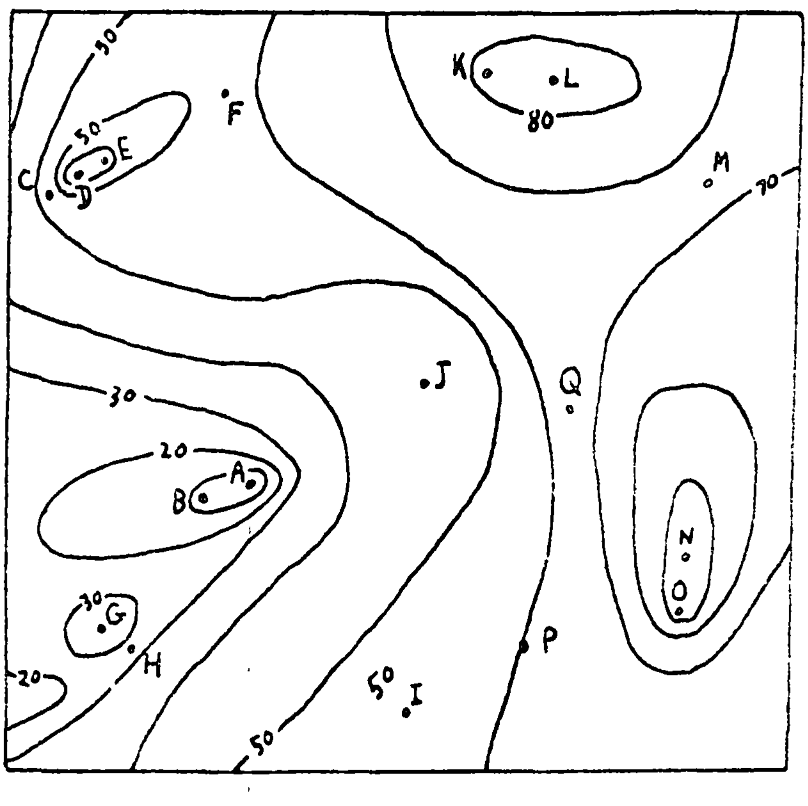


P = pit

⑤

Once again, the form of the surrounding slopes should be used to determine the exact location of the pit. Except in special types of areas (limestone areas or "karsts" and certain types of glacial topography are examples), pits are rather rare in temperate landscapes.

On the following map, determine which of the lettered points are ideal locations of peaks and of pits, and distinguish between peaks and pits. Depression contours are not hachured in this example.



⑥

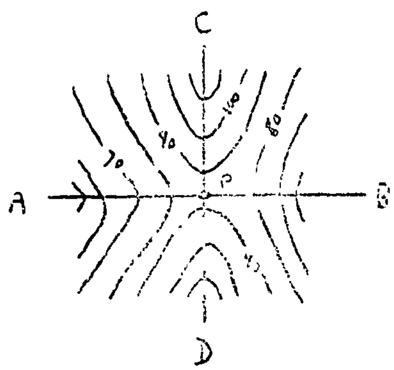
Answers

	Peak	Pit	Neither
A		X	
B			X
C			X
D		X	
E			X
F			X
G	X		
H			X
I		X	
J			X
K			X
L	X		
M			X
N			X
O	X		X
P			X
Q			X

Check your answers against the correct answers on the left. If you did not make <u>any</u> mistakes, go to part C.

If you thought that G was a pit, or that D or I was a peak, you have mis-read the contours. Consider point I: The area <u>outside</u> the closed 50 contour is between 50 and 60 (the point P is on the 60 contour); thus the area <u>inside</u> the closed 50 contour <u>must be below</u> 50 and I must therefore be a pit. The same argument holds for point D, while the reverse argument can be applied to show that G must be a peak. If you made any other mistakes, please re-read section B, and if you are still unsure, consult the instructor.

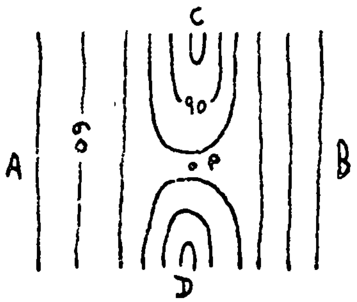
C: <u>PASSES</u>: A <u>pass</u> can be defined as a point which is a maximum along a line in one direction and at the same time minimum along a line at right angles. On a contour map, it usually appears as follows:

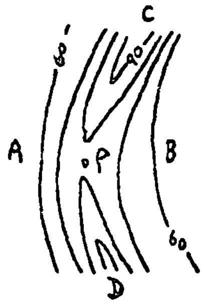


(7)

The point P is a pass: It is a maximum along line A-B and a minimum along C-D. Its elevation is about 85 units. Of course the profile lines A-B and C-D can both be curved, or one set of contours less curved than the other, producing similar but not identical apprearances on the contour map:

(8)





(9)

All of these passes have one feature in common: As one goes in a circle around the point, the land surface is, in turn, higher - lower - higher - lower, than the pass itself. The elevation of the pass can be estimated from the relative distances from the pass to the neighbouring higher and lower contours.

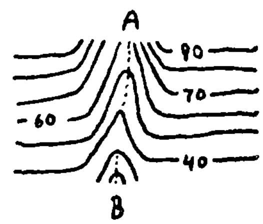
Turn back to the map on page 5 (diagram 6). Which of the lettered points on that diagram is a pass?

Anser: C, H, and M only. While F and Q may <u>look</u> like passes, examine the sequence of elevations as one goes around the point; you should see that these do not show the "higher - lower - higher - lower" sequence characteristic of a pass. If you are unsure of this, see the instructor.

D: <u>Course Lines and Ridge Lines</u>:

So far, we have examined three types of surface-specific <u>points</u>. This section looks at surface-specific <u>lines</u> of two sorts: Course lines and ridge lines.

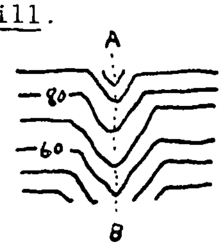
A course line will be failiar to most as the centre-line of a"valley." These valleys have sides higher than the centre-line, and in humid areas, these centre-lines are generally occupied by streams or "water-courses" (hence the name "course-line"). Even where streams are lacking, course-lines can be recognized by "V-shaped" contours in which the points of the "V"s point uphill.



(10)

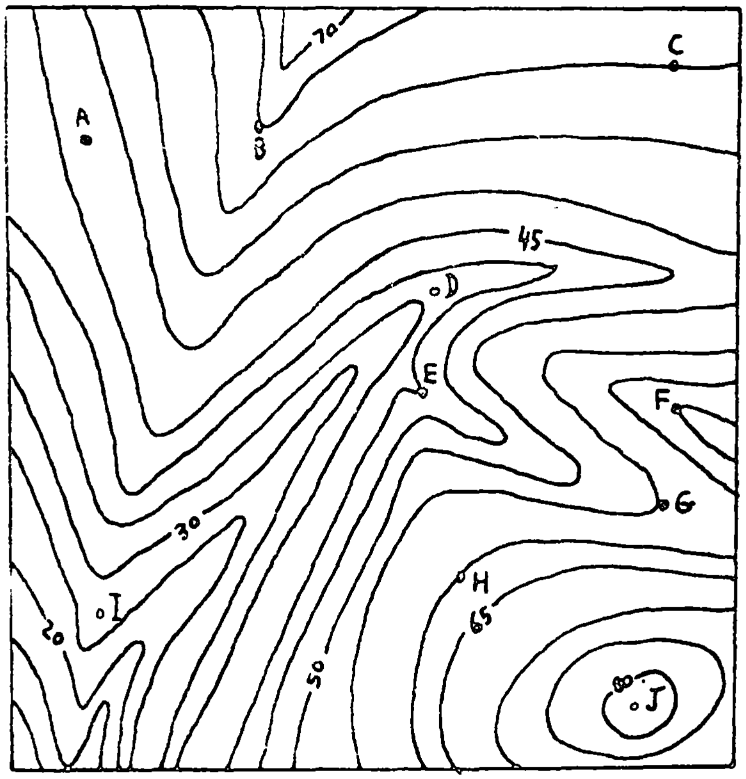
In the above diagram, A-B is a course-line. There are certain points along course-lines which are relatively more significant and which should be specified when a map is digitized. These include course-junctions, i.e., points where two course-lines merge or divide, points where a course-line bends, points where a course-line starts or ends, and points where a course-line enters or leaves a map sheet or study area.

Ridge lines are the exact inverse of course lines, but are not as obvious since they are not marked by obvious features (as are often course lines by streams). Once again, "V-shaped" contours mark a ridge, but in this case the "V"s point downhill.



(11)

In the above diagram A-B is a ridge line. As in the case of streams, ends, junctions and bends form important points along a ridge as well as points where ridges enter or leave the map. On the map below, sketch in the ridge lines and course lines using the indicated symbols, indicate which of the lettered points are ridge lines, course lines or neither.



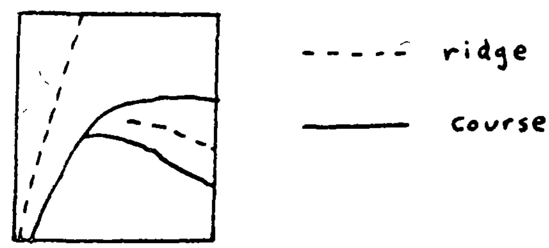
(12)

	Ridge	Course	Neither
A			
B			
C			
D			
E			
F			
G			
H			
I			
J			

Answers

	Ridge	Course	Neither
A			X
B	X		
C			X
D		X	
E		X	
F	X		
G		X	
H			X
I	X		
J			X

If you got any wrong, re-check the map. If you really want to argue, point H might be considered to be on a ridge, although this ridge, if it is a ridge, is not very well marked. If you are unsure of why the points are as they are, re-read section D and, if necessary, consult with the instructor.



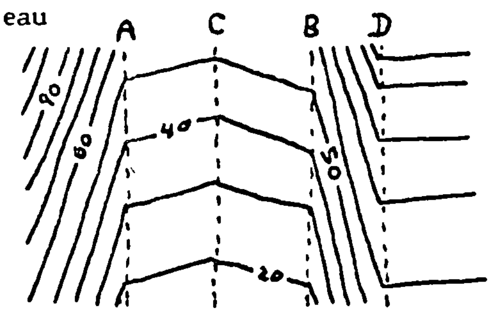
D 1: <u>Relationships between ridges and course lines, and peaks, pits and passes</u>

Course lines, when traced downhill, usually lead to either the ocean or a pit (often they simply lead off the map at hand). They may, especially in arid areas, simply end in an area of flat ground or an alluvial fan. Similarly, when ridge lines are followed uphill they often reach a peak, although sometimes they end on a slope or at a flat plateau. Ridges and course lines often bear a special relationship to passes: The two topographic "lows" encountered as one goes around the pass are often the beginnings of course lines, while the intervening highs are usually the starting points of ridges leading up to peaks. In fact, passes often form minima on continuous ridges passing between two peaks. Refer back to diagrams 7, 8 and 9.

The lines C-P-D form ridges on all three diagrams. PA and PB in 7 form well marked courses, while PB in 9 is a less well developed course line.

E: Breaks of slopes:

Another type of surface specific line is the line marking a break of slope, where the angle of slope of the land changes suddenly. This is reflected in a sharp change in the spacing of the contours. Examples are the lines where the steep slopes of a valley meet a broad, flat valley floor or gently sloping plateau



13

Here, C is a course line, while A, B, and D are important breaks of slope.

F: Digitizing a Map:

On the following attached section of a topographic map, do the following:

a) Place dots (•) at all the peaks, pits, and passes.

b) Sketch in the ridges, course lines and breaks of slope (if any).

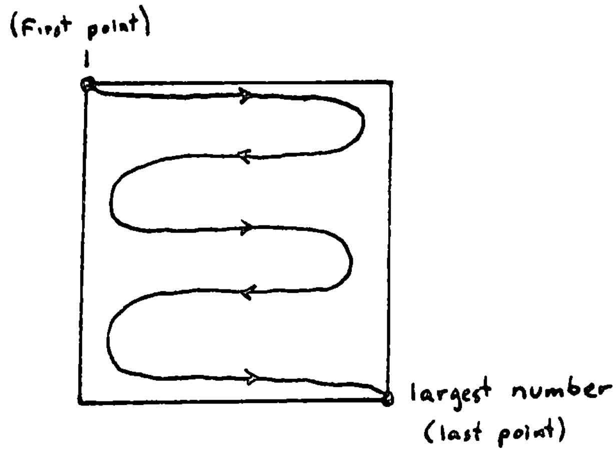
c) Place dots at significant points along the course lines and ridges.

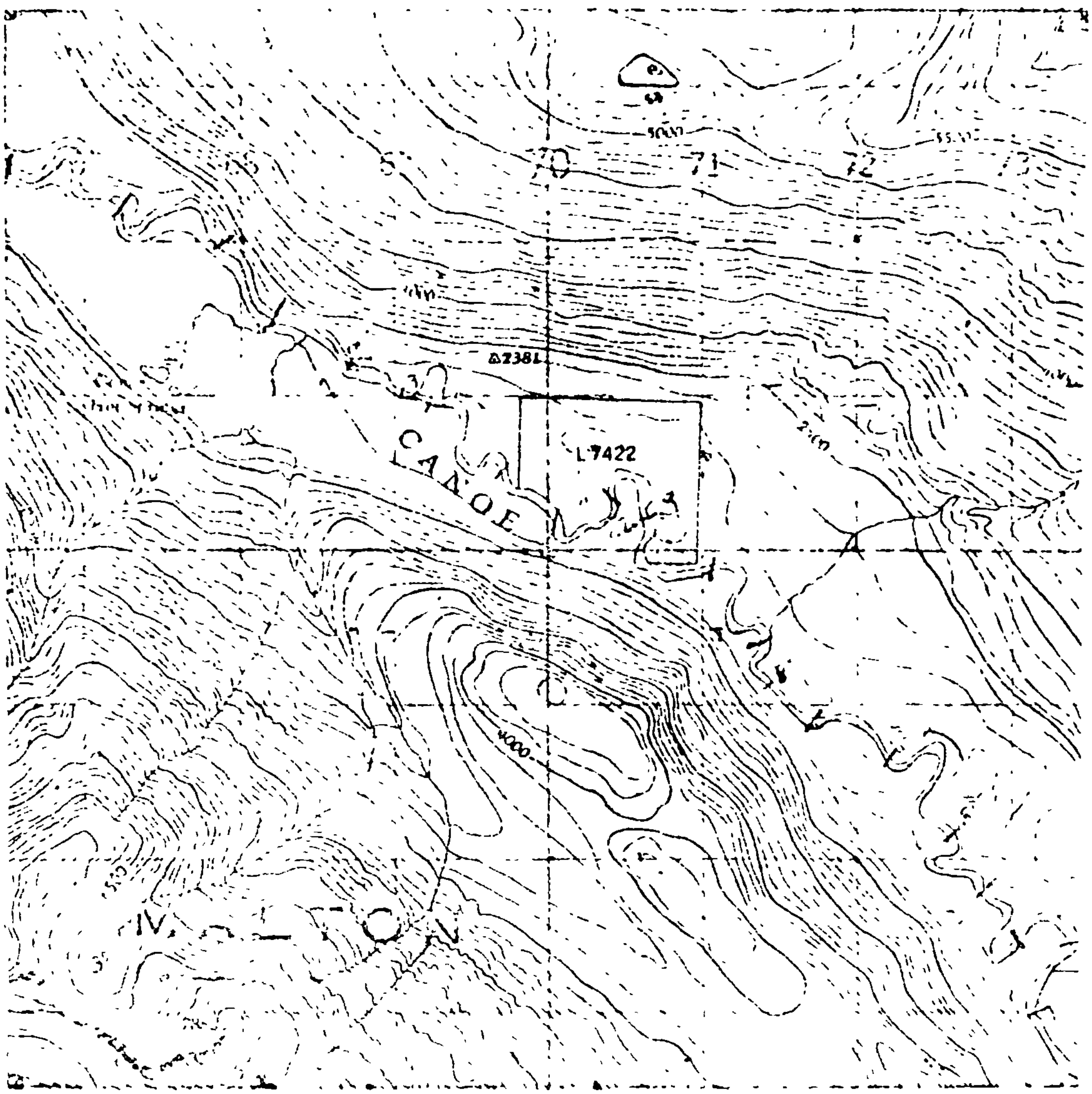
d) If there are any large areas of the map with no points, add one or two points at arbitrary locations within these areas. Include also points at corners or bends in the map boundary.

e) Number the points beginning in one corner and going back and forth in strips (see sketch below).

f) Make a list of the numbers and determine the elevations of each point by interpolation.

g) Hand in the map and a list of points to the instructor.





70
71
72
73
Δ2381
L 7422
CANOE

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

MANEJO DE SISTEMAS DE INFORMACION GEOGRAFICA EN PLANEACION

APENDICE BIBLIOGRAFICO

NOVIEMBRE, 1977.

UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE INGENIERIA
CENTRO DE COMPUTACION
LABORATORIO DE COMPUTACION GRAFICA

PROCESAMIENTO Y REPRESENTACION DE VARIABLES ESPACIALMENTE DISTRIBUIDAS POR MEDIO DE UN COMPUTADOR DIGITAL

Bernardo A. Paris

Caracas, 1.975

I N D I C E

Pág.

(continuación) Pág.

I.- RESUMEN

Se justifica la aplicación de técnicas de Computación Gráfica al despliegue de variables con significado espacial. Se comentan varios aspectos relacionados con el "HARDWARE" y el "SOFTWARE" necesarios. Algunas definiciones básicas sobre variables espaciales y su representación por medio de superficies son dadas y comentadas.

En la sección correspondiente a algorítmos y programas se comienza por dar dos implementaciones para el análisis y representación gráfica de variables unidimensionales como base para el tratamiento de variables multidimensionales.

Se introduce el concepto de matriz de superficie y algunos procedimientos para obtenerla por interpolación. Se establece que es necesario encontrar métodos más eficientes.

En el apartado 2.5 se comienza a reseñar las diversas formas de representación de la superficie. Se discuten las proyecciones ortogonales, oblícuas, en perspectiva, estereográficas y modelos tridimensionales. Los respectivos programas para computador son presentados.

En 2.6 se muestran algunas técnicas para la codificación y procesamiento de líneas y elementos de superficie junto con algunos

comentarios y descripciones referentes a los problemas de la recta y la superficie tapada.

El método de Monte Carlo es utilizado para la generación aleatoria de N caracteres en un área dada, siendo N proporcional a una variable bajo estudio.

Se discuten algunos intentos de generar secuencias animadas por computador para estudiar el aspecto dinámico de fenómenos.

En las conclusiones se comentan la utilidad de los desarrollos.

II.- INTRODUCCION

El fin de éste trabajo es presentar un conjunto de métodos, algorítmos y sus implementaciones desarrollados por el autor, y que estan destinados a la recolección, procesamiento y representación gráfica de información con significado espacial, es decir información de tipo social, geográfico, físico, etc. Inmediatamente se desprende del planteamiento anterior que, dado el volumen altamente elevado de datos a procesar con los correspondientes cálculos laboriosos y repetitivos, es indispensable la utilizacion de un Computador Digital.

Un Computador normalmente cuenta con medios de salida gráfica muy rápidos y exactos, tales como la impresora de línea, la pantalla de rayos catódicos y el delineador digital. Por lo tanto, es obvio que gran parte del esfuerzo de programar una máquina esté dirigido a procesar información para luego representarla en lo posible en forma de gráficos, diagramas, imágenes, mapas, etc. La razón de esto es inmediata; la especialidad llamada Computación Gráfica (Comunicación entre hombre-máquina en forma gráfica) utiliza el canal de entrada de mayor capacidad de recolección de información del hombre (la visión) junto a la excelente habilidad de cerebro humano para el reconocimiento de formas*. Brevemente se puede decir que una imagen producida por un Computador

* Traducción de los terminos en Ingles "Pattern Recognition".

maximiza la transmisión de información a un humano.

Tomando en cuenta lo anterior, se ha trato de crear un cuerpo coherente de métodos de despliegue gráfico, a fin de poder manejar y representar conjuntos de datos aparentemente informes, salidos o recolectados del medio ambiente, para asignarles su verdadero contenido espacial.

Desde un punto de vista mas formal, la Teoria de Autómata establece que cualquier problema que sea "computable" puede ser resuelto mediante una Máquina de Turing. La condición de computabilidad queda definida por el llamado Teorema de Turing y que dice, que si un proceso puede ser reducido a una secuencia finita de pasos perfectamente definidos, podrá ser automatizado.

La mayoria de los procesos de representación espacial de información pueden ser reducidos a un conjunto finito de operaciones simples y son por lo tanto, potencialmente susceptibles de ser llevados a cabo en forma automática.

Principalmente se tratará en este trabajo el problema de la generación automática de imágenes de conjuntos de resultados provenientes del procesamiento de datos recolectados en el mundo real. Estas imágenes vendrán en la forma de gráficos, de verdaderos modelos tridimensionales, proyecciones en dos dimensiones del tipo ortogonal, oblícua, axonométrica, etc. Finalmente se introducirá el aspecto dinámico del problema con la incorporación de la -

variable tiempo, es decir, la producción de secuencias conectadas temporalmente de un fenómeno.

III.- DESARROLLO

1. Aspectos Básicos.-

1.1. Equipo Necesario (Hardware)

El eje de los equipos necesarios lo constituye el computador. Este es un dispositivo capaz de aceptar información, procesarla bajo el comando de una secuencia de instrucciones (programa) almacenadas en la memoria, y devolver los resultados a través de un medio de salida. Cualquier computador de los que existen actualmente en el mercado es capaz de producir imágenes elementales de conjuntos de datos que representen distribuciones espaciales. El único requerimiento es que se disponga de una memoria lo suficientemente grande como para poder almacenar la gran cantidad de información necesaria o en su defecto un sistema operativo que utilice técnicas de memoria virtual (suponiendo la existencia de almacenamiento secundario).

Como medio de salida elemental debe disponerse de una impresora normal de 132 caracteres por línea, aunque si se está dispuesto a aceptar las lógicas limitaciones de tiempo y espacio, se puede hasta - utilizar la salida de un teletipo.

El dispositivo de salida más exacto y versátil que existe actualmente lo constituye el Delineador Digital (2), las características, ventajas y desventajas de cada uno de los dispositivos anteriormente mencionados estan reseñados por el autor en la Ref. 1.

1.2. Lenguajes de Programación Software)

Indudablemente que se necesita un lenguaje algorítmico de alto nivel tipo Fortran o Algol. La inmensa mayoría de los desarrollos a nivel mundial han sido hechos usando estos dos lenguajes. Existe en la literatura intentos experimentales de crear lenguajes orientados al problema (POL) (3) para el - tratamiento particularizado de líneas, superficies, contornos, etc. pero hasta la fecha no hay nada -- concluyente.

El lenguaje usado por el autor es el Fortran IV implementando mediante un compilador bastante poderoso y versátil por la Compañía Burroughs en un Computador B-6500. Un conjunto de seis (6) rutinas -

básicas de computación gráficas creadas por el grupo de Software de la Compañía Calcomp, permiten comandar adecuadamente el Delineador Digital anexado al B-5500

Existe en la U.C.V. otro poderoso computador, ubicado en la Facultad de Ciencias, pero por carecer de Delineador Digital no ha sido posible utilizar esta instalación en el diseño de técnicas de Computación Gráfica.

1.3. Variables Espacialmente Distribuidas. Superficies

Si se tiene una distribución espacial de valores sobre una región, cada punto puede ser identificado por sus coordenadas X e Y, relativos a un origen arbitrario. A cada punto se le puede hacer corresponder el valor de la variable de forma tal que le sea asignado una "altura" proporcional. Los tres conjuntos $\{X\}$, $\{Y\}$, $\{Z\}$ definidos en esta forma constituyen un conjunto $\{D\}$ de datos que posteriormente definirán una superficie $\{S\}$.

Existen varias formas de recolectar la información para construir la superficie $\{S\}$. La primera consiste en utilizar un retículo elementos cuadrados o rectangulares. Los datos se recojen por lo tanto a intervalos regulares. El segundo método sigue los cambios de dirección de una línea de altura constante para producir una curva de nivel. Usualmente la información es recolectada en puntos de coordenadas aleatoriamente distribuidos; esto puede producir que se "escapen" puntos singulares. En cuarto lugar, se recolecta solo los puntos de mayor contenido informativo, tales como picos, valles, pasajes, puntos singulares, etc.

El método más usado en este trabajo consiste en recolectar la información en forma aleatoria; eventualmente en el caso de estudios de población - se usa el método de los puntos con máximo contenido de información.

Dados los conjuntos $\{X\}$, $\{Y\}$, $\{Z\}$ que definen a $\{D\}$ regular o irregularmente distribuidos , se requiere procesar adecuadamente a $\{D\}$ para producir a $\{S\}$, de tal forma que sea contínua y funcional, es decir que para cada par (X, Y) exista un solo valor de $\{Z\}$.

Para manejar matemáticamente la superficie se pueden utilizar dos tipos diferentes de expresiones. La primera consiste en usar una matriz cuadrada o -

rectangular en donde cada elemento es asignado al e lemento correspondiente del retículo utilizado. El origen de coordenadas se toma en el ángulo superior izquierdo del retículo (ver Fig.).

El segundo tipo de representación viene dado por expresiones analíticas de tipo trigonométrico o polinómico. La utilización de cada técnica en particular dependerá de las premisas hechas inicialmente acerca del comportamiento de los puntos sin información o puntos del retículo {R} .

Una primera hipótesis es establecer que los puntos con información son representativos de los puntos - en su cercania, implicando esto que la superficie - está formada por un conjunto de superficies más pequeñas horizontales a diferentes alturas. Si un - punto cualquiera toma el valor del punto con información más cercano la representacion se llama de - proximidad. Si los puntos toman su valor de acuerdo a sublímites previamente establecidos, la representación es llamada de conformidad o coropletas. En ambos casos estamos asumiendo que la superficie es discontinua.

Si se establece que la superficie es continua esto implica que la primera derivada de la función es también continua y los puntos de retículo $\{R\}$ toman su valor por interpolación lineal a partir de los puntos con información $\{D\}$.

Una tercera posibilidad es suponer que los valores de los puntos de información tienen errores de medición y que por lo tanto la superficie no necesariamente pasa por ellos. Por métodos de mínimos cuadrados se puede ajustar polinomios en X e Y de primero, segundo o mas grados.

La casi totalidad de los experimentos hechos por el autor, utilizan el método matricial para representar la superficie. Esto se debe a que es sumamente cómodo la manipulación de arreglos usando el lenguaje Fortran IV, junto a la elaboracion de los métodos numéricos necesarios. Se ha seleccionado una matriz tipica de 50 filas por 120 columnas. Esto se debe al hecho de que los 6.000 elementos resultantes llenan casi completamente una hoja de salida de impresora (57 líneas de 132 caracteres) permitiendo así hacer adecuadas representaciones de la superficie utilizando la capacidad de sobreimpresión de caracteres para este dispositivo de salida.

2. Algoritmos y Programas.-

2.1. Representación Gráfica de Variables Unidimensionales y sus Análisis Estadísticos

Aunque este trabajo hace hincapié en el análisis y representación de variables en dos y tres dimensiones, a manera de paso introductorio se reseñarán algunos desarrollos para el caso unidimensional. Se ha creido conveniente hacerlo así motivado a que una buena parte de los algorítmos empleados en estos casos elementales son también utilizados en los casos mas avanzados.

2.1.1. Una Variable Simple

Se tiene una función $y = f(x)$, y se desea representar su gráfico mediante la impresora de línea del Computador. La idea consiste en convertir el rango de valores de $f(x)$ en un vector alfanumérico de longitud variable proporcional. Si los valores de $f(x)$ estan en los intervalos $\left[\pm a, \pm b\right]$, b a, para el caso general, $f(x)$ - Min (f (x)) representa una translación del origen de manera que los nuevos valores $f_1(x)$ variaran entre 0 y el valor del rango r que posea f(x). Dividiendo todos los valores de $f_1(x)$ por el máximo valor de $f_1(x)$ se obtiene una nueva función $f_2(x)$ normaliza da al intervalor $[0, 1]$. Basta entonces multiplicar los valores de $f_2(x)$ por el máximo valor deseado del vector alfanu mérico (generalmente 120) para que se obtenga un conjunto de valores enteros que varían proporcionalmente a la función - f(x) original.

El programa mostrado en la página 12 jun to con sus resultados implementa algunas de estas ideas. El eje x vá a lo largo - de la forma continua y el eje y perpendicular a ella.

El programa funciona así: se leen los da tos de entrada $\{D\}$ a partir de los cua les se calcula el promedio S , la varianza SD y la desviación típica DS. Se imprimen estos resultados y se calcula el máximo del conjunto. Las instrucciones:

```
      DØ 11 I = 1,N
   11 B(I) = A(I)/MAX*114+1
```

transforman el conjunto A(I) en otro B(I) que varía entre 1 y 115 en forma proporcional y en correspondencia biunívoca.

```
    DØ 13 J = 1, B(I)
13 DIB(J) = "X"
```

toman el valor actual de B(I) como valor variable del límite superior del rango de de la instrucción DØ, haciendo que el vector alfanumérico DIB(J) sea llenado de los caracteres X, H, O, I etc. para ser sobreimpresos bajo la dirección del lazo más exterior DØ 12 I = 1, N. El gráfico resultante puede verse en la pág.

La primera columna de cifras representa el número de secuencia en que vinieron originalmente los datos. El valor del dato propiamente viene en la segunda columna a cuyo caso aparece una hilera de pequeños rectangulos ennegrecidos cuya longitud es proporcional a la variable.

```
         B 5 7 0 0     F O R T R A N     C O M P I L A T I O N   XIII.11,    TUESDAY, 11/2

$ CARD LIST HOL SINGLE LIMIT 10                                               R  000(
                     SINGLE ILLEGAL CONSTRUCT ON DOLLAR CARD XXXX
FILE  3=OUT,UNIT=PRINTER                                                      R  000(
FILE  1=ALFA12 ,UNIT=READER                                                   R  000(
C                                                                             R  000(
C PROGRAMA PARA DESPLEGAR GRAFICAMENTE UNA VARIABLE UNIDIMENSIONAL.           R  000(
C                                                                             R  000(
                                                          START OF SEGMENT *******
      DIMENSION A(500)                                                        R  000(
      READ(1,1)N                                                              R  000(
    1 FORMAT(I3)                                                              R  001(
      READ (1,2) (A(I),I=1,N)                                                 R  001(
    2 FORMAT (10 F8.1)                                                        R  002?
      WRITE(3,9) A                                                            R  002?
    9 FORMAT (30X,"TEMPERATURA MEDIA"////,(10F10.3))                          R  003(
      S=0                                                                     R  003(
      DO 3 I=1,N                                                              R  003(
      S=S+A(I)                                                                R  004(
    3 CONTINUE                                                                R  004?
      PROM=S/N                                                                R  004?
      SD=0                                                                    R  004(
      DO 4 I=1,N                                                              R  004(
    4 SD=SD+(A(I)-PROM)**2                                                    R  005(
      DS=SQRT(SD/(N-1))                                                       R  005(
      WRITE (3,8) S,PROM,SD,DS                                                R  006(
    8 FORMAT("1",9X,"SUMA DE LOS DATOS="F10.3//10X,"PROMEDIO="F10.2//10X      R  007?
     *,"SUMA DE LAS DESVIACIONES="F10.3//10X,"DESVIACION TIPICA=",F10.3)      R  007?
      REAL MAX                                                                R  007?
      MAX= A(1)                                                               R  007?
      DO 5  I=2,N                                                             R  007(
      IF(A(I).LT.MAX) GO TO 5                                                 R  008(
      MAX= A(I)                                                               R  008(
    5 CONTINUE                                                                R  008(
      WRITE(3,10) MAX                                                         R  008(
   10 FORMAT ("0",9X,"ELEMENTO MAXIMO=",F10.3)                                R  010(
      WRITE(3,18)                                                             R  010(
   18 FORMAT("1",2X,"N",4X,"VARIABLE",30X,"GRAFICO",/3X,"-",4X,               R  010(
     *"--------",30X,"-------",//)                                            R  010(
      DIMENSION B(500),DIB(115)                                               R  010(
      DO 11 I=1,N                                                             R  010(
   11 B(I)=A(I)/MAX *114+1                                                    R  011(
      DO 12 I=1,N                                                             R  011(
      DO 7 K=1,115                                                            R  011(
    7 DIB(K)=" "                                                              R  012(
      DO 13 J=1,B(I)                                                          R  013(
   13 DIB(J)="X"                                                              R  013?
      WRITE(3,14) DIB                                                         R  014(
   14 FORMAT (17X,115A1)                                                      R  015(
      DO 15 J=1,B(I)                                                          R  015(
   15 DIB(J)="H"                                                              R  015(
      WRITE (3,16) DIB                                                        R  016(
   16 FORMAT("+",16X,115A1)                                                   R  017(
      DO 19 J=1,B(I)                                                          R  017(
   19 DIB(J)="O"                                                              R  018(
      WRITE (3,16) DIB                                                        R  018(
      DO 20 J=1,B(I)                                                          R  019(
   20 DIB (J) ="I"                                                            R  020(
      WRITE (3,16) DIB                                                        R  020(
      WRITE(3,17)I,A(I)                                                       R  021(
   17 FORMAT("+",I3"-",F10.2)                                                 R  023(
   12 CONTINUE                                                                R  023(
      STOP                                                                    R  023(
      END
                                                          SEGMENT    1 IS
                                                          SEGMENT    2 IS
```

N	VARIABLE
1-	3,10
2-	9,10
3-	38,70
4-	44,50
5-	8,10
6-	10,30
7-	9,44
8-	3,10
9-	40,10
10-	5,10
11-	48,40
12-	25,30
13-	28,60
14-	31,50
15-	41,60
16-	31,30
17-	31,60
18-	31,30
19-	71,30
20-	21,30
21-	80,30
22-	9,10
23-	48,50
24-	58,80
25-	31,30
26-	100,30
27-	99,60
28-	27,50
29-	48,40
30-	31,30
31-	31,50
32-	40,30
33-	71,40
34-	59,90
35-	38,10
36-	8,10
37-	81,10
38-	28,40
39-	95,50
40-	88,10
41-	26,60
42-	42,20
43-	72,10
44-	60,90
45-	39,40
46-	44,30
47-	82,30
48-	29,10
49-	97,60
50-	79,00
51-	27,70
52-	44,10
53-	73,30
54-	71,90
55-	40,60
56-	31,40
57-	85,40
58-	70,60

GRAFICO

2.1.2 Análisis Estadístico de una Variable Simple

El siguiente programa realiza un análisis estadístico de un vector de datos y genera un conjunto de tablas y gráficos que ayudan a la interpretación rápida y eficiente de los resultados.

Una vez que se introducen la cantidad n de datos a procesar y tres parámetros que definen un histograma, se obtendrá como salida la siguiente información:

a. Conjunto de datos originales

b. El mismo conjunto de datos originales pero ordenados en forma ascendentes (con fines de comprobación)

c. Promedio

d. Desviación típica

e. Suma total del conjunto de datos

f. Números n de datos

A continuación vendrá una tabla a 6 columnas con lo siguiente:

g. Clase

h. Límite superior de cada clase

i. Frecuencia dentro de la clase, es decir, cuantos elementos o datos caen dentro de ella.

j. Porcentaje individual

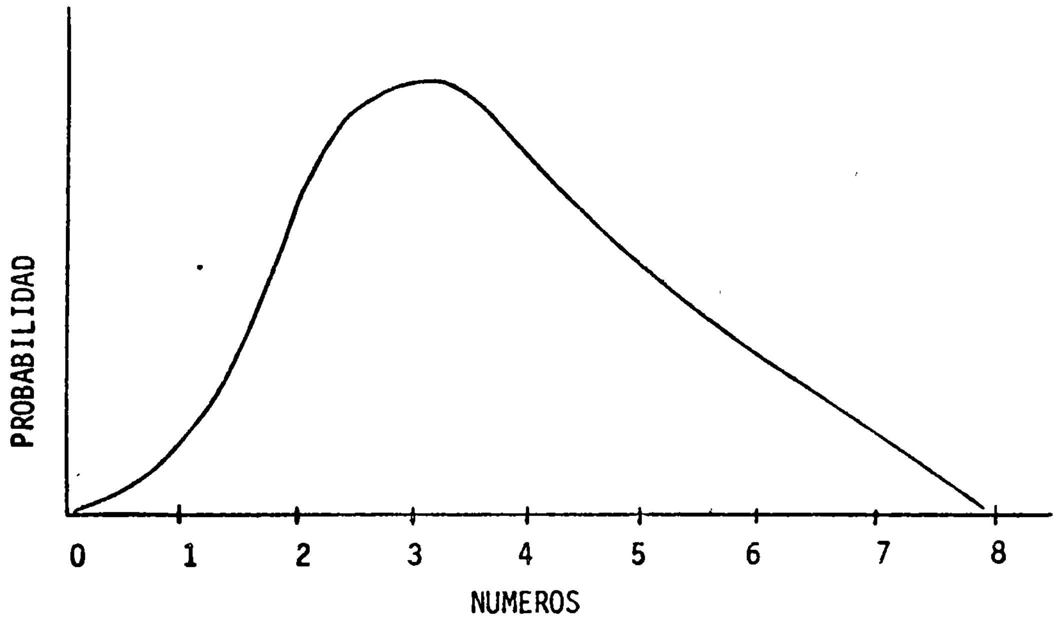
k. Porcentaje acumulativo

l. Resto acumulativo

m. Dibujo del histograma. Girando 90° a la izquierda el papel del Computador, se puede apreciar que hacia la derecha del usuario progresan positivamente las clases especificadas, y hacia arriba está la frecuencia. La altura de cada barra del histograma es igual o proporcional al número de rectángulos negros.

n. Finalmente se dibuja el gráfico de la distribución acumulativa relativa en el rango 0-100. Este gráfico permite (dentro de cierta precisión) el cálculo de la mediana, percentiles, cuartiles, etc.

Si tenemos un conjunto de datos siguiendo cierta distribución probabilística, podemos en este caso representarlos por la cur va contínua de dicha distribución, a saber:



Podemos ahora definir los tres parámetros generales que caracterizan un histograma.

A = Límite superior de la clase más inferior

B = Ancho de la clase

C = Número de clases

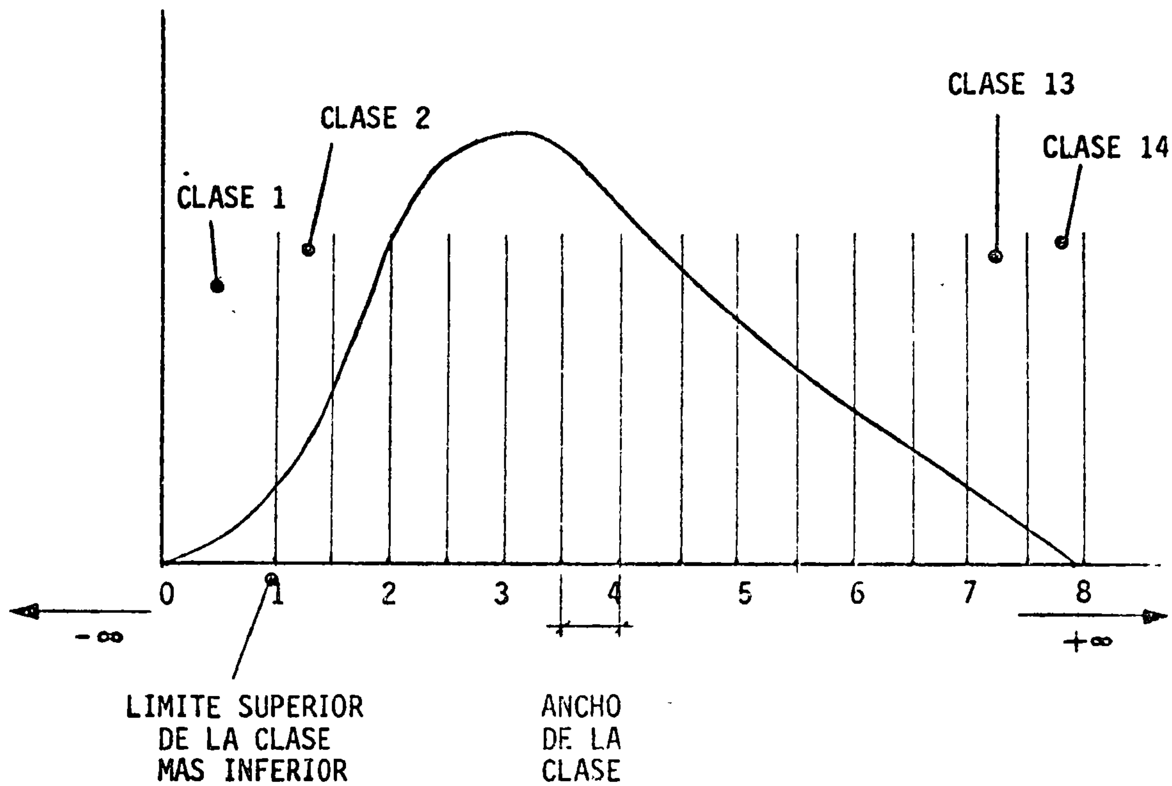
Si para la distribución probabilística de la figura definimos:

A = 1.5

B = 0.5

C = 14

Lo que graficamente será:



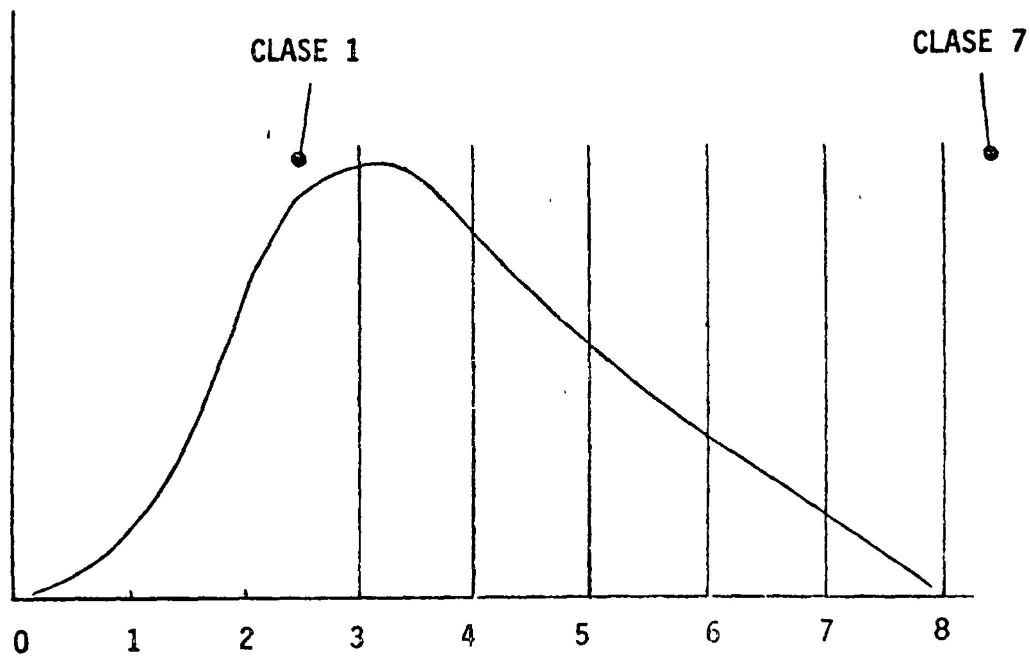
Con esta generalización, cualquier dato que se encuentre en la recta será catalogado en una clase, independientemente de como hayan sido definidos los tres parámetros del análisis; este caso puede presentarse cuando el usuario intenta hacer un análisis estadístico de un conjunto de datos que no conoce bien, en esta situación podrá hacer un tanteo preliminar. usando la distribución anterior podría especificar:

A = 3.

B = 1.

C = 7.

con lo que la situación sería:



Al correr el programa con estos parámetros vemos que el histograma obtenido tendrá la clase 1 un tanto recargada, mientras que - la clase 7 tendrá cero elementos. Un se-- gundo experimento le permitirá colocar me- jor la "trampa de números", logrando, por ejemplo, algo parecido a la fig.

Como ejemplo se desea analizar un conjunto de 536 datos donde el mayor vale 135 y el menor -8.

Se han seleccionado los siguientes grupos de parámetros:

a) A = 43.
   B = 42.
   C = 10.

b) A = -10.
   B = 8.
   C = 25.

c) A = 15.
   B = 7.5
   C = 30.

Puede observarse en la pág. 32 que primero aparece el vector original de datos. Para la pág. 33 corresponde presentar el vector ordenado de menor a mayor; por problemas de espacio no se han mostrado todos los datos (solo hasta el valor 41.). En la página siguiente se encuentran el análisis para el grupo de parámetros a) como puede deducirse, es muy poca la información que

puede conseguirse, dado que el ancho de la clase es excesivamente alto y el límite superior de la clase más inferior también.
Incidentalmente se notará que el eje de frecuencias del histograma fué graduado automáticamente entre 0 y 480 unidades. Esto se debe a que el programa detectó que una de las clases era mayor que 120, escala normal del eje de frecuencias.
Para los casos b) y c) puede notarse que la distribución tiende a ser normal; para el caso de b) la clase 8 es la de mayor frecuencia y está comprendida entre 38. y 46.: Las clases 16 y 17 están desiertas; el gráfico de la distribución acumulativa (al cual se le ha dibujado posteriormente una curva para facilitar la interpretación) muestra por ejemplo que el primer cuartil cae aproximadamente en la clase 6 (22. - 30.) etc.*
Queda al usuario estudiar detenidamente el ejemplo y sacar sus propias conclusiones.

* Si se desea mayor precisión en el cálculo de percentiles debe declararse un ancho de clase más pequeño.

Finalmente el listado del programa aparece de las página 27 a la 31 .

2.1.2.1. Comentarios Generales

El programa tiene una capacidad máxima de 960 datos y puede definir hasta 240 clases diferentes, sin embargo estas capacidades pueden variarse cambiando el dimensionamiento de las variables. El programa explora el vector de frecuencia de largo C para detectar el punto a partir del cual todas las frecuencias valen cero. De esta forma evitamos imprimir las porciones de las tablas y gráficos que valgan cero, esto viene indicado a la salida por un título que dice "el resto de las frecuencias valen cero".

La graduación de los ejes de frecuencia para el histograma es variable y automática. Se -

dispone de cuatro escalas: 120, 240, 480 y 960; sin embargo el usuario no interviene para nada en estas asignaciones.

Los cálculos del promedio, desviación típica, suma de los datos etc., se imprimen una sola vez, de forma tal que no aparecerán a partir de la segunda tabla.

Existe una segunda versión del programa que elimina la necesidad de declarar los tres parámetros A, B y C; y por lo tanto el análisis previo del conjunto de datos para obtener dichos parámetros. En este caso bastará con declarar el ancho de la clase, o sea el valor de B.

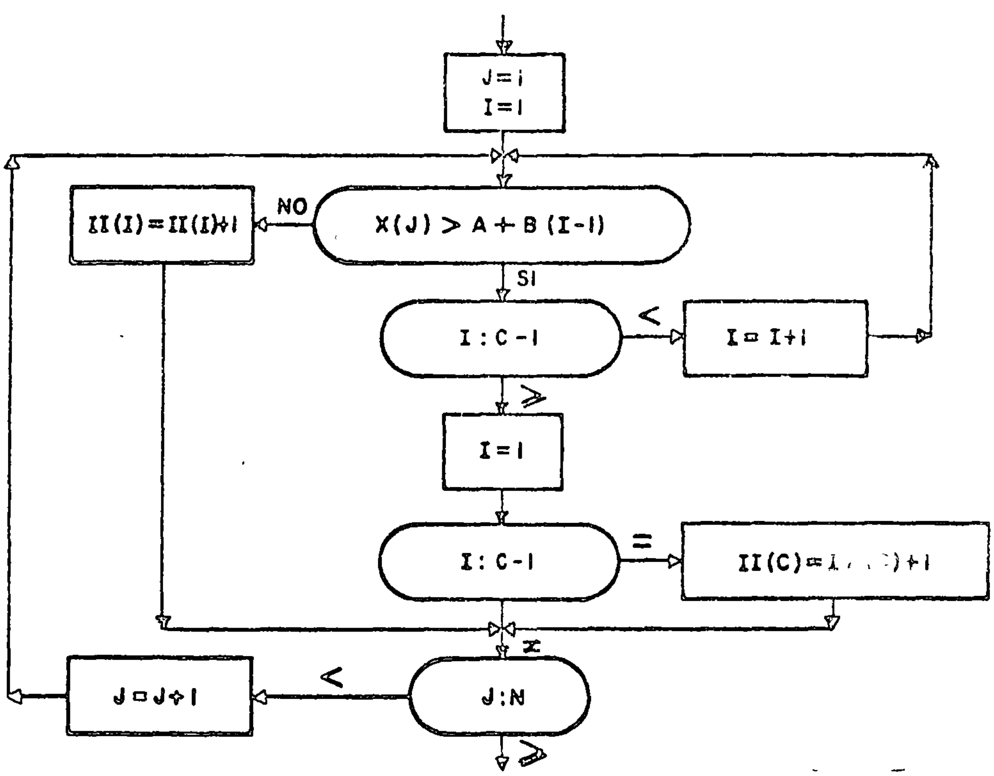
2.1.2.2. Funcionamiento del Programa (Versión II)

Después de leer los datos y el ancho de la clase B, se llama -

la rutina Sort que ordena el - conjunto X(I) de números y los imprime en su nuevo orden. Se cálcula el promedio y la des - viación típica. El rango se - cálcula mediante la operación AN = X(N) - X(1) del vector or denado. El número de clases - viene definido por la expre - sión C = AN/B + 3, es decir, - el rango dividido por el ancho de clase más un margen de tres. Una prueba lógica comprueba - que $3 \leq C \leq 240$. Si no se - cumple la condición el progra- ma automáticamente desiste del cálculo. El límite superior - de la clase mas inferior que - comprende todos los valores en $-\infty$ y el comienzo de la - "trampa de números" se cálcula por la expresión A = X(1) - B* 1.01, es decir el primer ele - mento ordenado menos el ancho de clase. El cálculo del número de elementos que cae en cada clase se hace mediante las instrucciones:

```
DØ 4 J = 1,N
DØ 5 I = 1, C - 1
IF (X (J). GT. (A + B * (I-1))) GØ TØ 5
II = II(I) + 1
GØ  TØ  4
5 CONTINUE
IF(I.EQ.C-1) II(C) = II (C) + 1
4 CONTINUE
```

el diagrama de flujo es:



En líneas generales la lógica del algorítmo establece si el elemento X(J), J = 1, N cae o no dentro de cada intervalo en partícular. La expresión A + B (I-1), I = 1, C - 1 va seleccionando cada intervalo en partícular. El vector - II(I) a través de la instrucción II(I) = II(I) + 1 va acumulando uno a uno, la cantidad de elementos que pertenecen a cada clase. Siendo esta parte del programa la más importante, lo que queda es operar con el vector II(I) -- (que representa la distribu - ción probabilística) para producir la distribución acumulativa, etc. El resto del programa se dedica a llamar un - conjunto de subrutinas que generan los gráficos y las ta-- blas. La primera es CUMUL -

que genera los gráficos de la distribución probabilística. Nótese que el aspecto de cada histograma varía de acuerdo - al ancho de clase escogido B.

```
B 5 7 0 0     F O R T R A N     C O M P I L A T I O N   XIII.11,     FRIDAY, 11/22/74

$ CARD LIST HOL SINGLE LIMIT 10                                           R  0000
              SINGLE ILLEGAL CONSTRUCT ON DOLLAR CARD XXXX
FILE  3=OUT,UNIT=PRINTER                                                  R  0000
FILE  1=TRES    ,UNIT=READER                                              R  0000
                                                START OF SEGMENT **********
      DIMENSION         II(240),PERTOT(240),       SUPE(240),IX(12),      R  0000
     6IIA(240),IZE(240) ,Y(240),X(2500), ZE(2500)                         R  0000
      INTEGER C,SWITCH,CONTAD                                             R  0000
      DATA SWITCH,CONTAD/-1,0/                                            R  0000
      READ(1,1) N                                                         R  0013
    1 FORMAT(I4)                                                          R  0023
      READ (1,91) (X(I),I=1,N)                                            R  0023
   91 FORMAT(8F10.0)                                                      R  0043
      WRITE(3,11) (X(I),I=1,N)                                            R  0043
   11 FORMAT(1H ,48X"VECTOR DE DATOS ORIGINAL"/40X,24("-")///(8X8F13.2))  R  0063
      CALL SORT(X,N)                                                      R  0063
      DO 12 I=1,N                                                         R  0065
   12 SUM=SUM+X(I)                                                        R  0072
      AVE=SUM/N                                                           R  0077
      DO 13 I=1,N                                                         R  0078
      ZE(I)=X(I)-AVE                                                         0083
   13 SUM2=SUM2+(ZE(I)**2)                                                R  0092
      STAND=SQRT(SUM2/N)                                                  R  0097
      AN=X(N)-X(1)                                                        R  0099
C          B=ANCHO DE  LAS CLASES.                                        R  0102
 1000 READ(1,3) B                                                         R  0105
    3 FORMAT(F10.0)                                                       R  0115
      IF(B.EQ.0) C A L L   E X I T                                        R  0115
      C=AN/B+3.                                                           R  0116
      IF(C.GT.240.AND.C.LT.3) GO TO 1000                                  R  0120
      A=X(1)-B*1.01                                                       R  0123
      DO 127 I=1,C                                                        R  0126
  127 II(I)=0                                                             R  0134
      ISUM=0                                                              R  0136
      L=1                                                                 R  0137
      CONTAD=CONTAD+1                                                     R  0137
      WRITE(3,55) CONTAD                                                  R  0139
   55 FORMAT("1",50X,"ESTADISTICAS CALCULADAS","(",I1,")",/51X,23("-")//  R  0149
     */)                                                                  R  0149
      DO 4 J=1,N                                                          R  0149
      DO 5 I=1,C-1                                                        R  0154
      IF(X(J).GT.(A+B*(I-1)))   GO TO 5                                   R  0161
      II(I)=II(I)+1                                                       R  0166
      GO TO 4                                                             R  0171
    5 CONTINUE                                                            R  0173
      IF(I.EQ.C-1) II(C)=II(C)+1                                          R  0173
    4 CONTINUE                                                            R  0179
      DO 10 I=1,C                                                         R  0179
      IF(II(I).EQ.0) GO TO 10                                             R  0184
      L=I                                                                 R  0187
   10 ISUM=ISUM+II(I)                                                     R  0189
      IF(L.LT.C) L=L+1                                                    R  0191
      WRITE(3,63) A,B,L                                                   R  0194
   63 FORMAT( 2X,"LIMITE SUPERIOR DE CLASE MAS INFERIOR A=",F8.2,5X,      R  0209
     *"ANCHO DE LAS CLASES B=",F8.2,5X,"NUMERO DE CLASES C=",I5,          R  0209
     */,2X,39("-"),14X,21("-"),14X,18("-"))                               R  0209
      WRITE(3,14) AVE,STAND,SUM,N                                         R  0209
   14 FORMAT(1H ,                                     /// 1X"PROMEDI      R  0225
     2O DE LOS DATOS",F10.3, 4X,"DESVIACION TIPICA",F10.3,4X,"SUMA DE LO  R  0225
```

```
  6S DATOS",F10.0,4X,"NUMERO DE DATOS",I5///)                      R  022
     II(C)=N-ISUM                                                  R  0225
     DO 16 I=1,C                                                   R  0227
  16 ZE(I)=II(I)                                                   R  0233
     AN=N                                                          R  0236
     DO 15 I=1,C                                                   R  0239
  15 PERTOT(I)=ZE(I)/AN*100.                                       R  0245
     Y(1)=PERTOT(1)                                                R  0251
     DO 17 I=2,C                                                   R  0253
  17 Y(I)=Y(I-1)+PERTOT(I)                                         R  0259
     DO 22 I=1,L                                                   R  0263
  22 SUPE(I)=A+(I-1)*B                                             R  0270
     SUPE(L)=999999.                                               R  0273
     DO 18 I=1,C                                                   R  0277
  18 ZE(I)=100.-Y(I)                                               R  0283
 59  WRITE(3,19)                                                   R  0289
     SWITCH=0                                                      R  0292
  19 FORMAT("+", 7X"CLASE"14X,"LIMITE",6X,"FRECUENCIA",9X,"PORCENTAJE",   R  0294
    39X,"PORCENTAJE"14X,"RESTO"//,25X,"SUPERIOR",7X,"OBSERVADA",9X,       R  0294
    4"INDIVIDUAL",8X,"ACUMULATIVO",8X,"ACUMULATIVO")               R  0294
     DO 20 I=1,L                                                   R  029'
  20 WRITE(3,21) I,SUPE(I),II(I),PERTOT(I),Y(I),ZE(I)              R  030
  21 FORMAT(10X,I3,10X,F10.2,13X,I3, 9X,F10.2, 9X,F10.2, 9X,F10.2)  R  0326
                                                    SEGMENT    2 IS   123
     IF(L.EQ.C) GO TO 551                                          R  0328
     WRITE(3,79)                                                   R  0330
  79 FORMAT(1H0,5X,"EL RESTO DE LAS FRECUENCIAS VALEN CERO"       )   R  0334
 551 WRITE(3,23)                                                   R  0334
  23 FORMAT(1H1)                                                   R  0336
     DO 101 I=1,C                                                  R  0336
 101 IZE(I)=Y(I)                                                   R  0342
     ITEMP=II(1)                                                   R  0347
     DO36 I=2,C                                                    R  0348
     IF(II(I).LT.ITEMP) GO TO 36                                   R  0353
     ITEMP=II(I)                                                   R  0352
  36 CONTINUE                                                      R  0359
     IF(ITEMP.GT.120) GO TO 38                                     R  0359
     DO 29 I=1,12                                                  R  0361
  29 IX(I)=I*10                                                    R  0366
     WRITE(3,28)  CONTAD,IX                                        R  0370
  28 FORMAT(1H0,55X,"HISTOGRAMA","(",I1,")",/56X,10("*")//99X,"FRECUENC   R  0384
    *IA"/7X,12I10,4/"+","CLASE")                                   R  0384
     DO 25 I=1,120                                                 R  0384
  25 Y(I)="*"                                                      R  0390
     DO 26 I=10,120,10                                             R  0394
  26 Y(I)="+"                                                      R  040
     WRITE(3,27)  (Y(I),I=1,120)                                   R  0409-
  27 FORMAT(6X,120A1/)                                             R  0421
     CALL HISTOG(II,L)                                             R  0421
     WRITE(3,31) (Y(I),I=1,120)                                    R  0422
     WRITE(3,30) IX                                                R  0440
  30 FORMAT("+",6X,12I10/99X, "FRECUENCIA")                        R  0451
  31 FORMAT("+",5X,120A1/)                                         R  0451
     CALL CUMUL(IZE,L)                                             R  0451
     GO TO 1000                                                    R  0452
  38 IF(ITEMP.GT.240) GO TO 52                                     R  0454
     DO 42 I=1,C                                                   R  0456
     IIA(I)=II(I)                                                  R  0461
  42 IIA(I)=IIA(I)/2                                               R  0465
     DO 44 I=1,12                                                  R  0468
  44 IX(I)=I*20                                                    R  0474
     WRITE(3,28) IX                                                R  0476
```

```
      DO 46 I=1,120                                        R 0487
   46 Y(I)="-"                                             R 0493
      DO 48 I=10,120,10                                    R 0497
   48 Y(I)="+"                                             R 0503
      WRITE(3,27) (Y(I),I=1,120)                           R 0507
      CALL HISTOG(IIA,L)                                   R 0524
      WRITE(3,31) (Y(I),I=1,120)                           R 0525
      WRITE(3,30) IX                                       R 0543
      CALL CUMUL(IZE,L)                                    R 0553
      GO TO 1000                                           R 0555
   52 IF(ITEMP.GT.480) GO TO 60                            R 0556
      DO 54 I=1,C                                          R 0556
      IIA(I)=II(I)                                         R 0563
   54 IIA(I)=IIA(I)/4                                      R 0567
      DO 49 I=1,12                                         R 0570
   49 IX(I)=I*40                                           R 0576
      WRITE(3,28) IX                                       R 0576
      DO 56 I=1,120                                        R 0589
   56 Y(I)="-"                                             R 0595
      DO 58 I=10,120,10                                    R 0599
   58 Y(I)="+"                                             R 0605
      WRITE(3,27) (Y(I),I=1,120)                           R 0609
      CALL HISTOG(IIA,L)                                   R 0626
      WRITE(3,31) (Y(I),I=1,120)                           R 0627
      WRITE(3,30) IX                                       R 0645
      CALL CUMUL(IZE,L)                                    R 0655
      GO TO 1000                                           R 0657
   60 IF(ITEMP.GT.960) GO TO 70                            R 0658
      DO 62 I=1,C                                          R 0660
      IIA(I)=II(I)                                         R 0665
   62 IIA(I)=IIA(I)/8                                      R 0669
      DO 64 I=1,12                                         R 0672
   64 IX(I)=I*80                                           R 0676
      WRITE(3,28) IX                                       R 0680
      DO 66 I=1,120                                        R 0691
   66 Y(I)="-"                                             R 0697
      DO 68 I=10,120,10                                    R 0701
   68 Y(I)="+"                                             R 0707
      WRITE(3,27)  (Y(I),I=1,120)                          R 0711
      CALL HISTOG(IIA,L)                                   R 0728
      WRITE(3,31) (Y(I),I=1,120)                           R 0729
      WRITE(3,30) IX                                       R 0747
      CALL CUMUL(IZE,L)                                    R 0757
      GO TO 1000                                           R 0759
   70 WRITE(3,72)                                          R 0760
   72 FORMAT("0",10X,"CLASE EXCEDE 960 ELEMENTOS")         R 0764
      GO TO 1000                                           R 0764
      END                                                  R 0764
                                              SEGMENT    1 IS  794
                                        START OF SEGMENT **********
      SUBROUTINE SORT(A,N)                                 R 0000
      DIMENSION A(960)                                     R 0000
      K=N-1                                                R 0000
      DO 6 J=1,N-2                                         R 0001
      DO 5 I=1,K                                           R 0007
      IF(A(I).LT.A(I+1)) GO TO 5                           R 0013
      TEMP=A(I+1)                                          R 0019
      A(I+1)=A(I)                                          R 0021
      A(I)=TEMP                                            R 0025
    5 CONTINUE                                             R 0026
      K=K-1                                                R 0026
```

```
    6 CONTINUE
      WRITE(3,10) (A(I),I=1,N)
   10 FORMAT(1H1,48X,"VECTOR DE DATOS ORDENADOS"/49X,25("-"),///(8X,8F1
     1.5))
      RETURN
      END
                                                                START C
      SUBROUTINE CUMUL(IE,C)
      DIMENSION X(100),IX(10),LINE(100)
      INTEGER C,IE(240)
      DO 3 I=1,100
    3 X(I)="-"
      DO 1 I=10,100,10
    1 X(I)="+"
      DO 30 I=1,10
   30 IX(I)=I*10
      WRITE(3,2) IX
    2 FORMAT("1",47X,"DISTRIBUCION ACUMULATIVA"/48X,24("-")///50X,"PORC
     1ENTAJE"/7X,10I10/ " CLASE")
      WRITE(3,4) (X(I),I=1,100)
    4 FORMAT("+",5X,100A1,/)
      DO 6 I=1,C
      DO 5 J=1,100
    5 LINE(J)=" "
      IF(IE(I).EQ.0)  GO TO 13
      DO 7 J=1,IE(I)
    7 LINE(J)="H"
      WRITE(3,8) I,LINE
    8 FORMAT("+",I3,1X,"I",100A1)
      DO 9 J=1,IE(I)
    9 LINE(J)="X"
      WRITE(3,10)  LINE
   10 FORMAT("+",5X,100A1)
      DO 11 J=1,IE(I)
   11 LINE(J)="O"
      WRITE(3,10) LINE
      DO 12 J=1,IE(I)
   12 LINE(J)="I"
      WRITE(3,10)  LINE
      GO TO 21
   13 WRITE(3,14)  I
   14 FORMAT("+",I3,1X,"I")
   21 WRITE(3,15)
   15 FORMAT(" ",4X,"+",/)

    6 CONTINUE
      WRITE(3,4) (X(I),I=1,100)
      WRITE(3,16)IX
   16 FORMAT("+",6X,10I10,/50X,"PORCENTAJE" )
      RETURN
      END
                                                                START C
      SUBROUTINE HISTOG(II,IC)
      DIMENSION II(IC),LINE(120)
      DATA ACHE,EQUIS,OH,AI/"H","X","O","I"/
      DATA BLANC/" "/
      DO 2 I=1,IC
      DO 7 K=1,120
    7 LINE(K)=BLANC
      IF(II(I).EQ.0)GO TO 12
```

```
    3 LINE(J)=ACHE                                      R  0054
      WRITE(3,10) I,LINE                                R  0055
   10 FORMAT("+",I3,1X,"I",120A1)                       R  0072
      DO 4 J=1,II(I)                                    R  0072
    4 LINE(J)=EQUIS                                     R  0080
      WRITE(3,11)LINE                                   R  0082
   11 FORMAT("+",5X,120A1)                              R  0096
      DO 5 J=1,II(I)                                    R  0096
    5 LINE(J)=OH                                        R  0104
      WRITE(3,11)LINE                                   R  0106
      DO 6 J=1,II(I)                                    R  0119
    6 LINE(J)=AI                                        R  0127
      WRITE(3,11)LINE                                   R  0129
      GO TO 21                                          R  0142
   12 WRITE(3,13)I                                      R  0144
   13 FORMAT("+",I3,1X,"I")                             R  0154
   21 WRITE(3,14)                                       R  0154
   14 FORMAT(" ",6X,"+",/)                              R  0156
    2 CONTINUE                                          R  0156
      RETURN                                            R  0156
      END                                               R  0161
                                          SEGMENT    6 IS   175 LONG
                                          SEGMENT    7 IS    36 LONG
                                          SEGMENT    8 IS   138 LONG
                                          SEGMENT    9 IS    17 LONG
                                          SEGMENT   10 IS     6 LONG
                                 START OF SEGMENT **********   11
                                          SEGMENT   11 IS    13 LONG

NUMBER OF SYNTAX ERRORS DETECTED = 0.

PRT SIZE = 108;  TOTAL SEGMENT SIZE = 1731 WORDS;  DISK SIZE = 72 SEGS;  NO. PRGM. SEGS = 15.

ESTIMATED CORE STORAGE REQUIREMENT = 6976 WORDS;  COMPILATION TIME = 4 MIN, 7 SECS;  NO. CARDS = 255.
```

VECTOR DE DATOS ORIGINAL

75.00	85.00	86.00	96.00	85.00	86.00	75.00	74.00
0.00	2.00	45.00	56.00	49.00	53.00	86.00	23.00
53.00	53.00	65.00	54.00	42.00	52.00	53.00	57.00
12.00	2.30	58.00	69.00	68.00	57.00	57.00	68.00
98.00	98.00	98.00	98.00	69.00	57.00	57.00	58.00
10.00	10.00	11.00	30.00	32.00	68.00	69.00	97.00
52.00	52.00	53.00	62.00	61.00	64.00	65.00	52.00
3.00	2.00	6.00	9.00	8.00	53.00	42.00	53.00
42.00	43.00	41.00	45.00	52.00	53.00	62.00	41.00
45.00	56.00	75.00	86.00	96.00	42.00	53.00	21.00
20.00	20.00	30.00	23.00	12.00	23.00	45.00	56.00
43.00	42.00	52.00	75.00	85.00	96.00	78.00	78.00
47.00	58.00	58.00	69.00	53.00	52.00	53.00	42.00
12.00	12.00	53.00	10.00	43.00	135.00	47.00	124.00
12.00	53.00	56.00	20.00	12.00	30.00	52.00	63.00
86.00	86.00	95.00	95.00	43.00	79.00	46.00	56.00
10.00	10.00	20.00	30.00	40.00	50.00	60.00	70.00
80.00	80.00	90.00	56.00	24.00	35.00	36.00	36.00
37.00	38.00	34.00	32.00	31.00	36.00	38.00	39.00
30.00	35.00	65.00	68.00	69.00	32.00	32.00	32.00
75.00	96.00	53.00	52.00	53.00	42.00	52.00	63.00
41.00	20.00	79.00	14.00	25.00	45.00	41.00	17.00
41.00	20.00	12.00	14.00	25.00	45.00	41.00	17.00
27.00	28.00	29.00	38.00	65.00	98.00	97.60	42.00
78.00	96.00	32.00	20.00	10.00	40.00	13.00	97.00
74.00	32.00	80.00	73.00	94.00	52.00	32.00	32.00
65.00	63.00	30.00	42.00	12.00	86.00	95.00	75.00
30.00	90.00	80.00	41.00	65.00	98.00	74.00	54.00
21.00	68.00	74.00	54.00	20.00	21.00	65.00	98.00
32.00	78.00	89.00	56.00	23.00	23.00	10.00	20.00
80.00	90.00	56.00	41.00	52.00	85.00	52.00	20.00
45.00	96.00	63.00	43.00	28.00	39.00	57.00	38.00
4.00	2.00	1.00	6.00	8.00	9.00	6.00	3.00
12.00	12.00	53.00	10.00	43.00	135.00	99.00	124.00
0.00	2.00	45.00	56.00	42.00	53.00	86.00	23.00
41.00	52.00	86.00	23.00	42.00	69.00	43.00	78.00
98.00	87.00	94.00	32.00	21.00	98.00	56.00	43.00
12.00	12.00	53.00	10.00	43.00	135.00	47.00	56.00
0.00	2.00	45.00	56.00	49.00	53.00	86.00	23.00
-1.00	-2.00	-3.00	-3.00	-4.00	-8.00	-1.00	-5.00
55.00	56.00	45.00	53.00	42.00	65.00	68.00	52.00
123.00	125.00	126.00	127.00	125.00	123.00	129.00	126.00
12.00	12.00	53.00	10.00	43.00	42.00	47.00	56.00
65.00	86.00	63.00	52.00	52.00	41.00	53.00	63.00
25.00	86.00	63.00	20.00	35.00	68.00	96.00	54.00
12.00	12.00	53.00	10.00	43.00	135.00	99.00	1.40
12.00	12.00	53.00	10.00	43.00	135.00	99.00	1.40
47.00	68.00	35.00	25.00	25.00	90.00	58.00	35.00
99.00	63.00	52.00	63.00	85.00	75.00	86.00	96.00
18.00	38.00	56.00	35.00	69.00	25.00	86.00	35.00
86.00	96.00	83.00	42.00	94.00	63.00	30.00	42.00
75.00	63.00	42.00	53.00	15.00	69.00	58.00	38.00
47.00	69.00	38.00	69.00	17.00	28.00	27.00	38.00
41.00	42.00	43.00	44.00	45.00	46.00	42.00	41.00

VEC[illegible] DE DATOS ORDENADOS

-8.00000	-5.00000	-4.00000	-3.00000	-3.00000	-2.00000	-1.00000	-1.00000
0.00000	0.00000	0.00000	1.00000	1.40000	1.40000	2.00000	2.00000
2.00000	2.00000	2.00000	2.30000	3.00000	3.00000	4.00000	6.00000
6.00000	6.00000	8.00000	8.00000	9.00000	9.00000	10.00000	10.00000
10.00000	10.00000	10.00000	10.00000	10.00000	10.00000	10.00000	10.00000
10.00000	10.00000	11.00000	12.00000	12.00000	12.00000	12.00000	12.00000
12.00000	12.00000	12.00000	12.00000	12.00000	12.00000	12.00000	12.00000
12.00000	12.00000	12.00000	12.00000	12.00000	12.00000	13.00000	14.00000
14.00000	14.00000	14.00000	14.00000	14.00000	15.00000	17.00000	17.00000
17.00000	17.00000	17.00000	17.00000	17.00000	18.00000	20.00000	20.00000
20.00000	20.00000	20.00000	20.00000	20.00000	20.00000	20.00000	20.00000
20.00000	20.00000	20.00000	20.00000	20.00000	20.00000	20.00000	21.00000
21.00000	21.00000	21.00000	23.00000	23.00000	23.00000	23.00000	23.00000
23.00000	23.00000	23.00000	23.00000	23.00000	23.00000	23.00000	24.00000
25.00000	25.00000	25.00000	25.00000	25.00000	25.00000	25.00000	25.00000
25.00000	25.00000	25.00000	25.00000	27.00000	27.00000	28.00000	28.00000
28.00000	29.00000	30.00000	30.00000	30.00000	30.00000	30.00000	30.00000
30.00000	30.00000	31.00000	32.00000	32.00000	32.00000	32.00000	32.00000
32.00000	32.00000	32.00000	32.00000	32.00000	32.00000	34.00000	35.00000
35.00000	35.00000	35.00000	35.00000	35.00000	35.00000	36.00000	36.00000
36.00000	36.00000	36.00000	36.00000	37.00000	38.00000	38.00000	38.00000
38.00000	38.00000	38.00000	38.00000	38.00000	39.00000	39.00000	40.00000
40.00000	41.00000	41.00000	41.00000	41.00000	41.00000	41.00000	41.00000
41.00000	41.00000	41.00000	41.00000	41.00000	41.00000	41.00000	41.00000
41.00000	41.00000	41.00000	41.00000	41.00000	41.00000	41.00000	41.00000
41.00000	42.00000	42.00000	42.00000	42.00000	42.00000	42.00000	42.00000
42.00000	42.00000	42.00000	42.00000	42.00000	42.00000	42.00000	42.00000
42.00000	42.00000	42.00000	42.00000	42.00000	42.00000	42.00000	42.00000
42.00000	42.00000	42.00000	42.00000	42.00000	42.00000	42.00000	43.00000
43.00000	43.00000	43.00000	43.00000	43.00000	43.00000	43.00000	43.00000
43.00000	43.00000	43.00000	43.00000	43.00000	43.00000	44.00000	45.00000
45.00000	45.00000	45.00000	45.00000	45.00000	45.00000	45.00000	45.00000
45.00000	45.00000	45.00000	45.00000	45.00000	45.00000	45.00000	45.00000
46.00000	46.00000	46.00000	46.80000	47.00000	47.00000	47.00000	47.00000
47.00000	47.00000	47.00000	47.00000	48.00000	49.00000	49.00000	49.00000
49.50000	50.00000	50.00000	50.00000	50.00000	52.00000	52.00000	52.00000
52.00000	52.00000	52.00000	52.00000	52.00000	52.00000	52.00000	52.00000
52.00000	52.00000	52.00000	52.00000	52.00000	52.00000	52.00000	52.00000
52.00000	52.00000	52.00000	52.00000	53.00000	53.00000	53.00000	53.00000
53.00000	53.00000	53.00000	53.00000	53.00000	53.00000	53.00000	53.00000
53.00000	53.00000	53.00000	53.00000	53.00000	53.00000	53.00000	52.00000
53.00000	53.00000	53.00000	53.00000	53.00000	53.00000	53.00000	53.00000
53.00000	54.00000	54.00000	54.00000	54.00000	54.00000	55.00000	56.00000
56.00000	56.00000	56.00000	56.00000	56.00000	56.00000	56.00000	56.00000
56.00000	56.00000	56.00000	56.00000	56.00000	56.00000	56.00000	56.00000
57.00000	57.00000	57.00000	57.00000	57.00000	57.00000	58.00000	58.00000
58.00000	58.00000	58.00000	58.00000	60.00000	61.00000	62.00000	62.00000
63.00000	63.00000	63.00000	63.00000	63.00000	63.00000	63.00000	63.00000
63.00000	63.00000	63.00000	63.00000	63.00000	63.00000	63.00000	63.00000
63.00000	63.00000	64.00000	65.00000	65.00000	65.00000	65.00000	65.00000
65.00000	65.00000	65.00000	65.00000	68.00000	68.00000	68.00000	68.00000
68.00000	68.00000	68.00000	68.00000	69.00000	69.00000	69.00000	69.00000
69.00000	69.00000	69.00000	69.00000	69.00000	69.00000	70.00000	73.00000
74.00000	74.00000	74.00000	74.00000	75.00000	75.00000	75.00000	75.00000
75.00000	75.00000	75.00000	75.00000	75.00000	77.00000	78.00000	78.000

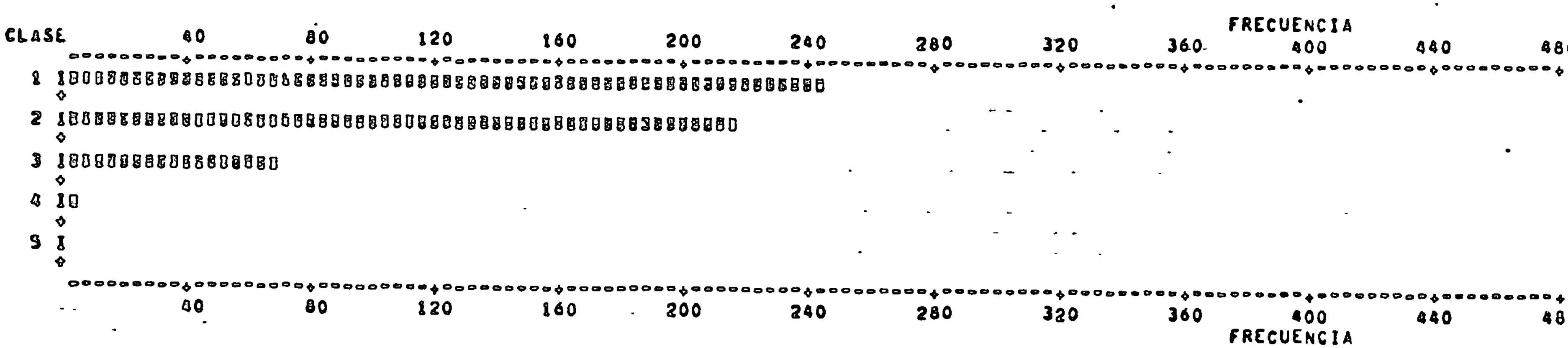
ESTADISTICAS CALCULADAS

PROMEDIO DE LOS DATOS 49,825 DESVIACION TIPICA 28,888 SUMA DE LOS DATOS 26706. NUMERO DE DATOS 536

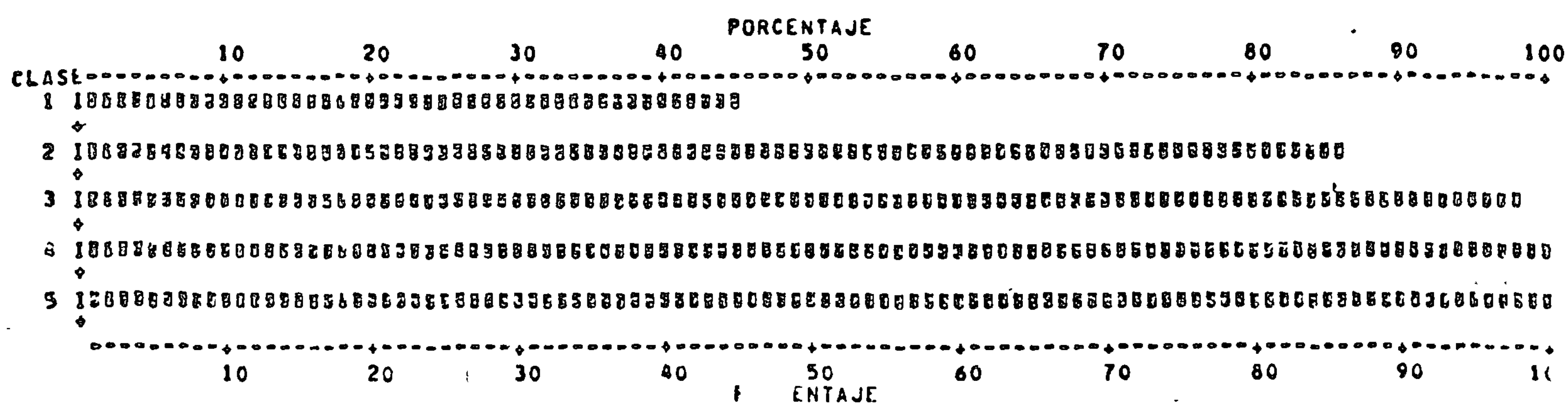
CLASE	LIMITE SUPERIOR	FRECUENCIA OBSERVADA	PORCENTAJE INDIVIDUAL	PORCENTAJE ACUMULATIVO	RESTO ACUMULATIVO
1	43.0	246	45.90	45.90	54.10
2	85.0	216	40.30	86.19	13.81
3	127.0	68	12.69	98.88	1.12
4	169.0	6	1.12	100.00	0.00
5	999999.0	0	0.00	100.00	0.00

EL RESTO DE LAS FRECUENCIAS VALEN CERO

HISTOGRAMA



DISTRIBUCION ACUMULATIVA

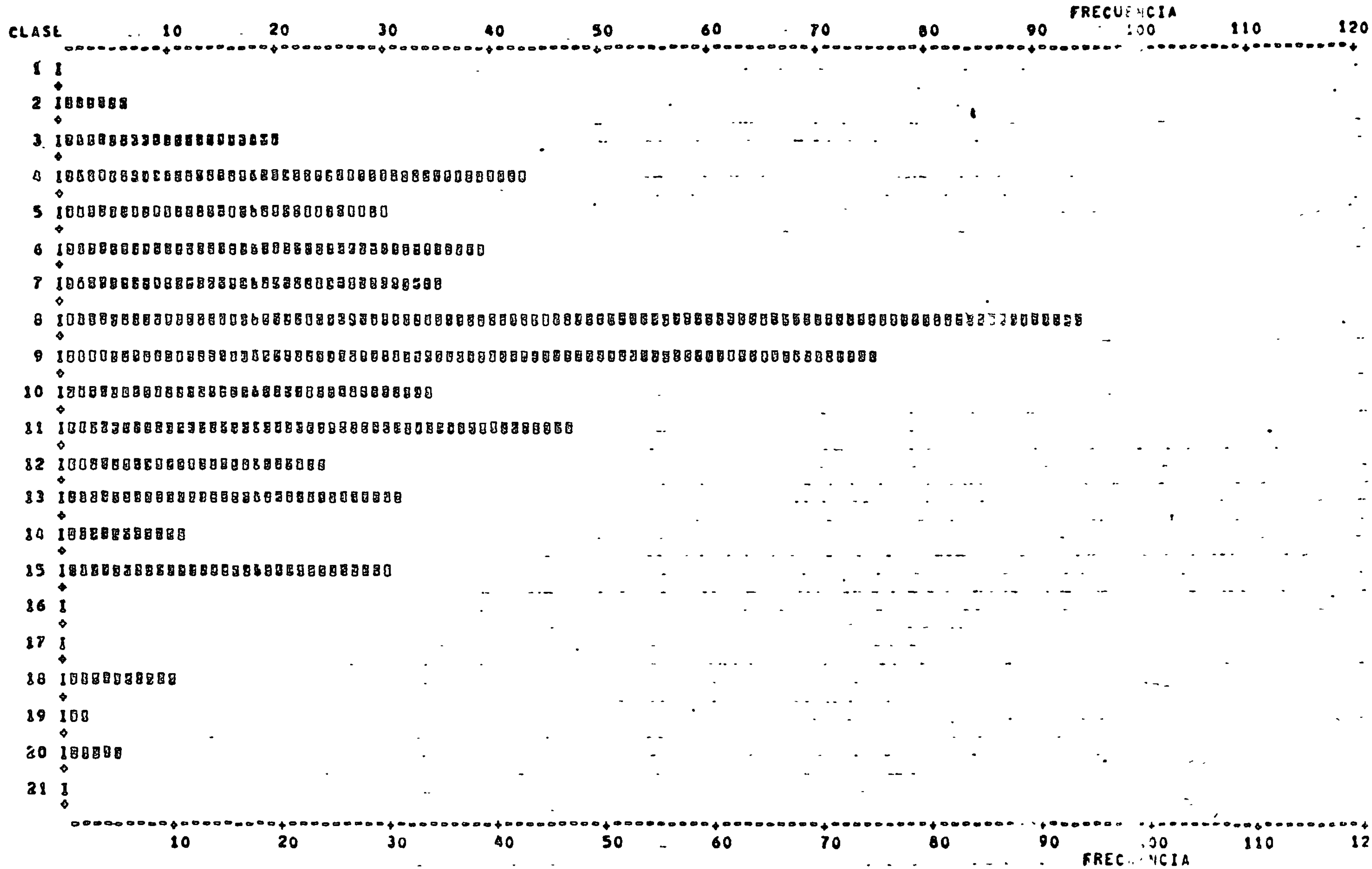


ESTADISTICAS CALCULADAS(1)

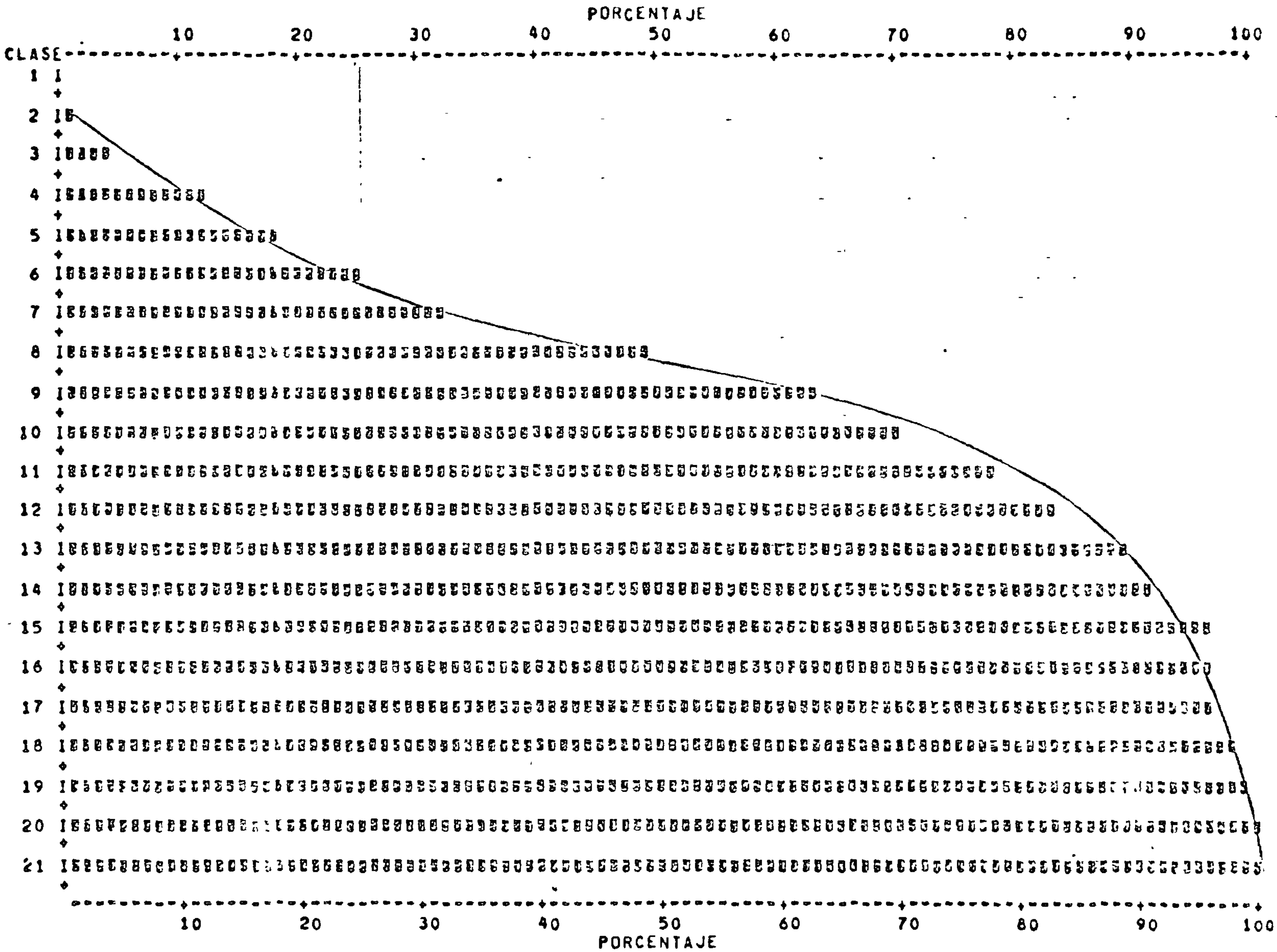
CLASE	LIMITE SUPERIOR	FRECUENCIA OBSERVADA	PORCENTAJE INDIVIDUAL	PORCENTAJE ACUMULATIVO	RESTO ACUMULATIVO
1	-10.0	0	0.00	0.00	100.00
2	-2.0	6	1.12	1.12	98.88
3	6.0	20	3.73	4.85	95.15
4	14.0	43	8.02	12.87	87.13
5	22.0	30	5.60	18.47	81.53
6	30.0	39	7.28	25.75	74.25
7	38.0	35	6.53	32.28	67.72
8	46.0	94	17.54	49.81	50.19
9	54.0	75	13.99	63.81	36.19
10	62.0	34	6.34	70.15	29.85
11	70.0	47	8.77	78.92	21.08
12	78.0	24	4.48	83.40	16.60
13	86.0	31	5.78	89.18	10.82
14	94.0	11	2.05	91.23	8.77
15	102.0	30	5.60	96.83	3.17
16	110.0	0	0.00	96.83	3.17
17	118.0	0	0.00	96.83	3.17
18	126.0	10	1.87	98.69	1.31
19	134.0	2	0.37	99.07	0.93
20	142.0	5	0.93	100.00	-0.00
21	999999.0	0	0.00	100.00	-0.00

EL RESTO DE LAS FRECUENCIAS VALEN CERO

HISTOGRAMA


FRECUENCIA
CLASE
10 20 30 40 50 60 70 80 90 100 110 120
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
10 20 30 40 50 60 70 80 90 100 110 120
FRECUENCIA

DISTRIBUCION ACUMULATI



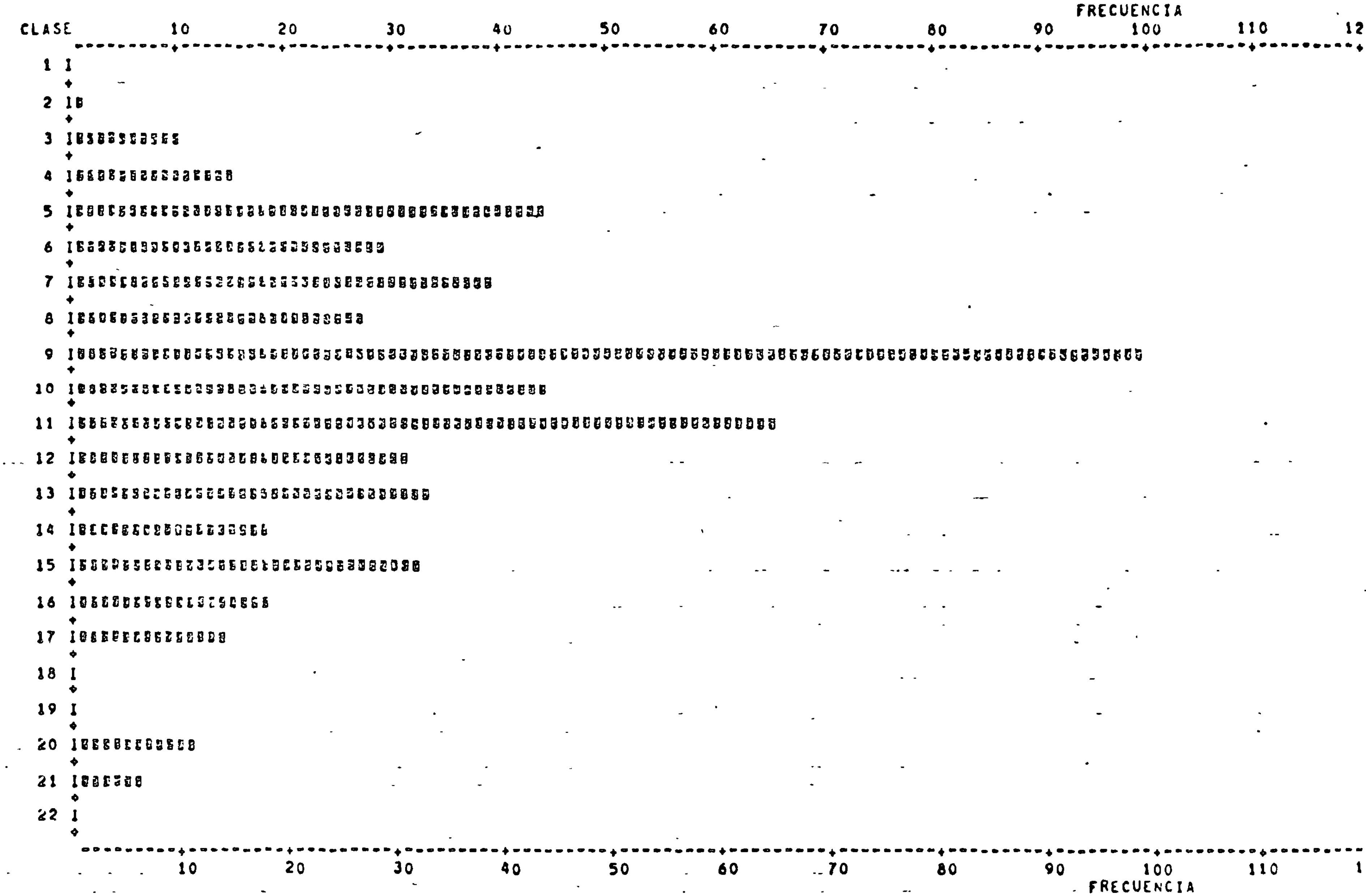
ESTADISTICAS CALCULADAS

PROMEDIO DE LOS DATOS 49.[illegible]25 DESVIACION TIPICA 28.888 SUMA DE LOS DATOS 26706. NUMERO DE DATOS 536

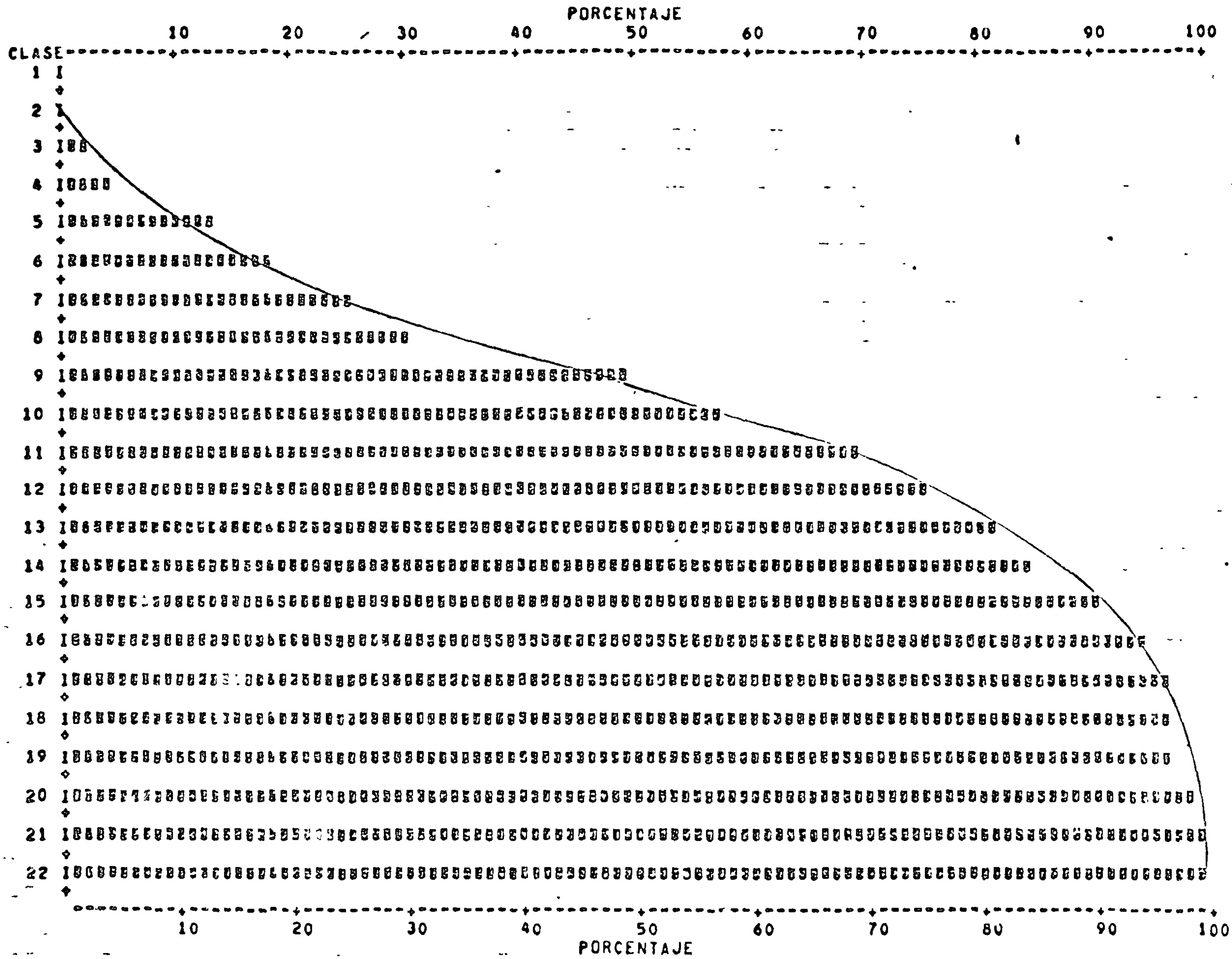
CLASE	LIMITE SUPERIOR	FRECUENCIA OBSERVADA	PORCENTAJE INDIVIDUAL	PORCENTAJE ACUMULATIVO	RESTO ACUMULATIVO
1	-15.0	0	0.00	0.00	100.00
2	-7.5	1	0.19	0.19	99.81
3	0.0	10	1.87	2.05	97.95
4	7.5	15	2.80	4.85	95.15
5	15.0	44	8.21	13.06	86.94
6	22.5	29	5.41	18.47	81.53
7	30.0	39	7.28	25.75	74.25
8	37.5	27	5.04	30.78	69.22
9	45.0	99	18.47	49.25	50.75
10	52.5	44	8.21	57.46	42.54
11	60.0	65	12.13	69.59	30.41
12	67.5	31	5.78	75.37	24.63
13	75.0	33	6.16	81.53	18.47
14	82.5	18	3.36	84.89	15.11
15	90.0	32	5.97	90.86	9.14
16	97.5	18	3.36	94.22	5.78
17	105.0	14	2.61	96.83	3.17
18	112.5	0	0.00	96.83	3.17
19	120.0	0	0.00	96.83	3.17
20	127.5	11	2.05	98.88	1.12
21	135.0	6	1.12	100.00	0.00
22	999999.0	0	0.00	100.00	0.00

EL RESTO DE LAS FRECUENCIAS VALEN CERO

HISTOGRAMA


FRECUENCIA
CLASE
10 20 30 40 50 60 70 80 90 100 110 12
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
10 20 30 40 50 60 70 80 90 100 110
FRECUENCIA

DISTRIBUCION ACUMULATIVA



2.2. Generación de la Matriz de Superficie

Como se planteó anteriormente, se desea obtener {S} a partir de {R} y {D} interpolando y extrapolando los valores de {R} a partir de los de {D} . Para esto se utilizará un caso partícular de la - .llamada Ecuación del Calor en dos dimensiones:

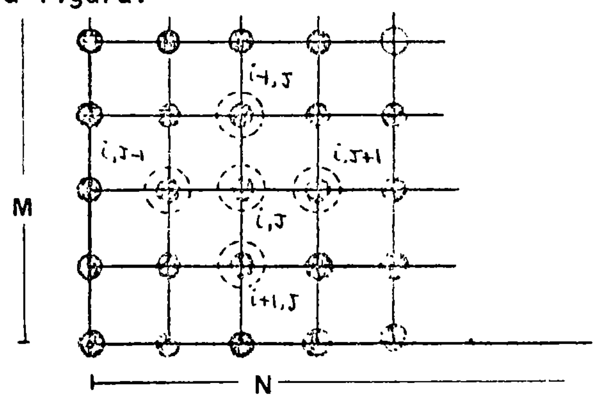
$$\frac{\partial^2 U}{\partial t^2} = \frac{K}{cp}\left[\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2}\right] \qquad (2.1)$$

Donde (2.1) describe la rata de cambio promedio de temperatura en un elemento diferencial de una lámina delgada de calor específico c y peso específico p. Resulta curioso que una formulación matemática desarrollada para estudiar el flujo dinámico del calor, sirva también para estudiar variables - de tipo social. Como estamos interesados en condiciones de estado estacionario (es decir, que aspecto tendrá la superficie de potencial de población en un momento determinado de una región dada) tendremos:

$\frac{\partial^2 U}{t^2} = 0$ y (2.1) queda reducida a:

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} = 0 \qquad (2.2)$$

y que recibe el nombre de Ecuación de Laplace. El desarrollo de la solución numérica de esta ecuación ha sido tratado numerosas veces en la literatura especializada y no se dará aquí (4). Si suponemos una región o una lámina de material delgado dividida en pequeños elementos finitos ocupando las intersecciones de un retículo cuadrado como se ilustra en la figura.



Un elemento cualquiera (i,j) está rodeado por cuatro elementos de coordenadas (i+1,j), (i-1,j), (i,j+1) y (i,j-1). La solución de la ecuación (2.2) es:

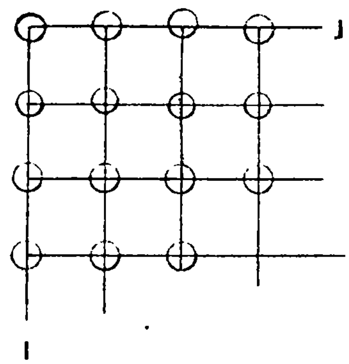
$$U_{i,j} = \frac{U_{i+1,j} + U_{i-1,j} + U_{i,j+1} + U_{i,j-1}}{4} \qquad \begin{matrix} 1 < i < M \\ 1 < j < N \end{matrix} \qquad (2.3)$$

Este resultado pone de manifiesto un hecho bastante intuitivo que establece que el potencial de un pun-o cualquiera dentro de la lámina es igual al promedio de los cuatro puntos circundantes. Normalmente se establecen unas condiciones de borde: si el fenómeno que intentamos visualizar es un flujo de calor o un campo eléctrico los bordes de la lámina pueden tomar valores fijos y los límites del proceso serán:

$$I = 1, M-1 \quad ; \quad I = \text{Subíndice de fila}$$
$$J = 2, N-1 \quad ; \quad J = \text{Subíndice de columna} \qquad (2.4)$$

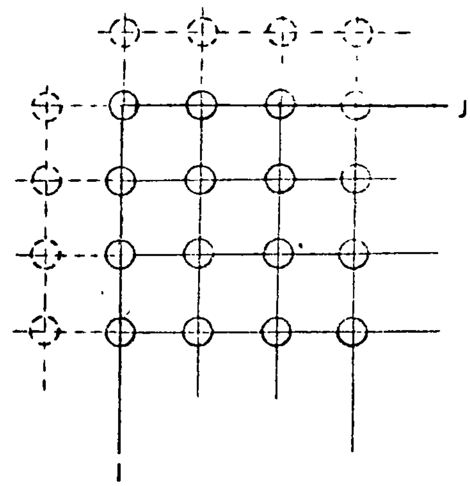
Como estamos interesados en generar la superficie de potencial producida por variables de tipo social, se deberá tomar en cuenta el caso en que los bordes de una región geográfica esten en contacto con un mar o un lago donde la población vale necesariamente cero. Para el resto de fronteras o límites la superficie de potencial debe "flotar" libremente; esta condición se denomina "bordes aislados". En las dos figuras siguientes se ilustran ambas situaciones:

Bordes con límites fijos



Caso A

Bordes con límites variables



Caso B

En el caso a) los bordes siempre permaneceran en un valor constante. Para el caso b) supondremos que existen los elementos ficticios indicados por las líneas punteadas. Estos van a tomar siempre el valor más nuevo calculado para el elemento simétrico respecto al borde.

Para implementar estas ideas en un programa de Computador se asigna a cada elemento finito de un retículo de tamaño M x N el, elemento respectivo de una matriz numérica de M x N filas y columnas. Aplicando la ecuación (2.3) a cada uno de los elementos $\{R\}$ interiores entre los límites establecidos por las relaciones (2.4) (se excluye $\{D\}$) se obtiene finalmente la expresión matricial $\{S\}$.

El autor ha encontrado que el número de iteracciones necesarias para generar una superficie que comunique un mínimo de información es por lo general menor que las indispensables para la solución numérica "exacta" de la ecuación de Laplace. De todas formas, como estamos en presencia de un algorítmo infinito, éste solo produce una respuesta aproximada, habiendo un acercamiento de tipo asintótico. Normalmente, para alcanzar una respuesta bastante exacta se requieren de 400 a 500 "barridos" de la matriz. Para el caso geográfico bastan 100 iteraciones.

2.3. Procesamiento de la Superficie

Una vez obtenida la matriz, debe ésta ser sometida a varios procesos algorítmicos con el fin de eliminar o aumentar ciertos detalles, desaparecer pequeñas variaciones producidas por incertidumbre en los datos de entrada, etc.

Como ejemplo solo se expondrá aquí el llamado "efecto de suavizamiento". Este consiste en eliminar las pequeñas variaciones o fluctuaciones de la superficie; sería equivalente a multiplicar por cero las componentes de alta frecuencia en un desa -

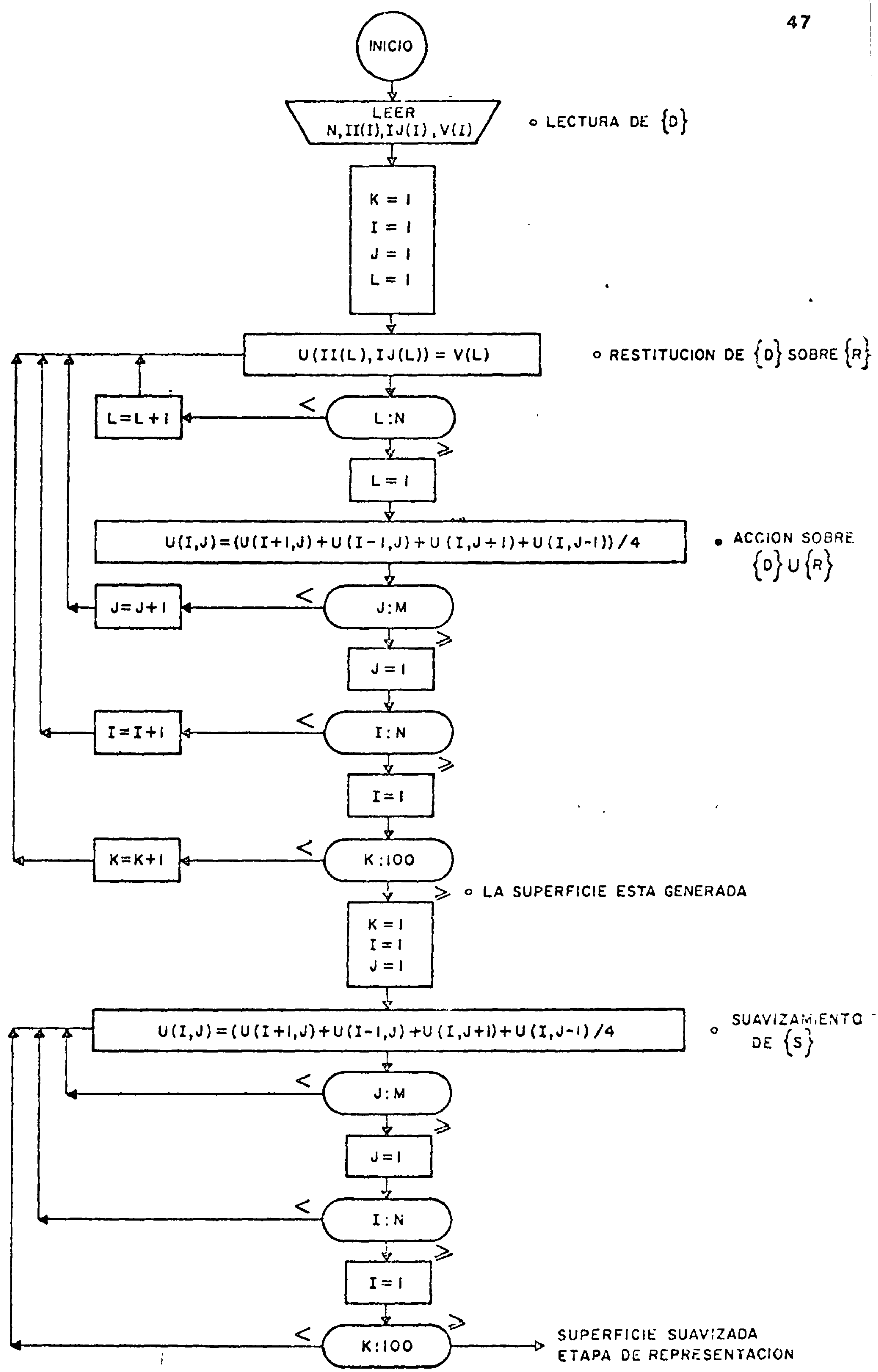
rrollo de la superficie en series dobles de Fourier. El proceso numérico consiste en aplicar nuevamente la ecuación (2.3) a la unión de R y D un número muy reducido de veces. Obviamente el proceso destruirá la información original D si no se toma previamente la precaución de almacenarlo en otros arreglos.

2.4. Implementación

El siguiente programa lee los datos de entrada $\{D\}$, genera $[S]$ a partir de $\{R\}$ y $\{D\}$ obteniéndose una nueva superficie suavizada:

```
      DIMENSION U(50,120),II(100),IJ(100),V(100)
C     PROGRAMA PARA GENERAR UNA MATRIZ DE SUPERFICIE A PARTIR DE TRES
C     VECTORES QUE CONTINENEN EL INDICE DE FILA,DE COLUMNA Y EL VALOR DEL PUNTO.
C     LECTURA DE DATOS
      READ(1,1 ) N
    1 FORMAT(I3)
      DO 2 I=1,N
    2 READ(1,3) II(I),IJ(I),V(I)
    3 FORMAT(2I3,4X,F10.0)
C     INICIO DEL BARRIDO (100 ITERACIONES.
      DO 5 K=1,100
      DO 5 I=2,49
      DO 5 J=5,119
C     ASIGNACION DE D A LA MATRIZ S.
      DO 6 L=1,N
    6 U(II(L),IJ(L))= V(L)
    5 U(I,J)= .25*(U(I+1,J)+U(I-1,J)+U(I,J+1)+U(I,J-1))
C     MATRIZ  S GENERADA
C     SE INICIA PROCESO DE SUAVIZAMIENTO.
      DO 7 K=1,3
      DO 7 I=2,49
      DO 7 J=2,119
    7 U(I,J)= .25*(U(I+1,J)+U(I-1,J)+U(I,J+1)+U(I,J-1))
C     MATRIZ LISTA.
      STOP;END
                                                              SEGMENT
```

El diagrama de flujo correspondiente es:


INICIO
LEER
N, II(I), IJ(I), V(I)
o LECTURA DE {D}
K = 1
I = 1
J = 1
L = 1
U(II(L), IJ(L)) = V(L)
o RESTITUCION DE {D} SOBRE {R}
L = L + 1
L : N
<
>=
L = 1
U(I,J) = (U(I+1,J) + U(I-1,J) + U(I,J+1) + U(I,J-1)) / 4
• ACCION SOBRE {D} U {R}
J = J + 1
J : M
J = 1
I = I + 1
I : N
I = 1
K = K + 1
K : 100
o LA SUPERFICIE ESTA GENERADA
K = 1
I = 1
J = 1
U(I,J) = (U(I+1,J) + U(I-1,J) + U(I,J+1) + U(I,J-1) / 4
o SUAVIZAMIENTO DE {S}
J : M
J = 1
I : N
I = 1
K : 100
SUPERFICIE SUAVIZADA
ETAPA DE REPRESENTACION

Los detalles de funcionamiento deben estar suficientemente claros del diagrama de flujo anterior. Basta solo decir que cada vez que se va a procesar un elemento cualquiera de la matriz, todos los valores de {D} -representados por los vectores II de fila, IJ de columna y V de valores- restituídos a sus posiciones y valores originales. La razón es que a objeto de no alterar los valores originales de {D} sobre {R} durante el proceso de iteración habría que poner una serie de condiciones lógicas (usando instrucciones IF) para detectar cuando estamos sobre un elemento que no debe ser procesado. Esta acción es mucho más lenta que asignar de nuevo los valores de {D} .

Para tener una idea de la cantidad de operaciones aritméticas, lógicas, de almacenamiento y de recuperación de información analizaremos nuevamente el diagrama de flujo. La operación definida por la ecuación (2.3) tiene 4 sumas y una división por cada elemento interno de la matriz, o sea 24.000 sumas y 6.000 divisiones en una iteración. Como tenemos 100 iteraciones el número total de operaciones aritméticas son 2.400.000 y 600.000 que totalizan 3.000.000. Siguiendo igual razonamiento, el número de recuperaciones aleatorías totalizan -supoiendo que el conjunto $\{D\}$ tiene en promedio 10 elementos - 9.000.000. Es comprensible que la generación de una matriz de superficie de 50 x 120 elementos tome entre 12 y 15 minutos de ejecución en un computador B-5500 sin contar el tiempo necesario para generar la cinta magnética para el Delineador Digital que contiene la imagen de la superficie o la generación de la misma por la impresora de línea.

Es imprescindible por lo tanto buscar nuevos algorítmos de búsqueda, asignación e interpolación de valores para tratar de reducir drasticamente el - tiempo de generación.

Para la búsqueda y asignación es evidente que hay que cambiar de una forma matricial de representación a una forma vectorial. Esto se justifica - porque a nivel interno del Computador es más eficiente el manejo de arreglos unidimensionales que multidimensionales.

En cuanto a los algorítmos de interpolación se están probando algunos que barren la matriz una sola vez. Los resultados será reportados en un futuro trabajo.

2.5. Representación de la Superficie

Una vez que se obtiene $\{S\}$ en forma de $[S]$ por el método (u otro cualquiera) descrito en 2.2 se entra en la etapa de producir distintos tipos de imágenes que dén al observador toda la información necesaria para interpretar el significado de los 18.000 valores numéricos que conforman a S .

Los métodos y algorítmos desarrollados hasta ahora por el autor, están en etapa experimental de desarrollo y por lo tanto no constituyen nada definitivo.

2.5.1. Despliegue Numérico de la Matriz

Este método se basa en la transformación del rango de los valores numéricos de la matriz en otro rango normalizado (0. A 1., A 10.0 etc.). El procedimiento, que ya fué explicado en el punto 2.1.1., consiste en encontrar los elementos de valores máximo y mínimo para efectuar los cambios de escala correspondientes al intervalo escogido. A continuación se transforma -

la matriz real normalizada en otra entera. Al imprimir por la impresora de línea el - resultado, cada uno de los valores origina les que ocupaban 10 espacios de impresión pasa a ocupar uno o cuando más, dos espa - cios; esto permite obtener una representa- ción (para el caso del rango 1 a 10) numé- rica,en la cual los elementos originales - de la matriz quedan categorizados o asigna dos a clases, cada una abarcando un subran go de valores. En la hoja resultante se - puede por lo tanto trazar a mano rápidamen te curvas de nivel que informan sobre las características de la superficie. Para - los números enteros con dos cifras (10, 11, 12 ......) se puede utilizar la convención de asignarlos a los caracteres alfanuméri- cos A, B ..... Z = , ., +, - etc. o sobre impresión de ellos. A continuación se - muestra un programa que efectua dicho pro- ceso sobre una matriz T de 50 x 50 ele - mentos utilizando un método ligeramente di ferente y menos general que el descrito an teriormente:

```
          B 5 7 0 0     F O R T R A N     C O M P I L A T I O N  XIII. 0,
FILE  1=DATA, UNIT=READER
FILE  3=OUTP, UNIT=PRINTER
                                                                START OF S
      INTEGER TGRAF(50),CODA(11)
      DIMENSION A(26,26),T(50,50)
      DATA CODA/"1","2","3","4","5","6","7","8","9","A","B"/
      DO 1 I=1,50
      T(1,I)=I*2
    1 T(50,I) =102-I*2
      DO 710 J=7,18
  710 T(J,26)=25
      DO 711 I=20,35
  711 T(I,1)=100
      DO 712 I=10,23
  712 T(I,50)=100
      T(35,40)=95
      T(8,8)=100
      T(20,20)=100
      DO 3 K=1,60
      DO 2 I=2,49
      DO 2 J=2,49
      IF(I.EQ.35.AND.J.EQ.40)  GO TO 2
      IF(I.EQ.8.AND.J.EQ.8) GO TO 2
      IF(I.EQ.20   .AND. J .EQ. 20 ) GO TO 2
      T(I,J)=  (T(I+1,J)+T(I-1,J)+T(I,J+1)+T(I,J-1))/4
    2 CONTINUE
    3 CONTINUE
      DO 9 I=1,50
      DO 12J=1,50
   12 TGRAF(J)= CODA(T(I,J)/9.9+1)
    9 WRITE(3,10)  TGRAF
   10 FORMAT(10X,50A2)
      CALL PARTIR(T)
      STOP;END
```

El programa comienza estableciendo los vec_ tores TGRAF (50) y CODA(11). El primero - para copiar los elementos de cada fila para poder procesarlos sin destruirlos; -- CODA(11) establece los caracteres que defi_ nen cada subrango a ser asignados a la matriz T original. El lazo DO 1 I = 1,50 establece dos distribuciones lineales de -

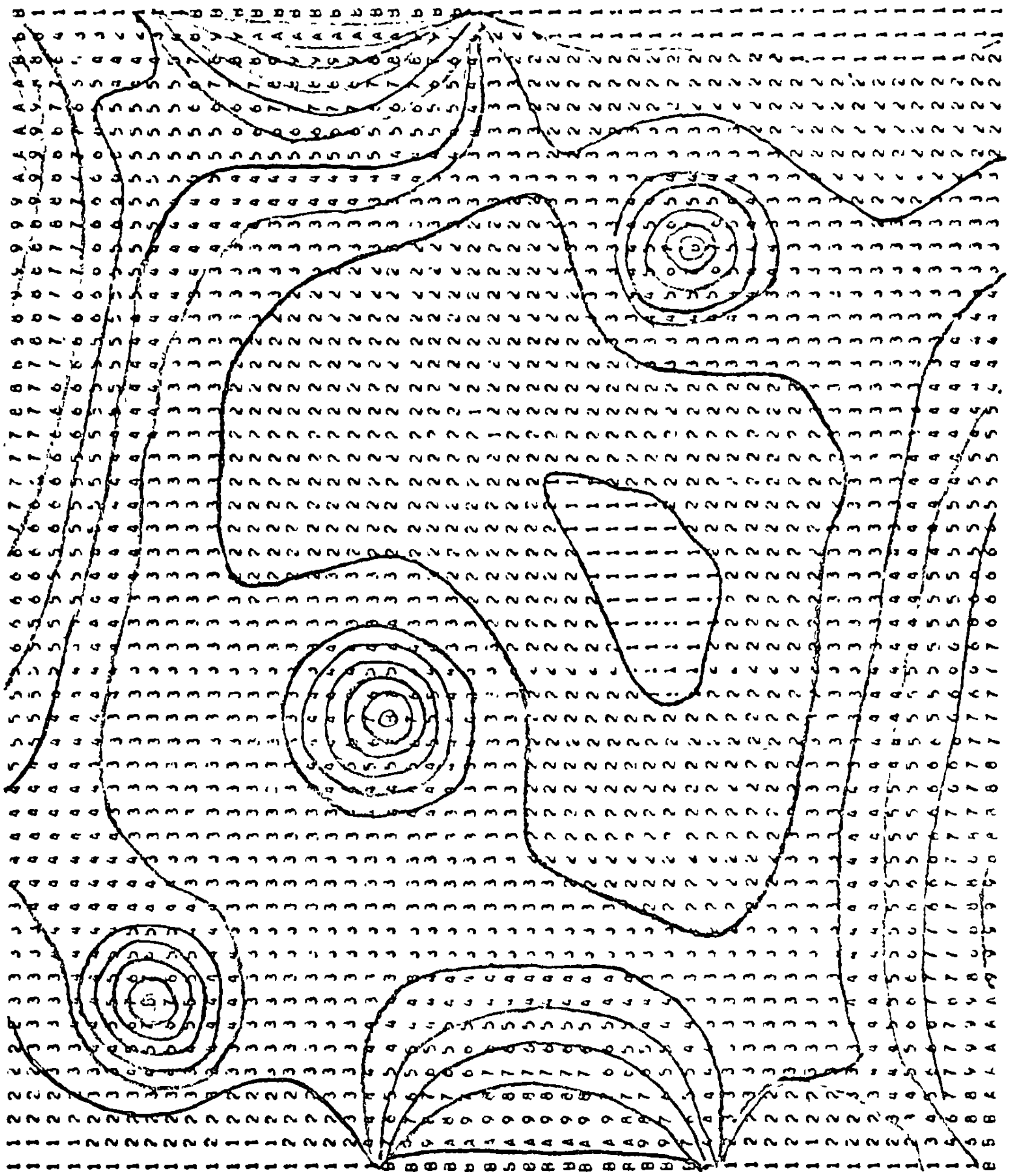
valores para la primera y última columna de la matriz que van en el rango 1 a 100 directa e inversa respectivamente.

Los lazos DØ 710, DØ 711, DØ 712 definen asignaciones de valores a sub-columnas. En los puntos (35, 40), (8, 8) y (20, 20) se colocan los valores 95, 100 y 100. A continuación se inicia el proceso de iteración aplicando 60 veces el algorítmo - del promedio móvil (DØ 3 K = 1,60) sobre la matriz T. Nótese que para no alterar los valores del conjunto $\{D\}$ se em-- plea una secuencia de IF lógicos para detectar en que momento se llega a estos elementos y evitar procesarlos. El nido de lazos DØ 9 I = 1,50 y DØ 12 J = 1,50 junto a la instrucción TGRAF(J) = CODA - (T (I,J)/9.9 + 1) divide cada elemento de T(I, J) entre 9.9 produciendo nuevos elementos en el rango 0.0 a 10,1, que al ser sumados con 1 caen entre 1 y 11. Este rango de valores enteros constituyen un subíndice variable para el vector - CODA(11) que contiene los caracteres alfanuméricos que van a ser asignados al vector de impresión TGRAF(50) por direccionamiento indirecto. TGRAF es impreso 50 veces para desplegar la matriz. Nótese que esta permanece inalterada durante el proceso. El resultado puede observarse en la pág. 55 y al cual se le ha trazado a mano algunas curvas de nivel.

Se puede observar que la matriz T(50, 50) tenía como máximo valor numérico 100, de tal forma que el programa anterior no era general. Un subprograma más versátil se define a continuación:

```
      SUBROUTINE ALFANU(T)
      REAL MIN,MAX
      DIMENSION  T(50,120),TGRAF(120),CODE(10)
      DATA CODE/"1","2","3","4","5","6","7","8","9","H"/
C  PROGRAMA PARA CONVERTIR UNA MATRIZ EN UN DESPLIEGUE DE CARACTERES ALFANUME
C -RICOS POR NORMALIZACION AL RANGO (1,10).
C          BUSQUEDA DEL ELEMENTO MINIMO.
      MIN= T(1,1)
      DO 1 I=1,50
      DO 1 J=1,120
      IF(T(I,J).LT.MIN) GO TO 1
      MIN=T(I,J)
    1 CONTINUE
C          BUSQUEDA DEL ELEMENTO MAXIMO
      MAX= T(1,1)
      DO 2 I=1,50
      DO 2 J=1,120
      IF(T(I,J).GT.MAX) GO TO 2
      MAX=T(I,J)
    2 CONTINUE
C  CONVERSION AL RANGO (1,10)
      DT=MAX-MIN
      DO 3 I=1,50

      DO 3 J=1,120
      TGRAF(J)=   CODE((T(I,J)-MIN)/(DT)+1)
    3 WRITE(3,4) TGRAF
    4 FORMAT(10X,50A1)
      RETURN,END
```

SEGMENT



Esta subrutina acepta una matriz de 50 x 120 elementos y la convierte en el tipo - de representación ya discutido. La única diferencia con el programa anterior es - que se busca el máximo y el mínimo de la matriz para reducirla al rango (1, 10) - sin destruir la información original.

2.5.2. Despliegue Simbólico de la Matriz

Usando una técnica basada en la anterior, se puede representar una matriz de superficie [S] mediante la sobreimpresión de caracteres alfabéticos que sean simé - tricos respecto al eje longitudinal. Si se combinan los caracteres:

.,-,=,I,O,Z,O,H,X,O,X
b̸,b̸,b̸,I,H,I,O,H,O,H,O
b̸,b̸,b̸,b̸,b̸,b̸,b̸,b̸,H,O,H (b̸ = Blanco)

De tal forma que se sobreimpriman aquello que ocupan una misma columna, se producirá una variación tonal que irá desde el - punto (.) hasta un pequeño cuadro negro - (▩). Una subrutina adecuadamente diseñada puede implementar esta idea para pro ducir un despliegue de $[S]$ mediante una variación gradual de claroscuro como una función aproximadamente líneal de la categorización dada a los valores de la matriz.

La característica funcional que permite tal posibilidad es el comando de impresora "+" en la primera columna de cada línea a imprimir, es decir, impedir el salto de línea. La subrutina mostrada seguidamente produce una verdadera imagen de isolíneas a partir de $[S]$.

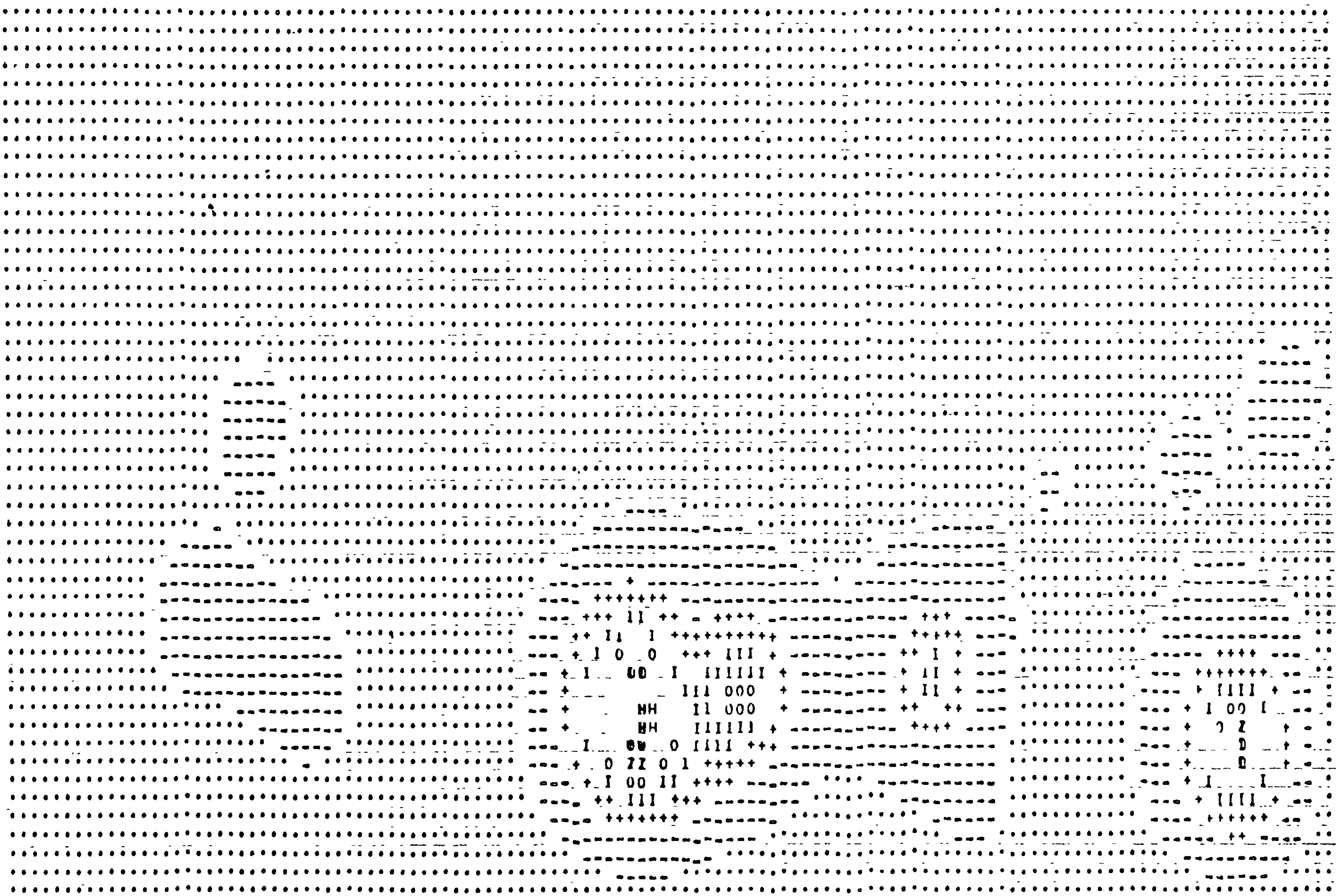
```
      SUBROUTINE IMPRIME(E,MAX)
      REAL MAX
      DIMENSION E(50,122),DIB(122),DIBU(122),CODF(10),CODE2(10),CODE3(10
     *)
      DIMENSION CODE1(10)
      DATA (CODE1(I),I=1,10) /".","-","+","I","O","Z","I","H","X","H"/
      DATA (CODE2(I),I=1,10) /" "," "," "," "," ","1","O","O","I","X"/
      DATA (CODE3(I),I=1,10) /" "," "," "," "," "," "," ","+",".","O"/
      WRITE(3,1)
    1 FORMAT("1")
      DO 8 I=1,50
      DO 2 J=1,122
    2 DIB(J)=E(I,J)/MAX*9+1
      DO 3 L=1,122
    3 DIBU(L)= CODE1(DIB(L))
      CALL LIMITE (DIBU)
      WRITE(3,4)DIBU
    4 FORMAT(" ",7X,122A1)
      DO 5 L=1,122
    5 DIBU(L)= CODE2(DIB (L))
      CALL LIMITE (DIBU)
      WRITE(3,9) DIBU
      DO 7 L=1,122
    7 DIBU(L)=CODE3(DIB (L))
      CALL LIMITE (DIBU)
      WRITE(3,9) DIBU
    9 FORMAT("+",7X,122A1)
    8 CONTINUE
      RETURN
      END
                                                                        S
                                                                 START OF
      SUBROUTINE LIMITE(DIBU)
      DIMENSION  DIBU(122)
      DO 10 K=2,122
      IF(DIBU(K).EQ.DIBU(K-1)) GO TO 10
      DIBU (K-1)=" "
   10 CONTINUE
      RETURN;END
```

La subrutina comienza por aceptar la matriz E(50, 122) y el valor máximo MAX. - Mediante la instrucción DIB(J) - E(I, J)/ MAX*9+1 en un lazo iterativo se toma la fila i-esima de la matriz para transfor - marla en un conjunto cuyos límites son 1 y 10. Es conjunto se copia en el vector DIB(J).

Aplicando tres veces la técnica descrita en 2.5.1 se sobreimprimen una secuencia partícular de caracteres de acuerdo a los vectores CODE1(10), CODE2(10), CODE3(10). El vector DIBU(I) podría tener sucesiva - mente tres secuencias del tipo:

```
XXXXXXXX......HHHHHHH--------000000000
00000000      XXXXXXIIIIIIII+++++++++
HHHHHHHH      IIIIII00000000IIIIIII
```

La subrutina límite (DIBU) procesa el vector DIBU y mediante la instrucción -- IF(DIBU(`).EQ.DIBU(K-1)) GØ TØ 10 detecta cuando hay un cambio de carácter. Si existe un cambio de carácter entre el K y el K-1, este es eliminado y sustituido - por un blanco. El resultado es que las - diferentes regiones del plano constituí -



das por un tipo especial de sobreimpresión de caracteres, estan separadas por una línea blanca que demarca perfectamente los límites. En la página se puede observar la imagen producida.

2.5.3. Modelos Tridimensionales

La forma más exacta de reproducir el contenido espacial de [S] es construir un modelo sólido en tres dimensiones. Tal técnica requiere el procesamiento adecuado de los 18.000 números que constituyan los 6.000 punto de [S] para producir una serie de instrucciones que comandan una fresadora a control numérico. La fresa de tal máquina herramienta talla sobre un bloque de material blando una superficie, fiel reflejo del fenómeno que se intenta estudiar.

Esta forma de representar la superficie es poco conveniente en la mayoria de los casos por lo incómodo de manejar el sólido resultante, además de que se necesita una máquina economicamente poco atractiva.

Sin embargo existen algunos casos en que se justifica tal procedimiento como por ejemplo, por motivos docentes, académicos, de demostración, etc.

El autor ha ideado un método mediante el cual, y con la ayuda de la impresora de línea, se puede construir el modelo tridimensional en cuestión. El procedimiento consiste en producir un despliegue alfanumérico de cada columna de la matriz $[S]$ y cortar la figura obtenida en láminas de algun material blando. El procedimiento es largo y tedioso pero es la única forma económica de producir en nuestro medio un modelo de tal naturaleza.

La parte principal del programa posee las siguientes instrucciones:

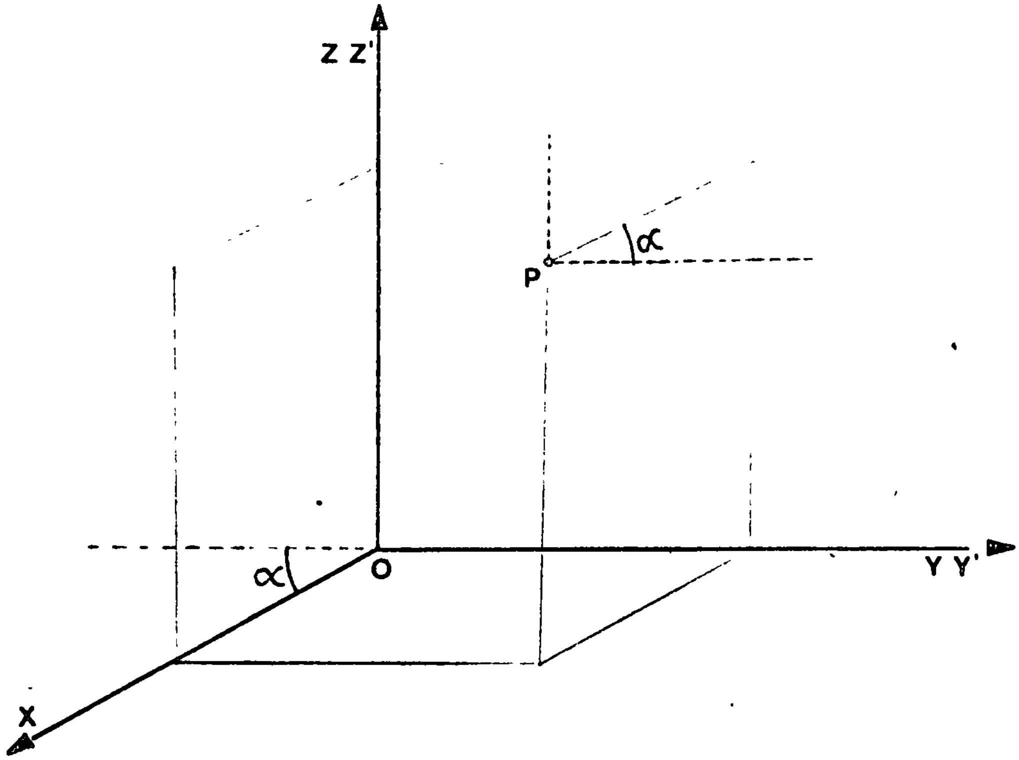
```
      .
      .
      .
      DØ 1 J=1,120
      DØ 2 I=1,50
      LIMIT=A(I,J)/MAX*99+1
      DØ 3 K=1,LIMIT
3     PERFIL(K)='H'
      WRITE(3,6) PERFIL
      DØ 4 K=1,100
4     PERFIL(K)=' '
2     CONTINUE
      WRITE(3,5)
1     CONTINUE'
5     FØRMAT('1')
6     FØRMAT(2X,100A1)
      .
      .
      .
      .
      .
```

El programa comienza explorando cada elemento de la matriz a lo largo de cada columna para obtener un valor proporcional en el rango [1, 100] . Este valor, almacenado en la variable entera LIMIT, sir ve como límite superior de un lazo iterativo que se encarga de colocar el carácter H en el vector PERFIL(I), I=1, 100. Este es impreso a razón de 50 veces por página para producir la silueta del vector columna. Las 120 páginas producidas en la impresora pueden ser procesadas por un dibujante que construirá el modelo.

2.5.4. Proyecciones Oblícuas

Trataremos el problema de la proyección de $[S]$ (que puede considerarse como una forma perteneciente al espacio tridimensional) en un espacio bidimensional de tal manera que la imagen producida proporcione la sensación de relieve y profundidad a un observador.

La primera consideración consiste en transformar los tres conjuntos $\{X\}$, $\{Y\}$, $\{Z\}$ en dos conjuntos $\{X'\}$, $\{Y'\}$ que definen puntos en un plano. De la figura anexa puede notarse lo siguiente:

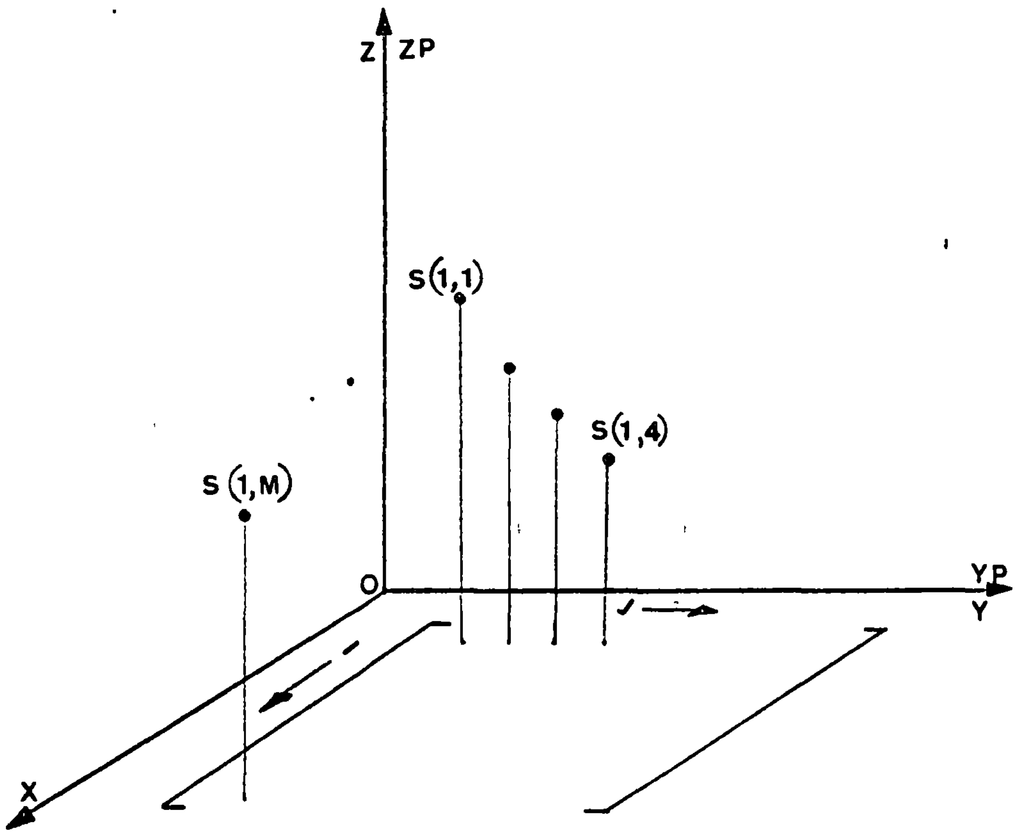


$$Z' = Z - X.\text{Cos}\alpha$$

$$Y' = Y - X.\text{Sen}\alpha$$

Es decir que la coordenada X de profundidad resta sus componentes proyectadas de la Z e Y correspondientes.

Para producir una proyección oblícua de u na superficie partiendo de la matriz co-- rrespondiente $[S]$ se deberá procesar adecuadamente los tres conjuntos $\{I\}$, $\{J\}$, $\{A(I, J)\}$ para proyectarlos en dos nuevos - conjuntos $\{YP\}$ y $\{ZP\}$. Para ello se procede de la siguiente forma (ver figura):



Suponiendo la matriz [S] en el plano XY puede notarse que existe la siguiente correspondencia entre la matriz, los índices I,J y las coordenadas X, Y, Z:

X = INDICE DE FILA I

Y = INDICE DE COLUMNA J

Z = VALORES DE LA MATRIZ S(I,J)

de tal manera que se puede deducir inmediatamente:

$$Y = J - I*COS\alpha$$

$$Z = A(I, J) - I*SEN\alpha$$

las siguientes instrucciones realizan la transformación:

```
      .
      .
      DØ 1 I=1,M
      DØ 1 J=1,N
      YP(J+(I-1)*M)= J-I*COS(ALFA)
1     ZP(J+(I-1)*M)= A(I,J)-I*SIN(ALFA)
      .
      .
      .
```

es decir, los 3M x N números de la matriz A(M, N) quedan convertidos en dos vectores de M x N elementos cada uno. Como subíndice de los vectores YP, ZP produce el orden exacto de secuencia lineal de un elemento de la matriz. En general:

(J +(I-1)*M) = 1,MxN   I = 1,M
                              J = 1,N

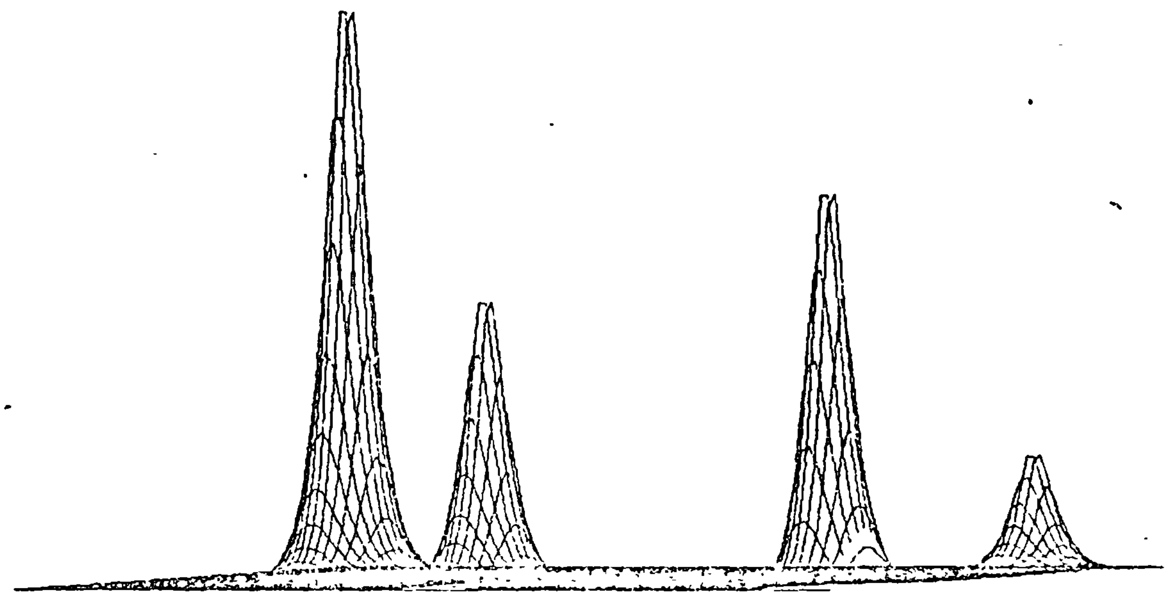
Colateralmente habrá que procesar los vectores YP y ZP para ajustarlos a las dimensiones del papel del Delineador. A conti nuación podrá examinarse la subrutina D3D2(A) destinada a generar un dibujo en proyección oblícua de una matriz de superficies:

```
      SUBROUTINE D3D2(A)
      DIMENSION A(50,120),YP(6000),ZP(6000) ,B(50,120)
      DATA FACY,FACZ /6,3/
      DATA N,M /50,120/
      DATA ALFA/45/
      REAL MAXY,MAXZ,MINY,MINZ,MAXA
      MAXA=B(1,1)
      DO 12 I=1,N
      DO 12 J=1,M
      B(I,J)=A(I,J)
      IF(MAXA.GT.B(I,J)) GO TO 12
      MAXA=B(I,J)
   12 CONTINUE
      DO 13 I=1,N
      DO 13 J=1,M
   13 B(I,J)=(B(I,J)/MAXA)*100
C TRANSFORMACION 3D A 2D.
      DO 999 ALFA= 5,175,20
      CALL PLOT(0,0,-3)
      ANG= ALFA*3.1416/180
      C=COS(ANG);S=SIN(ANG)
      DO 1 I=1,N
      DO 1 J=1,M
      YP(J+(I-1)*M)=J-I*C
    1 ZP(J+(I-1)*M)=B(I,J)-I*S
      L=N*M
      MINY=YP(1);MINZ=ZP(1)
      DO 2 I=2,L
      IF(YP(I).GT.MINY)GO TO 2
      MINY=YP(I)
    2 CONTINUE
      DO 3 I=1,L
    3 YP(I)=YP(I)-MINY
      DO 4 I=2,L
      IF(ZP(I).GT.MINZ) GO TO 4
      MINZ=ZP(I)
    4 CONTINUE
      DO 5 I=1,L
    5 ZP(I)=ZP(I)-MINZ
      MAXY=YP(1);MAXZ=ZP(1)
      DO 6 I=2,L
      IF(YP(I).LT.MAXY) GO TO 6
      MAXY=YP(I)
    6 CONTINUE
      DO 7 I=2,L
      IF(ZP(I).LT.MAXZ) GO TO 7
      MAXZ=ZP(I)
    7 CONTINUE
      DO 8 I=1,L
      YP(I)=YP(I)/MAXY*FACY
    8 ZP(I)=ZP(I)/MAXZ*FACZ
      WRITE(3,11)  (YP(I),I=1,500)
   11 FORMAT(20F6.2)
      WRITE(3,11)  (ZP(I),I=1,500)
      DO 10 I=1,N
      J=1
      CALL PLOT(YP(J+(I-1)*M),ZP(J+(I-1)*M),3)
      DO 9 J=2,M
    9 CALL PLOT(YP(J+(I-1)*M),ZP(J+(I-1)*M),2)
   10 CONTINUE
  999 CONTINUE
      RETURN
      END
```

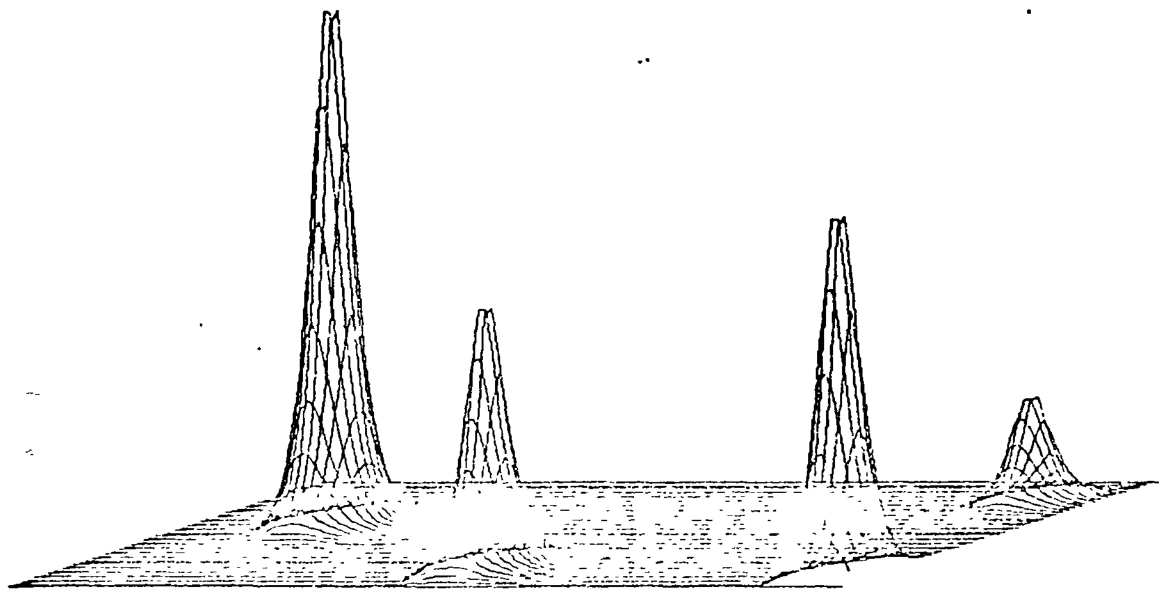
El subprograma comienza por aceptar la matriz A(50, 120), establecer un nuevo orígen según las coordenadas del "PLOTTER" y transofrmar el ángulo ALFA dé grados a radianes. A fin de salvar la información - de A(I,J), esta es copiada en el arreglo B(I,J), el cual es convertido en una ma-- triz proporcional a A(I,J) en el rango - [0,100] . Fila a fila B(I,J) es explorada para convertir cada elemento ma - tricial en dos elementos vectoriales -- YP(I) y ZP(I), tal como fué definido anteriormente. Los valores máximos y mínimos de estos vectores: MINZ, MAXZ, MINY, MAXY son calculados para reducir estos conjuntos al rango [0,1] . Una multiplica - ción por los factores FACY, FACZ ajustan el tamaño del dibujo en pulgadas. El nido de lazos iterativos DØ 10 I=1,N y - DØ 9 J=2,M funcionan en tal forma que el Delineador comienza a dibujar empezando - por la primera fila de la matriz tomando los primeros 120 elementos de cada vector YP y ZP, los segundos 120 elementos, etc.

Este tratamiento global del problema permitirá como un futuro desarrollo, incorporar el problema de la recta tapada. Para tal efecto se comienza a dibujar primero por la última fila de la matriz. La curva ploteada establece un "horizonte". La segunda curva, que correspondería a la penultima fila, es a su vez dibujada pero cuando trata de penetrar el "horizonte" establecida por la primera, la plumilla del Delineador es automáticamente levantada hasta emerger (en caso de que lo haga) de la zona prohíbida. Cada nueva curva va estableciendo un nuevo "horizonte" hasta finalizar el dibujo. En las superficies que siguen a continuación se presentó una matriz S(50,120) con cuatro puntos de información. El ángulo ALFA se hizo variar entre 10 y 170 grados.

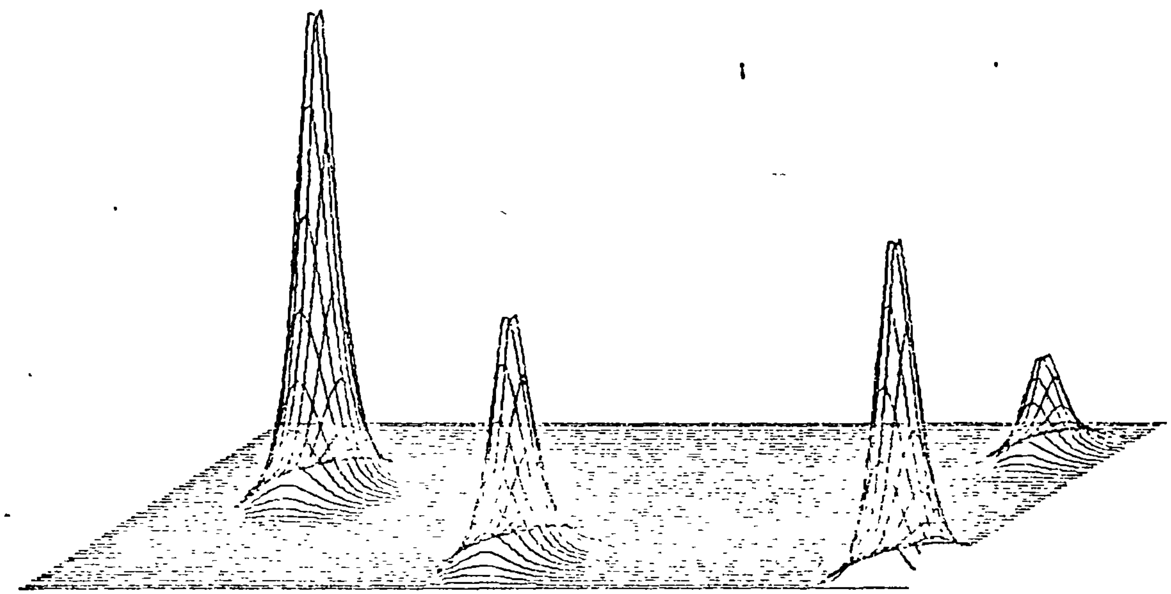
Nótese que el efecto es equivalente a una translación del observador alrededor de la superficie:



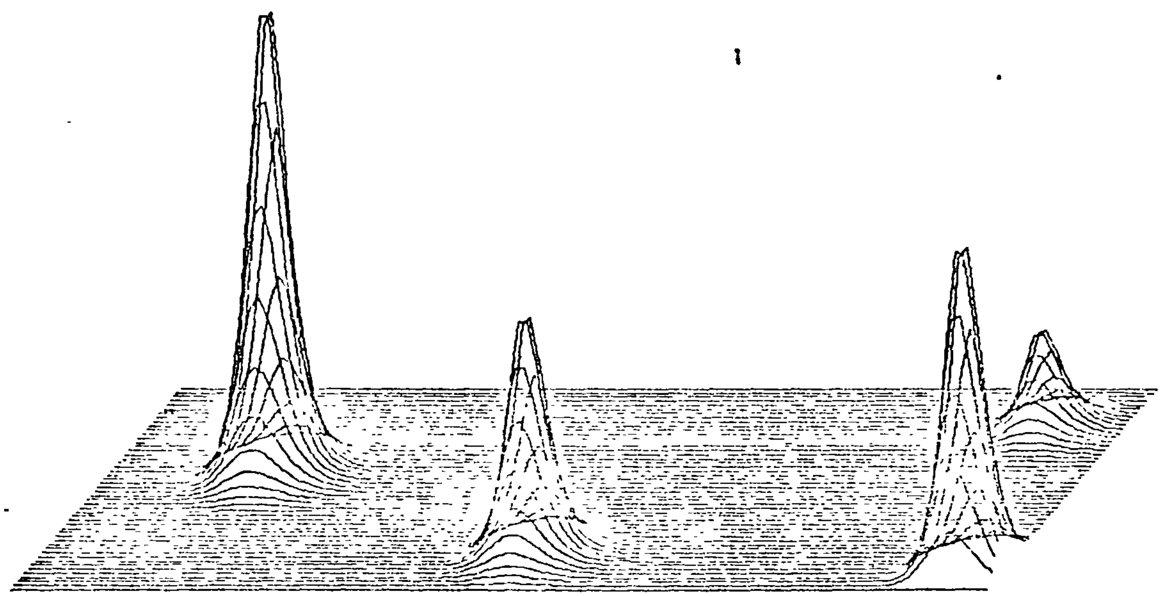
$\theta = 10^{\circ}$



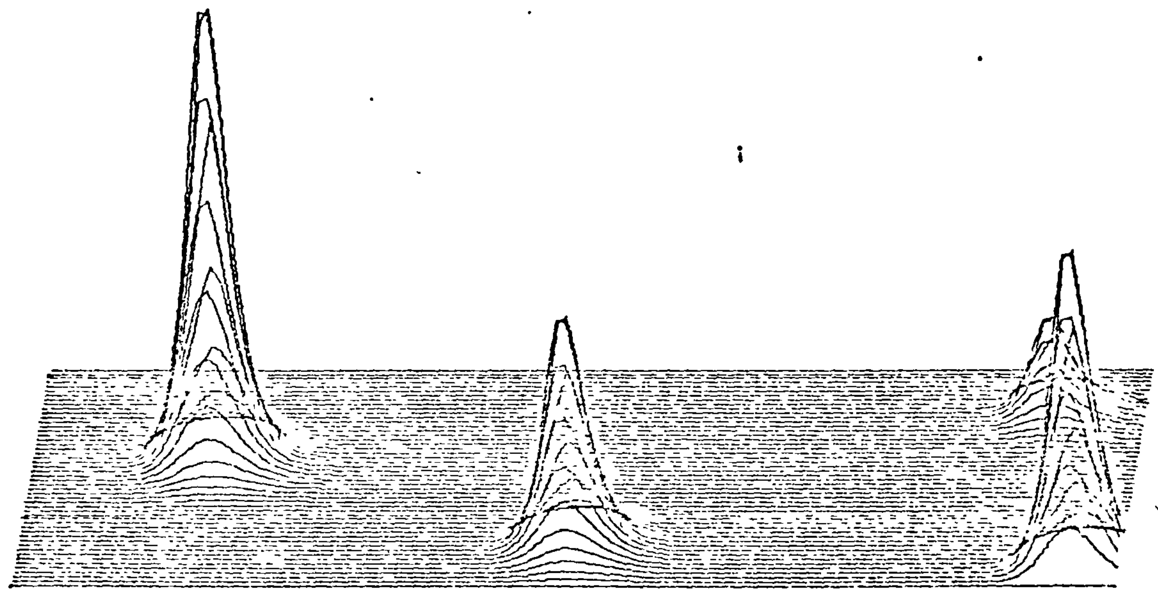
θ= 30°



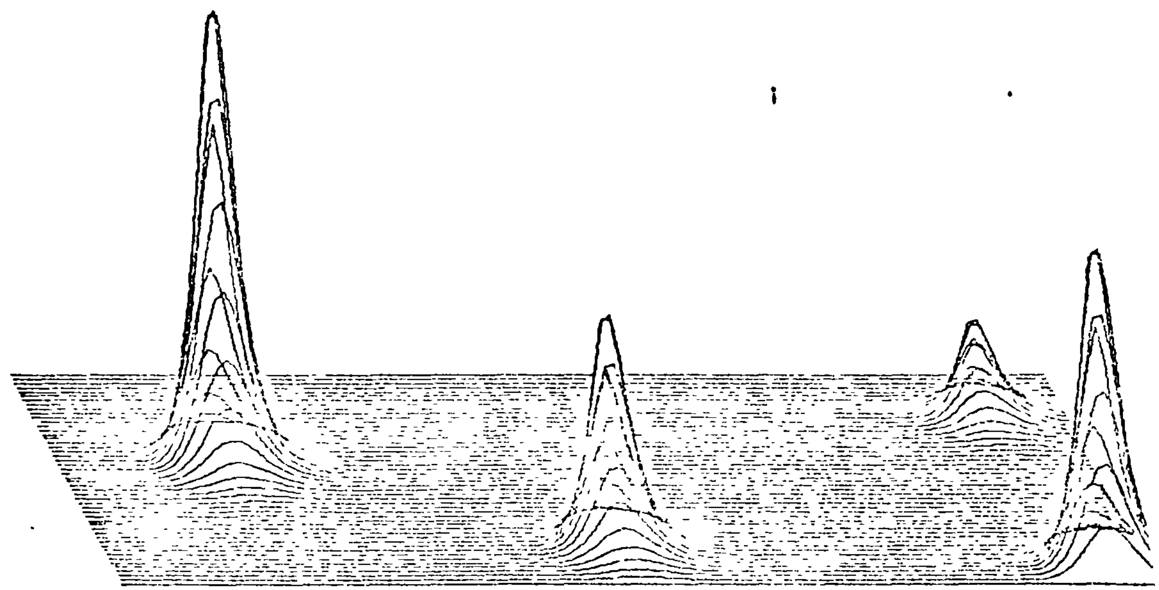
$\theta = 50^{\circ}$



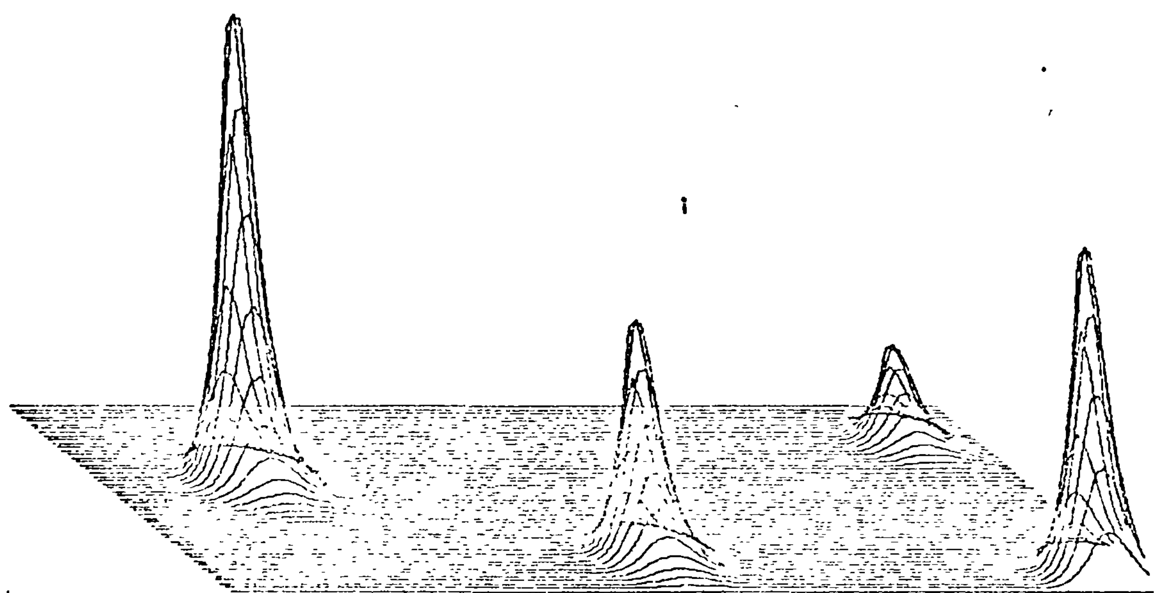
$\theta = 70^{\circ}$



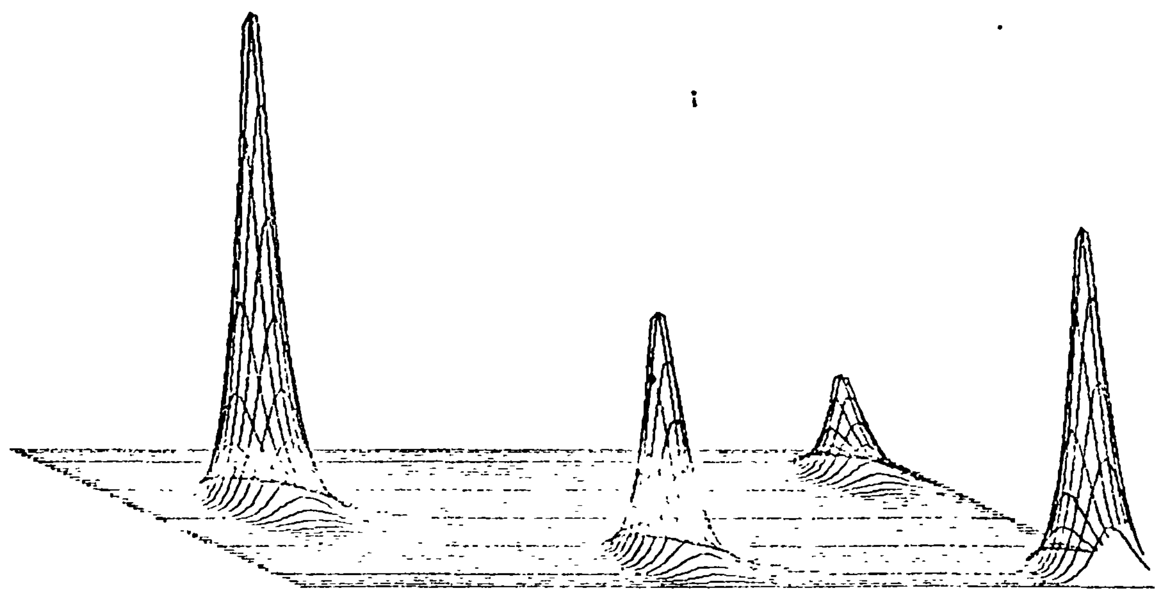
$\theta = 80^{\circ}$



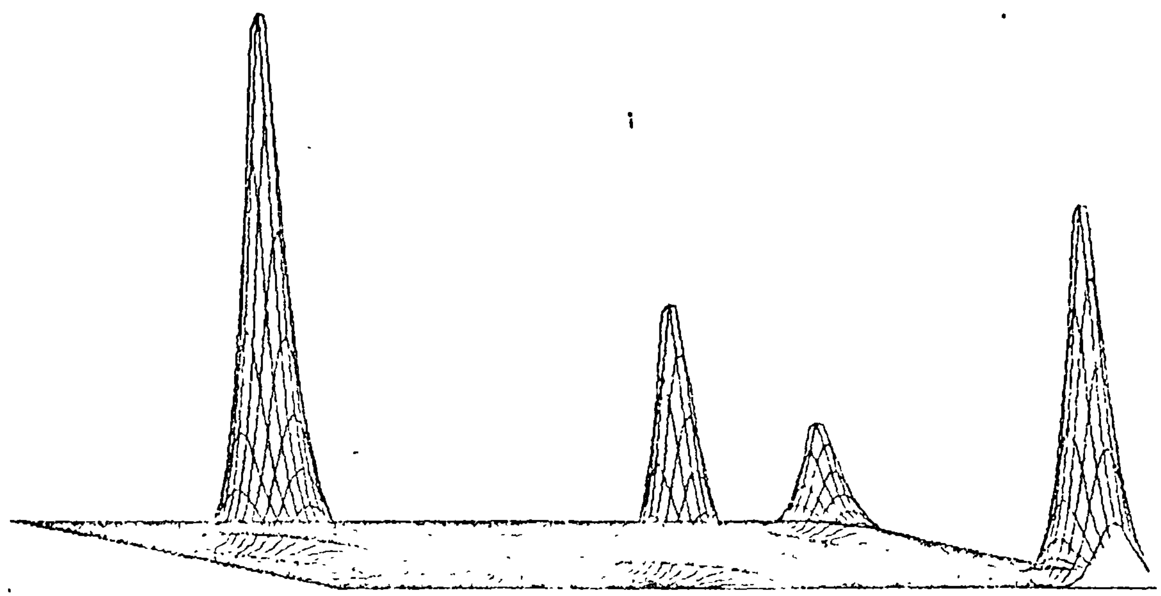
θ=110º



$\theta = 145^{\circ}$



$\theta = 160^{\circ}$

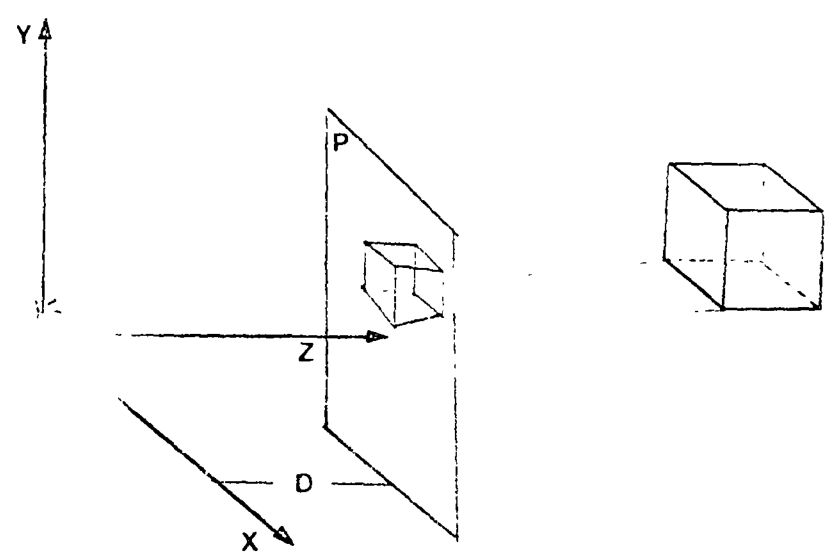


$\theta = 170^{\circ}$

2.5.5. Proyección Perspectiva (Un Punto de Fuga)

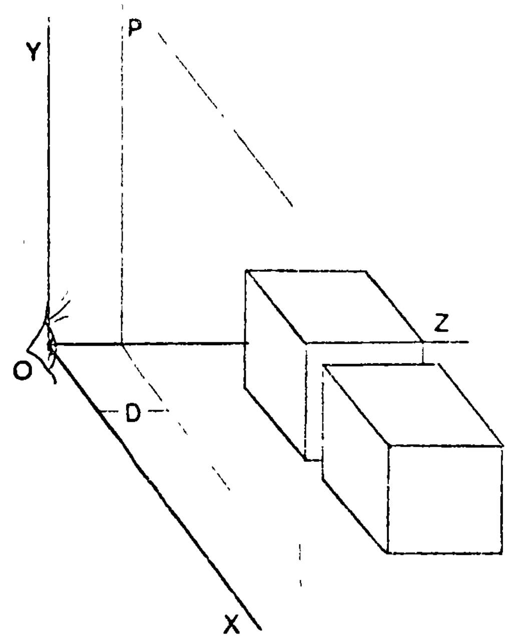
De las figuras anteriores podrá notarse que no existe un verdadero efecto de profundidad, es decir, la imagen es contemplada por el observador desde un punto en el infinito del tal forma que los rayos o visuales llegan paralelamente.

Un algorítmo que vale la pena implementar en un Computador es aquel que produce las transformaciones necesarias sobre tres conjuntos $\{X\}$, $\{Y\}$, $\{Z\}$ que describen un sólido en el espacio para producir una imagen bidimensional en perspectiva. Graficamente:



Es decir, se ha escogido como punto de origen del sistema, el aparato visual de un observador. Del objeto parten rayos que se proyectan en el plano de proyección situado a una distancia D del observador y paralelo al plano XY. Obviamente que cuando aumenta la profundidad (coordenada Z) las dimensiones X e Y correspondientes disminuyen tal y como ocurre en el mundo real.

Las transformaciones respectivas, aunque muy sencillas no se darán en este trabajo. Uno de los experimentos realizados muestran como lucirán dos cubos colocados a cierta altura sobre la zona positiva del plano ZX:



El plano de proyección, inicialmente muy cerca del observador, va alejándose de éste para producir intersecciones diferentes del cono de rayos. El efecto puede verse en la página 83. El dibujo fué producido en el Delineador Digital.

Si el plano de proyección P forma un ángulo diferente a cero con el plano XY la perspectiva tiene dos puntos de fuga y las transformaciones correspondientes son más complicadas.

Todas las consideraciones anteriores estan siendo utilizadas por el autor al despliegue de matrices de superficie aplicando los mismos lineamientos seguidos en el caso de la proyección oblícua.

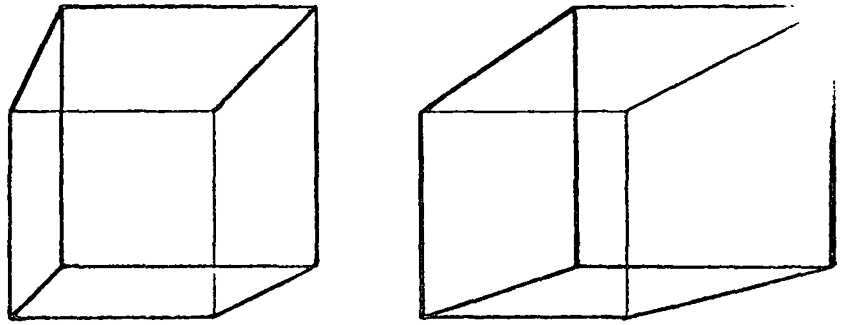
2.5.6. Proyección Estereográfica

Este tipo de proyección se basa en la generación de dos imágenes similares pero poniendo el observador en otro punto desplazado respecto al primero. Se trata en este caso, de producir una imagen para cada ojo de manera que el cerebro las interprete como una visión del mundo real. El par resultante se observa con un estereógrafo.

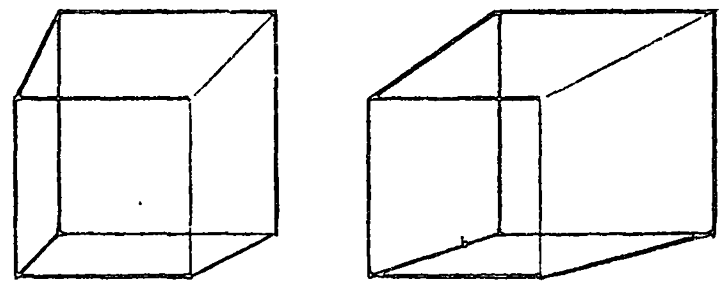
Para generar con un Computador Digital estos tipos de imágenes se hace uso de la proyección en perspectiva descrita anteriormente en 2.5.5. Bastará para ello dibujar la superficie en cuestión para formar la primera imagen. La segunda imagen se obtiene a partir de las coordenadas de la primera mediante la transformación:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} + \begin{bmatrix} \Delta X \\ 0 \\ 0 \end{bmatrix}$$

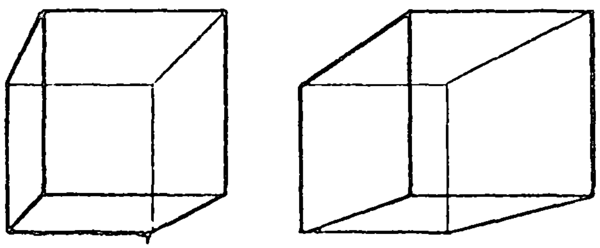
Donde el efecto neto consiste en desplazar una pequeña distancia ΔX el objeto en dirección perpendicular al observador.



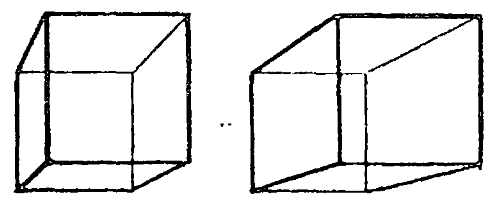
7



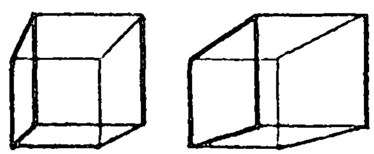
6



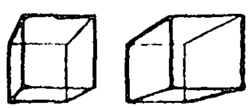
5



4



3



2

1

2.5.7. Proyección Ortogonal

Una de tales proyecciones ya fué analizada en las páginas 56 a 59 en la forma de una imagen de isolíneas obtenida en la impresora. Existen programas para alcanzar el mismo objetivo que utilizan como medio de salida al Delineador Digital.

Construir una imagen de contorno a partir de un retículo rectangular involucra tres aspectos: manejo de un elemento del retículo, creación de una hilera de líneas y el suavizamiento de las mismas.

Un elemento queda definido por las cuatro esquinas del rectángulo formado por cuatro posiciones de la matriz. Una de las ideas básicas consiste en detectar si uno cualquiera de los lados del rectángulo intercepta o toca una curva de nivel. En caso afirmativo se dibuja una línea a través del rectángulo. Otra forma consiste en dividir el elemento con una o dos diagonales para formar dos o cuatro triángulos en cuyos lados se interpolan los valo res de la curva de nivel.

Si se requiere mucha precisión se puede ajustar una superficie polinómica de 2do. y 3er. grado de los tipos:

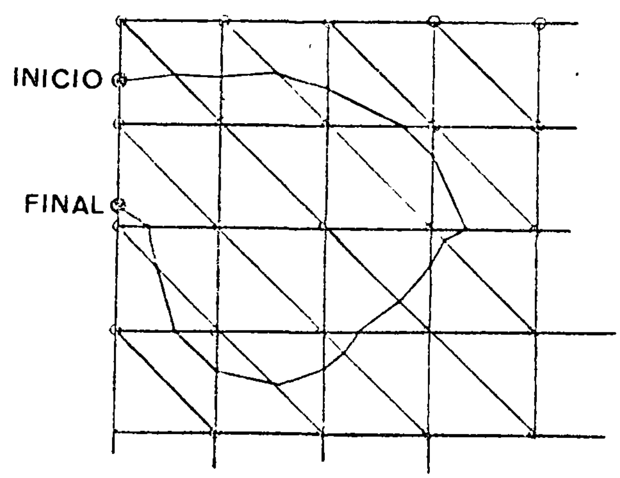
$$z = A + Bx + Cy + Dx^2 + Exy + Fy^2$$

$$z = A + Bx + Cy + Dx^2 + Exy + Fy^2 + Gx^3 + Hx^2y + Ixy^2 + Jx^3$$

A partir de los valores de las cuatro esquinas y otros puntos circundantes por métodos de mínimos cuadrados.

El primer método introducido, por lo sencillo, es el más rápido. El último es más lento como puede verse por las ecuaciones que repetidas veces tiene que manejar.

A nivel de programación, la forma más fácil de tratar la matriz [S] consiste en generar las curvas de nivel explorando columna a columna o fila a fila. Este enfoque utiliza la menor cantidad de memoria pero hace el proceso de suavizamiento muy difícil. Otra técnica de procesamien to verifica el borde de la matriz y busca a lo largo de cada columna el comienzo de la curva de nivel. Bajo escrutinio para seguirla dentro de la matriz hasta que se cierre como un lazo, o salga por un borde de [S] . Este algorítmo necesita un - Computador de acceso aleatorio muy rápido, ya que se tener toda la matriz - en memoria principal durante la corrida - del programa (ver figura), aunque el proceso de suavizamiento es ahora más sencillo.



Adicionalmente, el Delineador Digital ejecutará más rápido el dibujo debido al hecho de que una vez que se ha seleccionado una curva en partícular, esta es trazada en forma total.

En general, la experiencia demuestra que en la generación de imágenes de curvas de nivel en forma automatizada no vale la pena refinar demasiado los algorítmos debido al hecho de que la información adicional obtenida no justifica el costo en tiempo de máquina necesario.

Una forma de generar a $\{S\}$ consiste en asignar a cada punto de $\{R\}$ el valor del elemento de $\{D\}$ mas cercano. Obviamente que ya no se dispone de una superficie continua sino de una serie de subsuperficies horízontales escalonadas, en cuyo centro geométrico se encuentra el punto de información respectivo.

Esta constituye una forma rápida de generar una imagen de $[S]$. Obviamente, al sacrificar la continuidad de la superficie ya no es posible estudiar las tendencias y gradientes del fenómeno.

Un programa que implementa la técnica anterior para producir imágenes de proximidad se dá a continuación:

```
          B 5 7 0 0     F O R T R A N    C O M P I L A T I O N  XIII. 0,
FILE  1=DATA, UNIT=READER
FILE  3=OUTP,UNIT=PRINTER
                                                                    START OF
      INTEGER VX(100),VY(100)
      DIMENSION U(50,50),V(100),D(100)
      WRITE(3,4)
    4 FORMAT(1H ,20X,"VALORES DE ENTRADA"//)
      READ(1,1) N
    1 FORMAT(I3)
      DO 2 I=1,N
    2 READ(1,3) VX(I),VY(I), V(I)
    3 FORMAT(2I3,2X,F8.0)
      DO 6 K=1,N
    6 U(VX(K),VY(K))=V(K)
      DO 9 I=1,50
      DO 9 J=1,50
      DO 8 K=1,N
      IF(I.EQ.VX(K).AND.J.EQ.VY(K)) GO TO 9
    8 CONTINUE
      DO 10 K=1,N
   10 D(K)= (I-VX(K))**2+(J-VY(K))**2
      MIN=D(1)
      MINK=1
      DO  11 K=2,N
      IF(MIN.LT.D(K))GO TO 11
      MIN=D(K)
      MINK=K
   11 CONTINUE
      U(I,J)= V(MINK)
    9 CONTINUE
C LA MATRIZ NUMERICA HA SIDO GENERADA
      WRITE(3,13)
   13 FORMAT("1")
      WRITE(3,12)  ((U(I,J),J=1,25),I=1,50)
      WRITE(3,13)
      WRITE(3,12)   ((U(I,J),J=26,50),I=1,50)
   12 FORMAT(" ",25F5.0)
      CALL PROXIM(U,V,N,VX,VY)
      CALL EXIT
      END
```

Los vectores que se leen al principio del programa, VX(I), VY(I) y V(I), contienen respectivamente el índice de fila, el de columna y un valor numérico. Las tres cantidades definen un punto del conjunto {D} . Este conjunto es unido con el conjunto {R} del retículo para producir la matriz U(I,J) a través del lazo DØ 6 K = 1, N. Los lazos DØ 9 I = 1,50; DØ 9 J = 1,50; DØ 8 K = 1, N y DØ 10 K = 1, N exploran cada elemento de la matriz U(I,J) para calcular la distancia a cada uno de los puntos de {D} . Las instrucciones siguientes calculan el mínimo del conjunto {D} y lo identifican plenamente mediante el subíndice MINK. El algorítmo termina asignando al elemento bajo análisis de la matriz U(I,J), el valor de su punto de información mas cercano V(MINK).

Una vez obtenida la matriz, esta pasa a ser procesada por la subrutina PROXIM, que utilizando algunos de los métodos de despliegue gráfico ya descrito, genera -

por la impresora de línea la proyección ortogonal correspondiente (ver figura - página No. 91). En la siguiente tabla se muestran los puntos de información - usados. N es número de secuencia; F Fila. C Columna; V Valor del Punto y CL Clase Estadística. Al fondo de la página aparece la escala de ennegrecimiento junto con la correspondencia de clases.

N	F	C	V	CL
1	45	9	30.000	2
2	25	40	40.000	3
3	9	13	0.000	1
4	1	18	90.000	5
5	25	93	165.000	10
6	5	25	100.000	6
7	3	25	0.000	1
8	3	40	150.000	9
9	40	17	90.000	5
10	47	16	140.000	8
11	47	16	140.000	8
12	12	33	5.000	1
13	14	36	10.000	1
14	4	47	10.000	1
15	30	30	50.000	3
16	10	40	125.000	7
17	40	10	60.000	4
18	33	60	163.000	10
19	25	25	50.000	2
20	35	35	110.000	7
21	39	45	35.000	2
22	45	42	10.000	1
23	45	48	45.000	3
24	17	15	75.000	5
25	19	45	85.000	5
26	5	10	95.000	6
27	29	45	105.000	6
28	37	10	10.000	1
29	20	33	130.000	8
30	21	9	125.000	7
31	12	60	120.000	7
32	5	100	25.000	2
33	2	115	175.000	10
34	45	100	180.000	10
35	36	89	10.000	1
36	2	75	0.000	1
37	45	114	23.000	2
38	25	100	0.000	1
			180.0000	0.0000

1	2	3	4	5	6	7	8	9	10
.	-	+	=	I	X	H	Ø	8	8

Existe otro tipo de representación para [S] de tipo discontinuo que implica la descripción cuidadosa de una serie de subcontornos de [S] , y a los cuales se les ha asignado un valor numérico representativo del fenómeno bajo estudio. En efecto, las divisiones existentes en una región geográfica generalmente son de tipo político (países, estados, distritos, etc.) o administrativo (zona de desarrollo). Interesa entonces poder visualizar el estado actual de una variable en cada subdivisión del área.

Se puede implementar la producción automática de este tipo de imagen a través de la impresora y el delineador digital. Para el primer caso los algorítmos son bastantes complicados y no se discutirán aquí. Para el segundo caso el principal problema consiste describir adecuadamente el contorno de cada subconjunto [S] usando los vértices que lo componen.

Los algorítmos que sirven para la descripción y procesamiento de vértices, líneas y superficies serán definidos en las páginas siguientes.

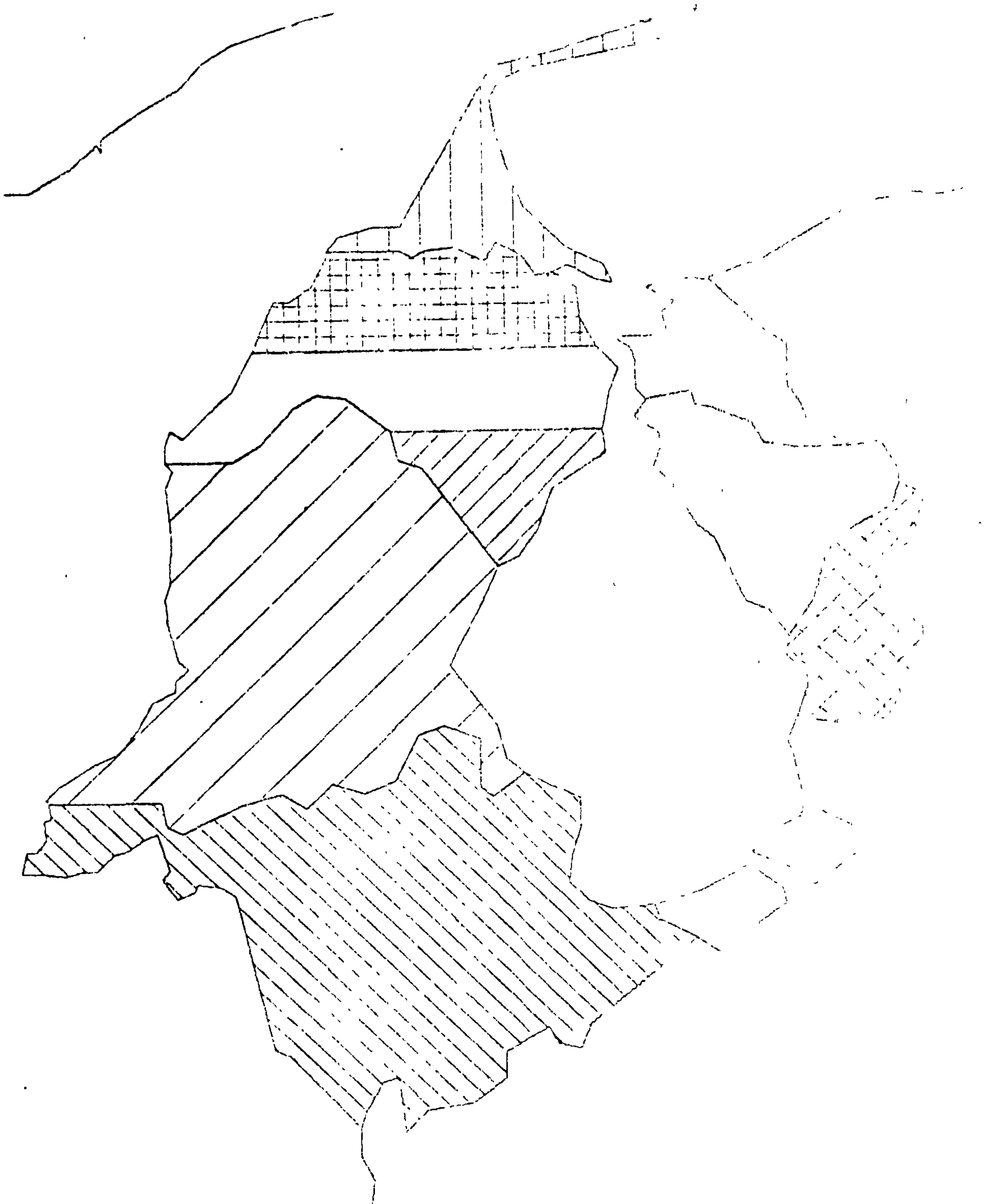
Una vez que se tiene cada contorno, este debe ser rayado en una o dos direcciones a fin de producir una graduación de tonos que van desde el blanco hasta el negro, - proporcional al valor de la variable. La subrutina básica de dibujo consiste en una serie de instrucciones que ejecuta 10 tipos de rayado diferentes sobre un contorno. Cada tipo de rayado es asignado - por un número en el rango [1, 10] y - que corresponde a la categorización otorgada a la información de entrada, es decir, a cada subcontorno se le dá una clase. Los límites entre clases pueden ser constantes (se obtienen de acuerdo a lo - descrito en la página 24) o variables (el usuario los proporciona de acuerdo a su - criterio).

En la figura de la página 94 puede verse el dibujo hecho por el "PLOTTER" de una -

serie de cons[illegible]

de acuerdo a [illegible]

va a cada políg[illegible]



Se hace evidente de que se trata de un mapa de conformidad del Estado Zulia y que los contornos corresponden a los 10 Distritos que lo componen.

2.6. Codificación y Procesamiento de Líneas y Elementos de Superficies

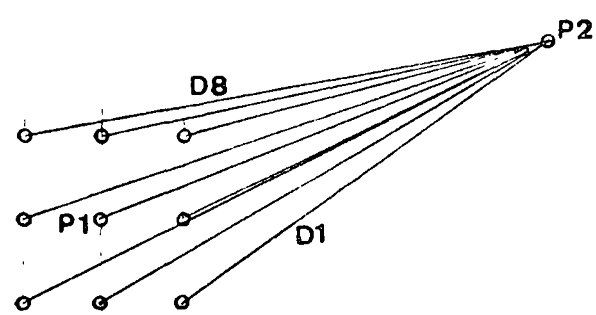
Se presentará brevemente en esta sección la forma de almacenar en la memoria del Computador las líneas y elementos de superficie que constituyen los diversos tipos de imágenes discutidos en las secciones anteriores. Las técnicas variarán grandemente según se trate de una impresora de línea, de una proyección ortogonal, perspectiva, etc.

2.6.1. Construcción de un Contorno sobre la Matriz [S]

El problema consiste en sobreimponer en una matriz un contorno o poligonal que delimite un área geográfica predeterminada. Todos los elementos que caen fuera del contorno son excluídos de procesamiento e impresión.

Como un contorno puede estar compuesto de centenares de elementos S(I,J), se necesita un método para recolectar solo los vértices de la poligonal para que un proceso automático interpole los puntos intermedios. La unión del conjunto {V} de vértices y {I} de puntos interpolados produce el conjunto {P} o poligonal en la forma de dos vectores o arreglos PI(I), PJ(I), I = 1,N con N-1 puntos diferentes (PI(1) = PI(N); PJ(1) = PJ(N)).

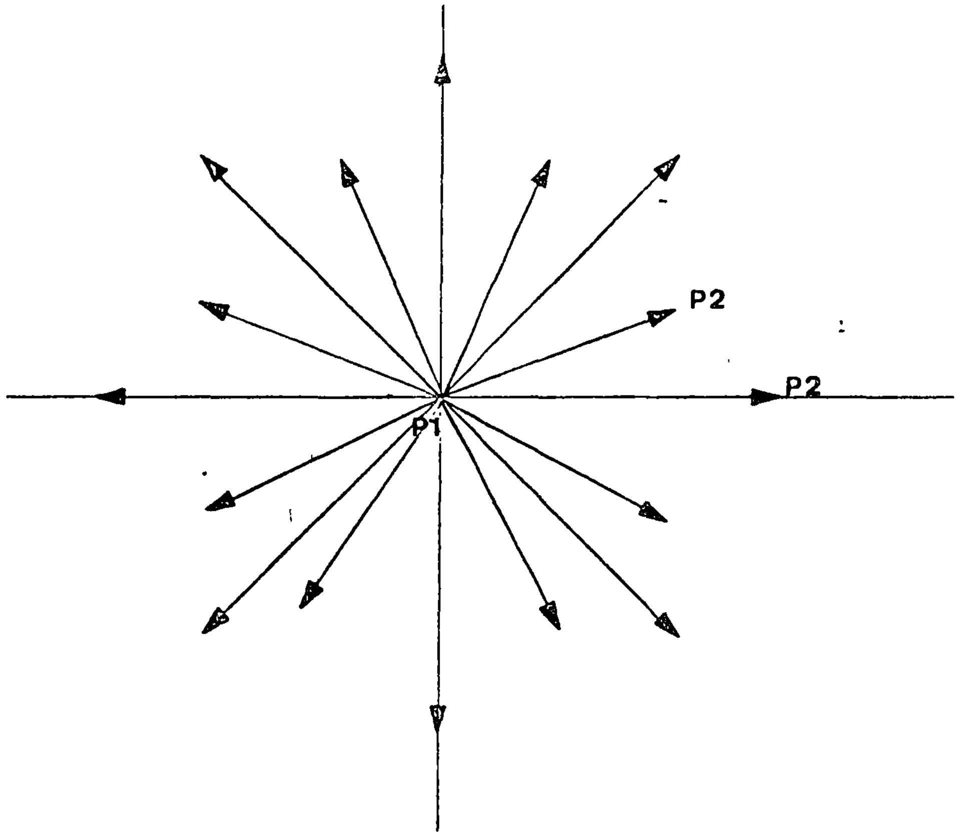
Un primer método implementado por el autor consiste en explorar los ocho puntos circundantes de un punto P1 y averiguar cual de ellos está mas cerca al punto P2 como lo indica la figura:



El punto más cercano pasa a ser ahora el nuevo P1 y se vuelve a repetir el proceso de cálculo de los ocho puntos circundantes hasta que $P1 \neq P2$. El procedimiento se ejecuta hasta que se cumpla $P1 = P(N)$, en cuyo caso ya se ha recorrido todos los puntos del contorno.

Al implementarse este algorítmo se pudo comprobar con sorpresa que los puntos seleccionados siguen trayectorias de 0º, 45º, 90º, 135º, 180º, 225º, 270º y 315º o sea que dos puntos P1, P2 dentro de la matriz quedan unidos por una secuencia de segmentos siguiendo las direcciones arriba descritas. Obviamente el procedimiento no es apto para los fines perseguidos.

Un segundo método, que representa una aplicación curiosa del Método de Monte Carlo a la Computación Gráfica, se describe a continuación: dos puntos P1 y P2 definen un vector cuya dirección puede ubicarse en 16 zonas notables de la circunferencia:



Ocho de ellos siguen las direcciones fijas 0º, 45º, 90º, 135º, 180º, 225º, 270º y - 315º. Los ocho restantes son variables - dentro de los límites impuestos por los fi jos. El algorítmo comienza por consiguiente. identificando la dirección en particular que sigue el vector que une dos elementos particulares P1. P2 de un contorno. - partiendo de los índices de fila y columna.

Los elementos de la matriz que unen en línea recta a P1 y P2 de coordenadas F1, C1 y F2, C2 respectivamente tienen la particularidad de que sus índices de fila y columna siguen una distribución probabilística uniforme. Es inmediato que utilizando un generador de números pseudoaleatorios se puede obtener el conjunto de índices necesarios para cada par sucesivo de puntos de una poligonal.

El programa mostrado en la página 101 comienza por leer las coordenadas del conjunto de puntos $\{P\}$. La subrutina CLAVE(F1, C1, F2, C2, L) es llamada para identificar a través del parámetro de retorno L = 1 hasta 16, que dirección sigue el vector P1, P2. El GØ TØ computado siguiente envia el control a la sección correspondiente encargada de calcular adecuadamente los límites LF y LC que van a ser utilizados por el generador de números aleatorios FUNTION RANDOM(A,B,0) para calcular los índices necesarios. Como los índices son generados uniformemente -

distribuidos pero en forma desordenada, es tos son ordenados de menor a mayor a través de la subrutina SORT1. El lazo iterativo DØ 96 M = 1, N - 1 asegura de que el proce so se lleve a cabo para los N puntos del - contorno.

En la página 105 puede observarse el resulta do obtenido para una matriz MATRIX(50,122) a la cual se le ha asignado 23 puntos de u- na polígonal. Esta ha sido dibujada por la impresora con el caracter X y el resto de - la matriz por signos.-.

En la página 106 aparece el diagrama de flu jo de la subrutina CLAVE, la cual analiza - el signo y magnitud de los vectores D1 = F2 - F1 y D2 = C2 - C1 para decidir la direc - ción del vector P1, P2.

```
          B 5 7 0 0    F O R T R A N    C O M P I L A T I O N   XIII.11,     FRIDA
$ CARD LIST HDL SINGLE LIMIT 10
                  SINGLE ILLEGAL CONSTRUCT ON DOLLAR CARD XXXX
FILE   5=OUT,UNIT=PRINTER
FILE   1=PRINTTR,UNIT=READER
                                                                  START OF SEGMENT
      DIMENSION  KF(99),KC(99),LC(99),LF(99),LIMF(700),LIMC(700),
     *MATRIX(30,122)
      ..RANDOM(1,10,523466512J5)
      READ(1,88)  KF(1),KC(1)
      READ(1,88)  KF(2),KC(2)
      DO 771 I=3,500
      READ(1,88)  KF(I),KC(I)
      IF(KF(I).EQ.KF(1).AND.KC(I).EQ.KC(1))  GO TO 772
  771 CONTINUE
  772 N=I
   88 FORMAT(2I3)
      DO 96 M=1,N-1
      F1=KF(M)
      F2=KF(M+1)
      C1=KC(M)
      C2=KC(M+1)
      CALL CLAVE(F1,C1,F2,C2,L)
      GO TO(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16),L
    1 N2=C2-C1-1
      DO 20 I=1,N2
      LC(I)=C1+I
   20 LF(I)=F1
      GO TO 888
    2 N2=C1-C2-1
      DO 21 I=1,N2
      [illegible]2+I
   21 [illegible]

    3 N[illegible]-1
      DO 22 I=1,N2
      LF(I)=F2+I
   22 LC(I)=C1
      GO TO 888
    4 N2=F2-F1-1
      DO 23 I=1,N2
      LF(I)=F1+I
   23 LC(I)=C1
      GO TO 888
    5 N2=F1-F2-1
      DO 24 I=1,N2
      LF(I)=F2+I
   24 LC(I)=C2-I
      GO TO 888
    6 N2=F2-F1-1
      DO 25 I=1,N2
      LF(I)=F1+I
   25 LC(I)=C1-I
      GO TO 888
    7 N2=C1-C2-1
      DO 26 I=1,N2
      LF(I)=F2+I
   26 LC (I)=C2+I
      GO TO 888
    8 N2=C2-C1-1
      DO 27 I=1,N2
```

```
      LF(I)=F1+I
   27 LC (I)=C1+I
      GO TO 888
    9 N2=F2-F1-1
      DO 50 I=1,N2
      LF(I)=F1+I
   50 LC(I)=RANDOM(C1,C2,0)
      CALL SORT1(LC,N2,1)
      GO TO 888
   10 N2=F1-F2-1
      DO 40 I=1,N2
      LF(I)=F2+I
   40 LC(I)=RANDOM(C2,C1,0)
      CALL SORT1(LC,N2,1) .
      GO TO 888
   11 N2=C2-C1-1
      DO 42 I=1,N2
      LC(I)=C1+I
   42 LF(I)=RANDOM(F1,F2,0)
      CALL SORT1(LF,N2,1)
      GO TO 888
   12 N2=C1-C2-1
      DO 43 I=1,N2
      LC(I)=C2+I
   43 LF(I)=RANDOM(F2,F1,0)
      CALL SORT1(LF,N2,1)
      GO TO 888
      [illegible]=F1-F2-1
      [illegible]7 I=1,N2
      LF(I)=F2+I
   47 LC(I)=RANDOM(C1,C2,0)
      CALL SORT1(LC,N2,2)
      GO TO 888
   14 N2=F2-F1-1
      DO 48 I=1,N2
      LF(I)=F1+I
   48 LC(I)=RANDOM(C2,C1,0)
      CALL SORT1(LC,N2,2)
      GO TO 888
   15 N2=C2-C1-1
      DO 44 I=1,N2
      LC(I)=C1+I
   44 LF(I)=RANDOM(F2,F1,0)
      CALL SORT1(LF,N2,2)
      GO TO 888
   16 N2=C1-C2-1
      DO 45 I=1,N2
      LC(I)=C2+I
   45 LF(I)=RANDOM(F1,F2,0)
      CALL SORT1(LF,N2,2)
  888 L1=L2+1
      L2=L2+N2
      DO 30 K=L1,L2
      LIMF(K)=LF(K-L1+1)
   30 LIMC(K)=LC(K-L1+1)
   39 WRITE(3,95) N,F1,C1,F2,C2,L
   95 FORMAT (6I4)
   96 CONTINUE
      DO 31 I=1,50
      DO 31 J=1,122
   31 MATRIX(I,J)="-"
```

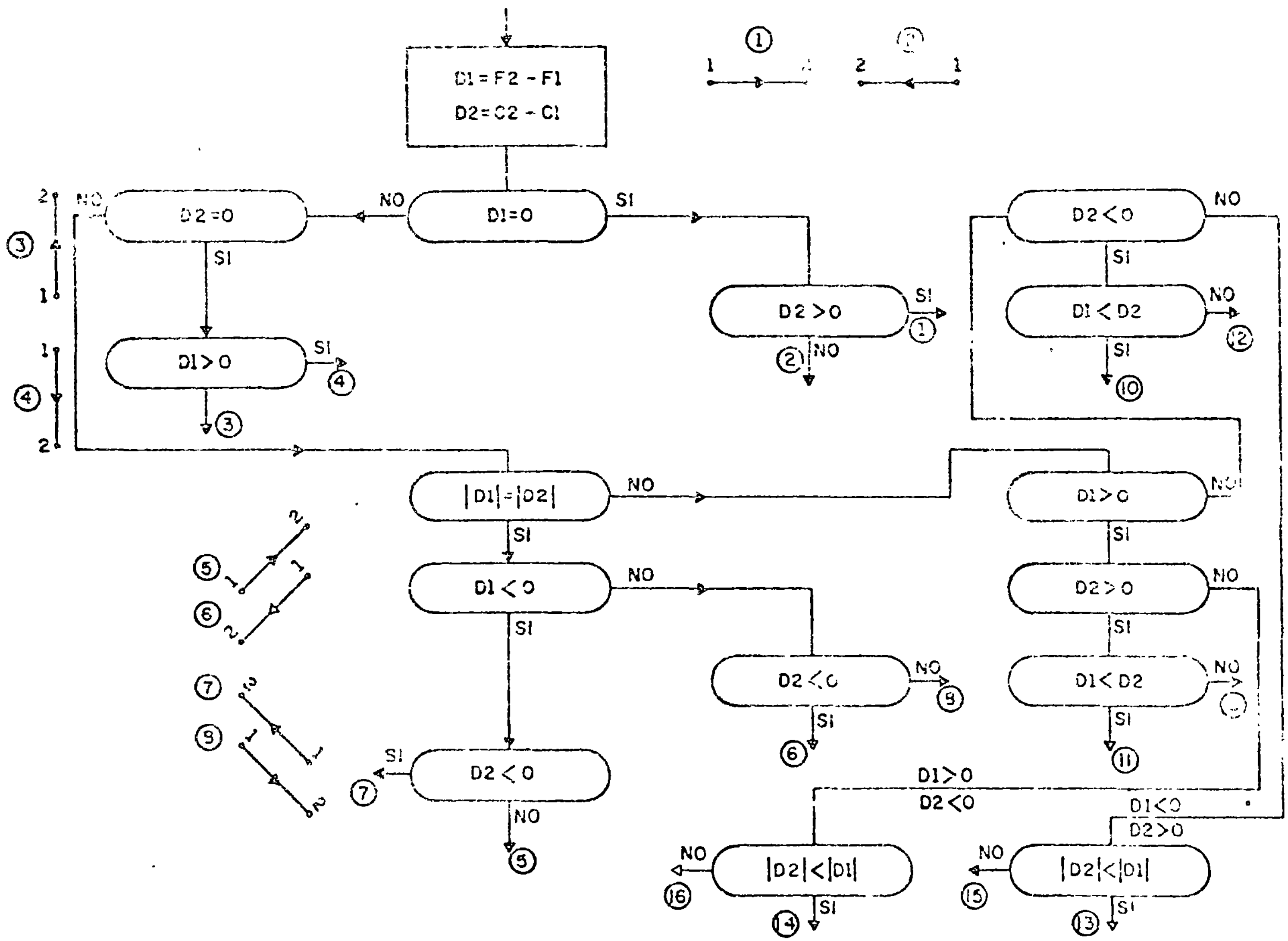
```
      DO 32 I=L2+1,L2+N
      LIMF(I)=RF(I-L2)
   32 LIMC(I)=RC(I-L2)
      DO 33 I=1,L2+N
   33 MATRIX(LIMF(I),LIMC(I))="X"
      WRITE(3,34) ((MATRIX(I,J),J=1,122),I=1,50)
   34 FORMAT("1"/(10X,122A1))
      STOP;END
```

SEGMENT
START OF SEGMENT

```
      SUBROUTINE CLAVE(F1,C1,F2,C2,L)
      D1=F2-F1
      D2=C2-C1
      IF(D1.EQ.0) GO TO 1
      IF(D2.EQ.0)  GO TO 2
      IF(ABS(D1).EQ.ABS(D2)) GO TO 3
      IF(D1.GT.0)  GO TO 4
      IF(D2.LT.0)  GO TO 5
      IF(ABS(D2).LT.ABS(D1))  GO TO 6
      L=15
      RETURN
    1 IF(D2.GT.0)  GO TO 100
      L=2
      RETURN
    2 IF(D1.GT.0)  GO TO 7
      L=3
      RETURN
    3 IF(D1.LT.0)  GO TO 8
      IF(D2.LT.0)  GO TO 9
      L=6
      RETURN
    4 IF(D2.GT.0)  GO TO 10
      IF(ABS(D2).LT.ABS(D1)) GO TO 11
      L=16
      RETURN
    5 IF(D1.LT.D2)  GO TO 12
      L=12
      RETURN
    6 L=13
      RETURN
  100 L=1
      RETURN
    7 L=4
      RETURN
    8 IF(D2.LT.0)  GO TO 13
      L=5
      RETURN
    9 L=8
      RETURN
   10 IF(D1.LT.D2)  GO TO 14
      L=9
      RETURN
   11 L=14
      RETURN
   12 L=10
      RETURN
   13 L=7
      RETURN
   14 L=11
      RETURN
      END
```

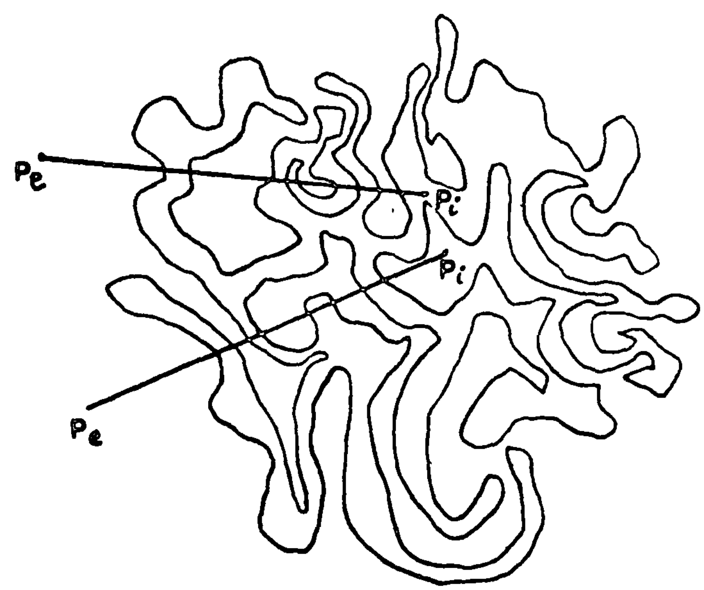
SEGMENT

```
                                                      START OF SEGMENT
      FUNCTION RANDOM (A,B,INIC)
      COMMON /AZAR/ I
      IF(INIC.NE.0) I=INIC
      I=CONCAT(0,I,13,13,35)
      I=I+13
      RANDOM=I/4866765968/84.*(B-A)+A
      RETURN,END
                                                           SEGMENT
                                                      START OF SEGMENT
      SUBROUTINE SORT1(A,N,IND)
      INTEGER A(100)
      K=N-1
      DO 6 J=1,N-2
      DO 5 I=1,K
      IF(A(I).LT.A(I+1)) GO TO 5
      TEMP=A(I+1)
      A(I+1)=A(I)
      A(I)=TEMP
    5 CONTINUE
      K=K-1
    6 CONTINUE
      IF(IND.EQ.1)  RETURN
      MIND=(N+1)/2
      DO 3 I=1,MIND
      TEMP=A(I)
      A(I)=A(N+1-I)
    3 A(N+1-I)=TEMP
      RETURN;END
                                                           SEGMENT
                                                           SEGMENT
                                                           SEGMENT
                                                      START OF SEGMENT
                                                           SEGMENT

NUMBER OF SYNTAX ERRORS DETECTED = 0.

PRT SIZE = 58;  TOTAL SEGMENT SIZE = 861 WORDS;  DISK SIZE = 37 SEGS;  NO. PRGI

ESTIMATED CORE STORAGE REQUIREMENT = 5760 WORDS;  COMPILATION TIME = 1 MIN, 7
```


D1 = F2 - F1
D2 = C2 - C1
D1 = 0
D2 = 0
D1 > 0
D2 > 0
\|D1\| = \|D2\|
D1 < 0
D2 < 0
D1 < D2
\|D2\| < \|D1\|
D1 > 0
D2 < 0
D1 < 0
D2 > 0
SI
NO

2.6.2. <u>Diferenciación entre Puntos Externos e Internos en un Contorno</u>

Una vez obtenidos los elementos del contorno debe buscarse alguna forma de poder detectar que objeto de poder eliminar todos los puntos externos a la poligonal. La idea básica a implementar consiste en la aplicación de un conocido teorema topológico y que dice que un punto se encuentra encerrado en un contorno si, una recta trazada a él desde otro punto cualquiera exterior corta el contorno un número impar de veces:



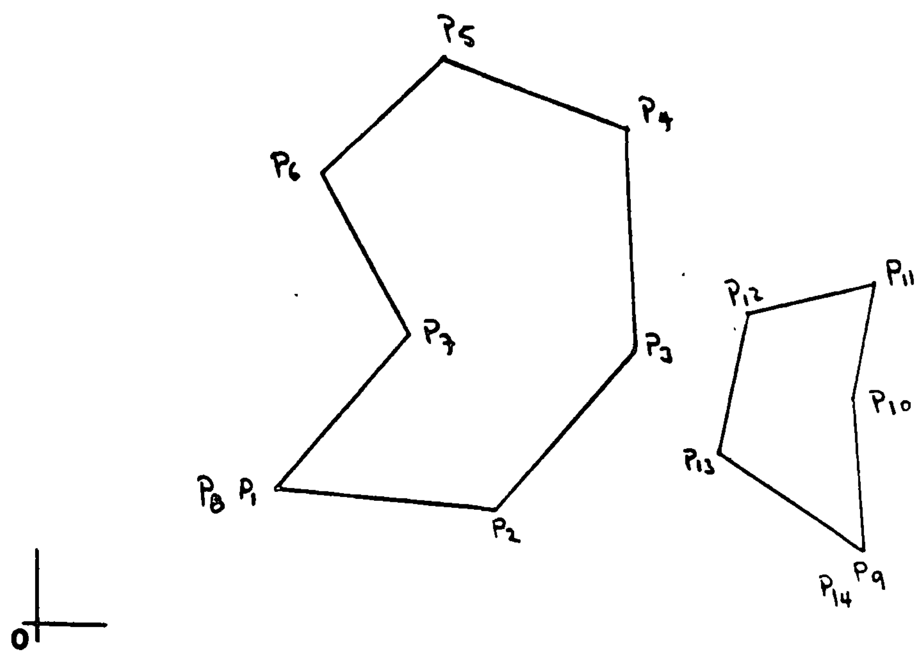
Para detectar si un elemento de una matriz cae en un contorno previamente definido - bastará con analizar la fila correspondiente del elemento. Se considera:que se traza una línea imaginaria desde el margen izquierdo del papel al elemento en cuestión ubicado en la fila H y columna Q. Explorando ahora los vectores que contienen la descripción del contorno se calcula cuantos elementos pertenecientes a la fila H - tienen un índice de columna menor que Q. Si el número obtenido es impar, el punto - se encuentra dentro del contorno,

2.6.3. Codificación de Líneas

La exposición presentada en este punto se aplicará principalmente al delineador digital, es decir ya no se considera el procesamiento de elementos finitos en una matriz sino el tratamiento de puntos y líneas a través de funciones continuas.

Un primer método para codificar una poligonal consiste en describir la secuencia de

puntos que la componen: para el caso de contornos no conectados se debe crear otro conjunto que describa la conectividad. Para implementar esta sencilla idea simplemente se hace uso de tres arreglos unidimensionales, uno para la coordenada X, otro para la coordenada Y, y otro para describir la conectividad. Por ejemplo los contornos:



Compuestos por 14 puntos de coordenadas: X1, X2...X14; Y1, Y2,...Y14 y el vector de conectividad 1,0,0,0,0,0,0,0,1,0,0,0,0,0 pueden ser procesados (Cálculo de Areas,

Centro de Gravedad, Dibujo, etc.) mediante las siguientes instrucciones:

```
     DIMENSION X(1000), Y(1000), K(1000)
     READ(1,1) N
     DØ 2 I = 1,N
2    READ(1,2) X(I), Y(I), K(I)
     .
     .
     .
     DØ 4 I = 1,N
4    CALL PLØT(x(I), Y(I), K(I)
     .
     .
     .
```

Un segundo método más eficiente consiste en utilizar direccionamiento indirecto. Para basta marcar cuidadosamente cada vertice de la poligonal con un número de secuencia. Este número indica la posición del punto en dos vectores X(I) e Y(I) que contienen las coordenadas. Los vertices de las figuras

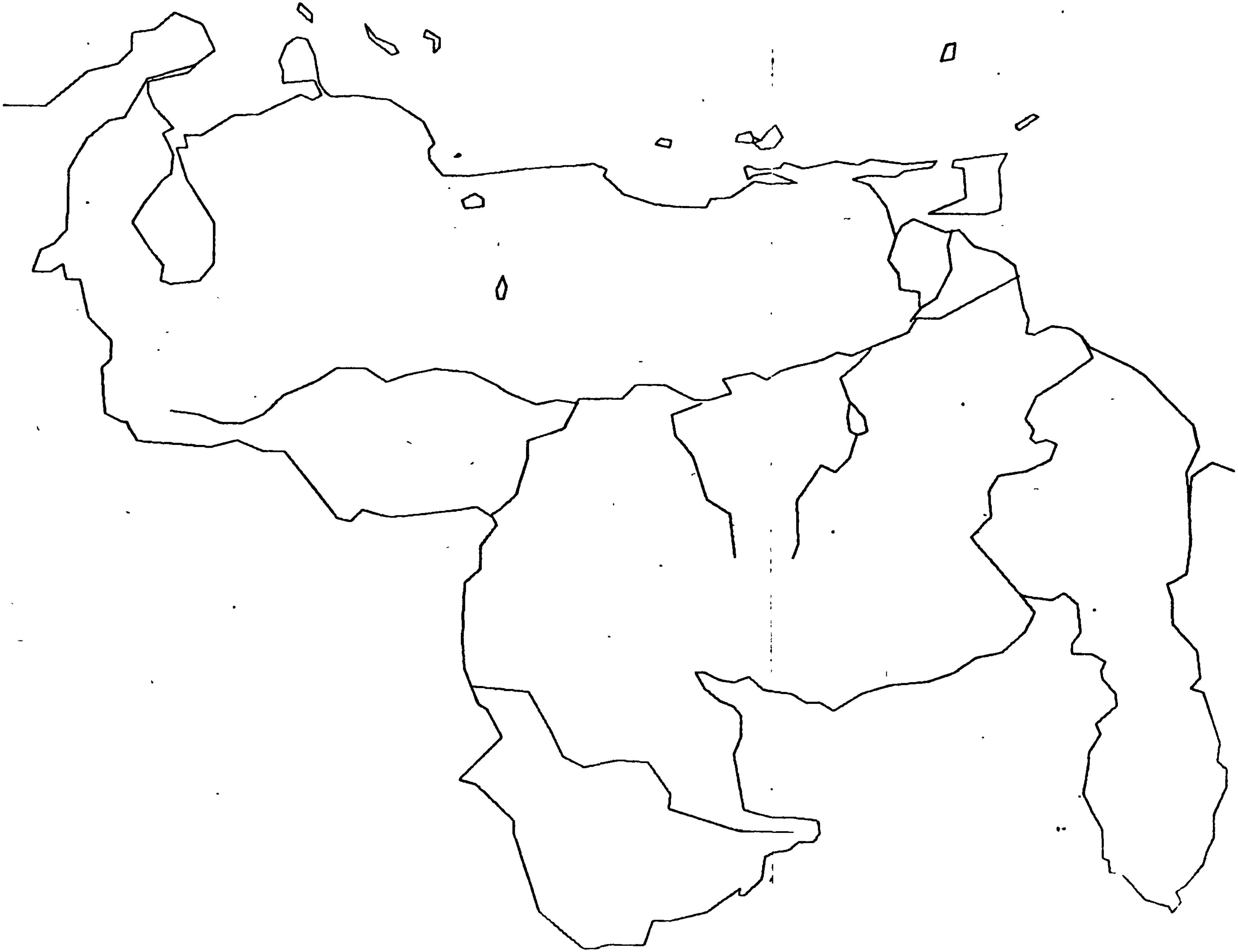
vendrán descritas por un vector que contiene una secuencia de números enteros indicadores de los puntos particulares que componen la poligonal. Para referirse a una figura se escribe X(P(I)), Y(P(I)), donde P(I) es el vector de contorno.

Un ejemplo del primer método lo constituye la subrutina VENEZUE(I) que sirve para dibujar un contorno de Venezuela. El parámetro I de entrada sirve para producir un dibujo simple sin división política interna (I = 1). Si I = 2 el mapa se genera con la división entre Estados. En las páginas 112 y 113 aparece el listado. Los dibujos producidos en las páginas 114 y 115 .

El subprograma comienza por leer las secuencias de coordenadas A(I) y B(I). El vector L(I) de conectividad se construye a continuación asignando el valor 3 a ciertas posiciones previamente predeterminadas. Una serie de llamadas a la subrutina PLOT construyen el contorno pedido.

```
     B 5 7 0 0   F O R T R A N   C O M P I L A T I O N   XIII.II,   TUESDAY, 09/03/74, 1010
FILE  1=VENE,UNIT=READER                                                R 0000
FILE  3=PRIN,UNIT=PRINTER                                               R 0000
FILE  9=PARIS,UNIT=TAPE,SAVE=1,RECORD=750                               R 0000
C PLOTLIB/PLOT VERSION 02 LEVEL 02 09 07 74 JH                          L 0667
                                                  START OF SEGMENT **********   2
      CALL PLOTS(I,I,9)                                                 R 0000
      CALL FACTOR(0.50)                                                 R 0001
      I=0                                                               R 0004
      CALL VENEZUE(I)                                                   R 0005
      CALL PLOT (0.,0.,999)                                             R 0005
      STOP                                                              R 0007
      END                                                               R 0008
                                                  SEGMENT    2 IS    13 LONG
                                                  START OF SEGMENT **********   3
      SUBROUTINE VENEZUE(KLAVE)                                         R 0000
      DIMENSION  A(783),B(783),L(783)                                   R 0000
      READ(1,3) N                                                       R 0000
    3 FORMAT(I3)                                                        R 0010
      READ(1,1) (A(I),I=1,N)                                            R 0010
      READ(1,1) (B(I),I=1,N)                                            R 0027
    1 FORMAT(14F5.0)                                                    R 0044
      DO 2 I=1,N                                                        R 0044
    2 L(I)=2                                                            R 0050
      L(1  )=3                                                          R 0052
      L(214)=3                                                          R 0053
      L(225)=3                                                          R 0054
      L(253)=3                                                          R 0056
      L(259)=3                                                          R 0057
      L(271)=3                                                          R 0058
      L(274)=3                                                          R 0059
      L(288)=3                                                          R 0061
      L(303)=3                                                          R 0062
      L(315)=3                                                          R 0063
      L(320)=3                                                          R 0064
      L(320)=3                                                          R 0066
      L(327)=3                                                          R 0067
      L(334)=3                                                          R 0068
      L(340)=3                                                          R 0069
      L(346)=3                                                          R 0071
      L(351)=3                                                          R 0072
      L(354)=3                                                          R 0073
      L(362)=3                                                          R 0074
      L(372)=3                                                          R 0076
      L(379)=3                                                          R 0077
      L(386)=3                                                          R 0078
      L(404)=3                                                          R 0079
      L(456)=3                                                          R 0081
      L(458)=3                                                          R 0082
      L(460)=3                                                          R 0083
      L(462)=3                                                          R 0084
      L(464)=3                                                          R 0086
      L(466)=3                                                          R 0087
      L(468)=3                                                          R 0088
      L(470)=3                                                          R 0089
      L(472)=3                                                          R 0091
      L(474)=3                                                          R 0092
      L(476)=3                                                          R 0093
      L(478)=3                                                          R 0094
      L(480)=3                                                          R 0096
      L(482)=3                                                          R 0097
```

```
      L(484)=3                                                  R--009
      L(486)=3                                                  R--009
      L(519)=3                                                  R--010
      L(522)=3                                                  R--010
      L(539)=3                                                  R--010
      L(567)=3                                                  R--010
      L(600)=3                                                  R--010
      L(606)=3                                                  R--010
      L(615)=3                                                  R--010
      L(622)=3                                                  R--010
      L(654)=3                                                  R--011
      L(668)=3                                                  R--011
      L(696)=3                                                  R--011
      L(722)=3                                                  R--011
      L(742)=3                                                  R--011
      L(757)=3                                                  R--0117
      L(765)=3                                                  R--0118
      L(770)=3                                                  R--0119
      L(775)=3                                                  R--0121
      L(778)=3                                                  R--0122
      DO 5 I=1,455                                              R--0123
      WRITE(3,6) A(I),B(I),L(I),I                               R--0128
    6 FORMAT(10X,F8.1,F8.1,I5,I4)                               R--0146
      A(I)=-A(I)/2.54                                           R--0146
      B(I)=-B(I)/2.54                                           R--0152
    5 CALL PLOT(A(I),B(I),L(I))                                 R--0156
      CALL PLOT(0.,0.,3)                                        R--0164
      CALL PLOT(0.,0.,-3)                                       R--0165
      DO 7 I=456,485                                            R--0167
      WRITE(3,6) A(I),B(I),L(I),I                               R--0172
      A(I)=-A(I)/2.54                                           R--0191
      B(I)=-B(I)/2.54                                           R--0196
    7 CALL PLOT(A(I),B(I),L(I))                                 R--0202
      IF (KLAVE.EQ.1) RETURN                                    R--0208
      CALL PLOT(0.,0.,3)                                        R--0212
      CALL PLOT(0.,0.,-3)                                       R--0214
      DO 8 I=486,782                                            R--0215
      WRITE(3,6) A(I),B(I),L(I),I                               R--0220
      A(I)=-A(I)/2.54                                           R--0239
      B(I)=-B(I)/2.54                                           R--0244
    8 CALL PLOT(A(I),B(I),L(I))                                 R--0250
      RETURN                                                    R--0258
      END                                                       R--0259
                                                   SEGMENT   3 IS
                                                   SEGMENT   4 IS
                                                   SEGMENT   5 IS
                                                   SEGMENT   6 IS
                                            START OF SEGMENT *******
                                                   SEGMENT  13 IS
                                                   SEGMENT   7 IS

NUMBER OF SYNTAX ERRORS DETECTED = 0.

PRT SIZE = 86; TOTAL SEGMENT SIZE = 1203 WORDS; DISK SIZE = 51 SEGS; NO. PRGM. SEGS =

ESTIMATED CORE STORAGE REQUIREMENT = 10496 WORDS; COMPILATION TIME = 45 SECS; NO. CARD
```



2.6.4. Codificación de Superficies

De la misma forma que un conjunto de segmentos de recta conectados por sus extremos componen un contorno, un conjunto de superficies unidas por los bordes constituyen un poliedro. En el caso bidimensional una curva continua se aproxima mediante segmentos. Para el caso tridimensional una superficie se aproxima mediante planos.

Para describir y manejar adecuadamente sólidos por intermedio de un Computador se reducen estos a un conjunto de planos elementales. Cada plano es entonces descrito por un vector de contorno que contiene el número asignado a cada vértice. Las coordenadas rectangulares de cada punto se describen a través de tres vectores X(I), Y(I), Z(I). Nuevamente se usa el direccionamiento indirecto para procesar cada elemento de superficie.

Para el dibujo de un poliedro habrá que transformar las coordenadas tridimensiona

les a coordenadas bidimensionales por medio de proyecciones oblícuas, perspectiva, axonométricas, etc.

Con respecto al problema de la visibilidad, es decir, la elimación de todas aquellas líneas que estan ocultas por el poliedro a un observador, se puede decir que es una cuestión compleja de resolver en general.

Una solución partícular, para el caso de un poliedro*se dá a continuación: como cada elemento de superficie viene descrito por los puntos que lo forman, se pueden calcular fácilmente el vector unitario normal que apunta hacía el exterior del sólido. Si el ángulo que forma la dirección en que mira el observador y la normal al plano es menor o igual a noventa grados, la superficie no es visible y no se dibuja:

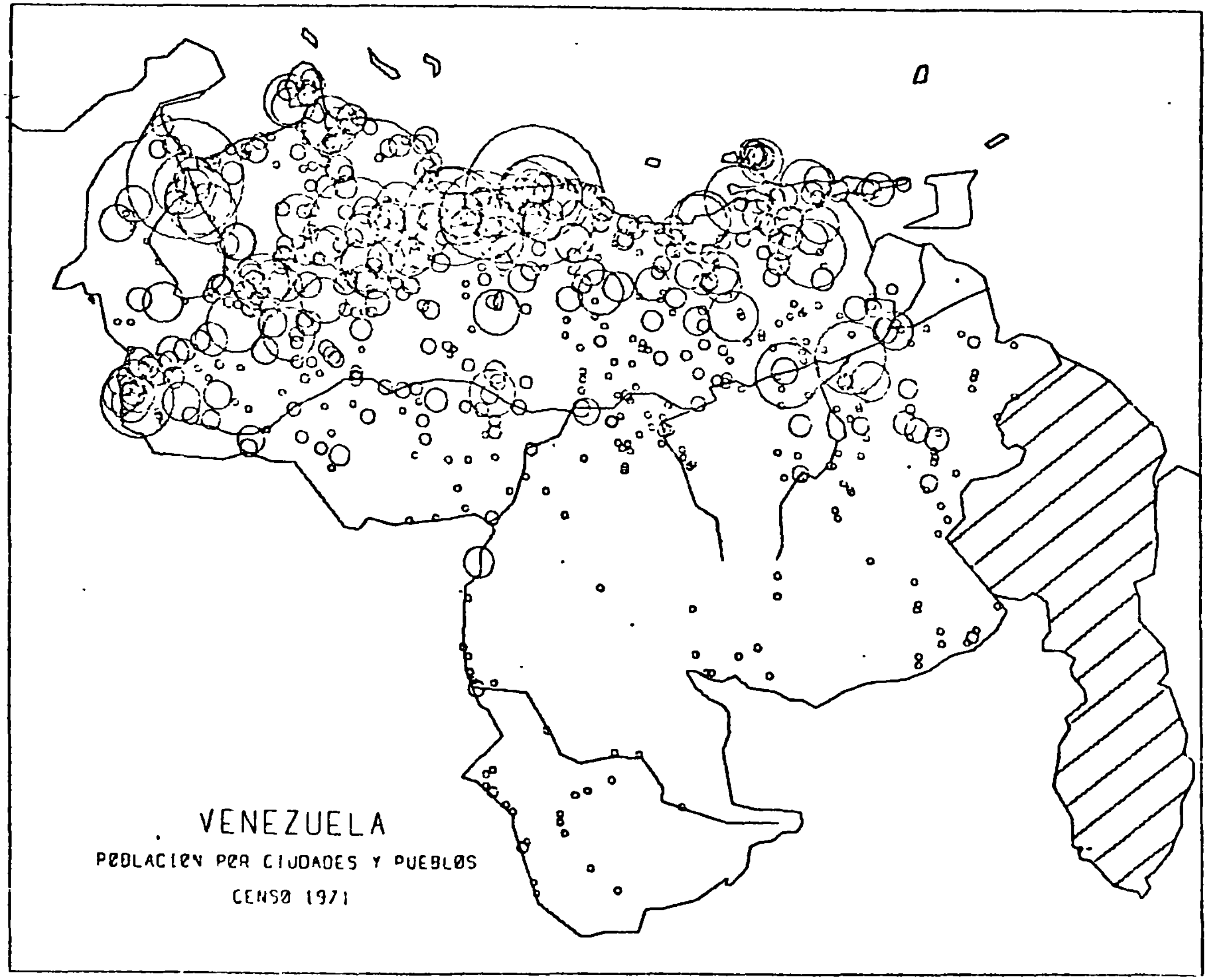
*Convexo

La implementación de esta idea usando cálculo vectorial y direccionamiento indirecto no presenta ninguna dificultad.

2.7. Otras Formas de Representación

Otra manera de visualizar el tipo de variable analizado hasta ahora consiste en asignar un símbolo o figura geométrica de área o volumen proyectado, proporcional al valor medido, sobre el contorno geográfico en partícular, El algorítmo ubicará los puntos donde existe información y generará los comandos necesarios para dibujar el símbolo.

El dibujo de la página 119 representa un mapa de Venezuela de la población de 617 pueblos y ciudades para 1.971. En este caso el programa dibujó el contorno del país y luego trazó 617 círculos cuya área era proporcional a la variable. Conceptualmente no existe dificultad en la implementación del procedimiento, pues se trata de hacer una subrutina que dibuje círculos de diámetros variables dadas las coordenadas de los mismos.

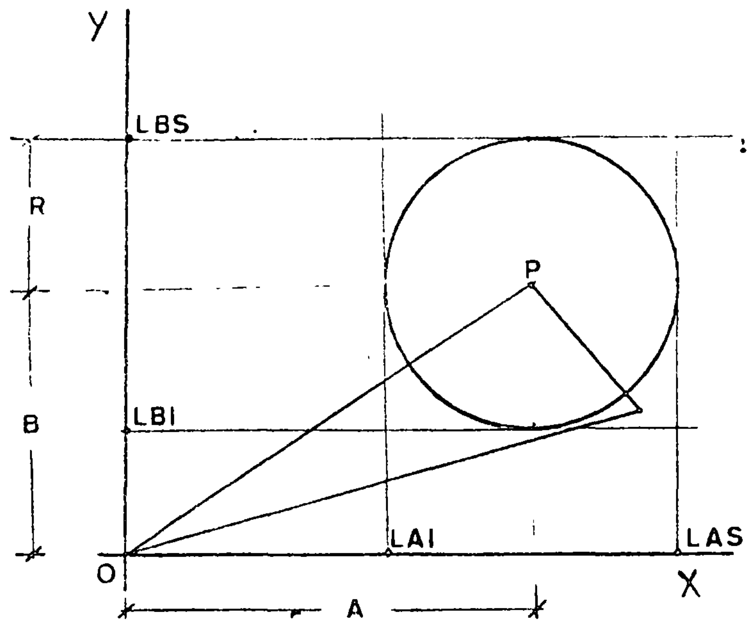

VENEZUELA
POBLACION POR CIUDADES Y PUEBLOS
CENSO 1971

2.7.1. El Método de Monte Carlo

Se puede utilizar esta técnica para implementar algunos procedimientos en el campo de la Computación Gráfica. Un primer ejemplo fué dado en la página 97 donde el autor lo utilizó para construir un contorno a partir de sus vértices, usando un generador de números aleatorios.

El método puede ser utilizado por consiguiente en el despliegue gráfico de variables espacialmente distribuidas. Para ello debe idearse alguna forma de aglomerar algun tipo de símbolo proporcionalmente al valor bajo estudio. Se puede producir por ejemplo, las coordenadas X e Y de N puntos utilizando para ello una subrutina de números aleatorios como la mostrada en la página 104. Se puede utilizar una distribución probabilística uniforme o de tipo normal. Obviamente que para el segundo caso los puntos tenderán a concentrarse hacia el centro del círculo que contenga el conjunto de caracteres empleados.

Graficamente se tiene:



En donde A y B son las coordenadas del punto con información alrededor del cual se van a dibujar al azar N caracteres.

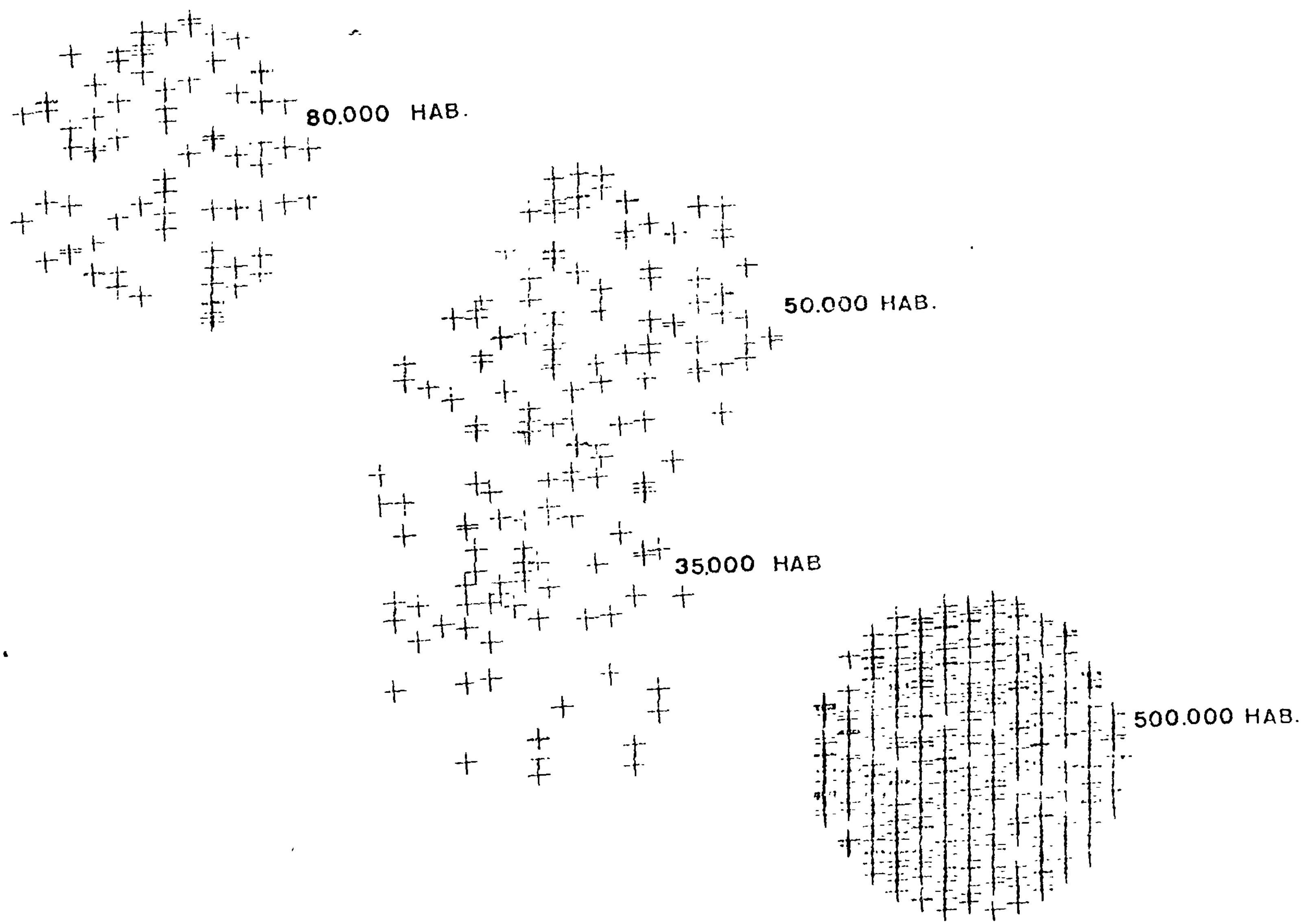
En la página 123 puede examinarse el programa principal y la subrutina ALEATO(A, B, N, INC). Se comienza el proceso por leer las coordenadas de los puntos respecto a un origen en centímetros junto al valor numérico del fenómeno. Para nuestro caso se seleccionó ficticiamente la población de varias ciudades y se tomó la relación de -

un caracter por cada 1000 habitantes. A continuación se preparan adecuadamente los parámetros A, B, M e INC dentro del lazo DO 5 I = 1, N; A y B son las coordenadas del círculo de contención; M es el número de caracteres de posición aleatoria a generar; INC es un incremento de dos unidades asignado a la semilla del generador de números al azar FUNCTIØN RANDOM. Una vez que el control del programa ha sido transferido a la subrutina ALEATO se calculan los límites superiores e inferiores de las coordenadas X e Y a generar. Cada vez que se crea un nuevo punto, se chequea que esté dentro de un círculo de radio una pulgada. Si cae fuera se rechaza. Si cae dentro se acepta, incrementándose el contador I y verificando si se han generado los M valores necesarios.

En la página 125 puede examinarse el dibujo resultante ejecutado por el delineador.

```
        B 5 7 0 0    F O R T R A N    C O M P I L A T I O N   XIII.11.
  $ CARD LIST HOL SINGLE LIMIT 10
                   SINGLE ILLEGAL CONSTRUCT ON DOLLAR CARD XXXX
  FILE   3=OUT,UNIT=PRINTER
  FILE   1=AZAR    ,UNIT=READER
  FILE   9=ESE,UNIT=TAPE,SAVE=1,RECORD=750
  C PLOTLIB/PLOT VERSION 02 LEVEL 02 09 07 74 JM
  C PLOTLIB/SYMBOL VERSION 01 LEVEL 04 010472 PA
                                                              START OI
        DIMENSION CX(100),CY(100),V(100),NSYMB(100)
        CALL FACTOR(1.0)
        CALL PLOTS (I,T,9)
  C   LECTURA DE LOS DATOS
        READ(1,2) N
        DO 1 I=1,N
      1 READ(1,3) CX(I),CY(I),V(I)
  C   TRANSFORMACION DE LA VARIABLE V EN UN NUMERO ENTERO PROPORCIONAL N.
  C   CADA 1000 HABITANTES ESTAN REPRESENTADOS POR UN CARACTERS.
        DO 4 I=1,N
      4 NSYMB(I)= V(I)/1000
  C   INICIO DEL PROCESO DEL DIBUJO.
        DO 5 I=1,N
        INC=INC+2
        A=CX(I)
        B=CY(I)
        M=NSYMB(I)
      5 CALL ALEATO(A,B,M,INC)
      2 FORMAT(I3)
      3 FORMAT(3F10.0)
        CALL PLOT(0.,0.,999)
        STOP
        END
                                                              START
        SUBROUTINE ALEATO(A,B,N,INC)
        REAL LIX,LIY,LSX,LSY
        DIMENSION IBCD(1)
        IBCD(1)=7
  C  ESTE SUBPROGRAMA DIBUJA UN CONJUNTO DE SIMBOLOS CENTRADOS POR
  C  GENERACION ALEATORIA DE LAS COORDENDAS X E Y . A Y B SON LAS
  C  COORDENADAS DEL PUNTO DE INFORMACION.  N ES EI NUMERO DE
  C  SIMBOLOS A DIBUJAR Y INC ES LA SEMILLA DEL GENERADOR DE NUMEROS
  C  ALEATORIOS. EL RADIO DEL CIRCULO DE CONTENCION ES UNA PULGADA.
        A=A/2.54
```

```
      B=B/2.54
      DATA RADIO/1/
      LSX=A+RADIO
      LSY=B+RADIO
      LIX=A-RADIO
      LIY=B-RADIO
      X= RANDOM(0,1,55555555555+INC)
      I=1
    2 X=RANDOM(LIX,LSX,0)
      Y=RANDOM(LIY,LSY,0)
      D= SQRT((X-A)**2+(Y-B)**2)
      IF(D.GT.1)  GO TO 2
      CALL SYMBOL( X,Y,0.15,IBCD,0.0,-1)
      I=I+1
      IF(I.GT.N)RETURN
      GO TO 2
      END
                                                              SEGMENT
                                                     START OF SEGMENT
      FUNCTION RANDOM (A,B,INIC)
      COMMON /AZAR/ I
      IF(INIC.NE.0) I=INIC
      I=CONCAT(0,I,13,13,35)
      I=I*13
      RANDOM=I/4466765598784.*(B-A)+A
      RETURN;END
                                                              SEGMENT
                                                              SEGMENT
                                                              SEGMENT
                                                              SEGMENT
                                                     START OF SEGMENT
                                                              SEGMENT
                                                              SEGMENT

UMBER OF SYNTAX ERRORS DETECTED = 0.

NT SIZE = 98;  TOTAL SEGMENT SIZE = 1727 WORDS;  DISK SIZE = 71 SEGS;  NO. PRGM.

STIMATED CORE STORAGE REQUIREMENT = 9408 WORDS;  COMPILATION TIME = 3 MIN, 13 SE
```


80.000 HAB.
50.000 HAB.
35.000 HAB
500.000 HAB.

2.7.2. Secuencias Animadas

Una de las extensiones lógicas que se pueden hacer de la automatización del proceso de representar variables espaciales, consiste en generar una secuencia de imágenes de un fenómeno en función del tiempo o la posición. Si el conjunto de dibujos obtenidos es fotografiado cuadro a cuadro por una cámara de cine, la película resultante presenta la evolución dinámica de la variable.

La secuencia de dibujos que van de las páginas 70 a la 78 muestran una superficie con cuatro puntos de información en proyección oblícua y que dá la impresión de estar rotando con respecto a un observador. La secuencia completa filmada por el autor abarcó un giro completo de 360º en incrementos de 10º. El total de 36 imágenes fueron generadas en un tiempo de una hora con cuarenta minutos. El proceso de filmación en si tomó una hora.

Pueden combinarse variaciones de posición

como la descrita anteriormente con variaciones en el tiempo para imprimir una mayor sensación de realismo. En tal caso la programación se complica pues además de intervenir matrices de rotación de coordenadas, hay que introducir ecuaciones dependientes del tiempo.

3. Conclusiones y Comentarios.-

No hay duda de que a escala mundial se han realizado numerosos trabajos relacionados con la aplicación de los computadores digitales al manejo y representación de variables espacialmente distribuidas. Algunas razones que motivan esta tendencia son:

a. Proliferación de Centros de Computación Académicos. Durante los últimos catorces años, los Centros de Computación han pasado a formar una parte indispensable de toda Universidad. La disponibilidad y alta accesibilidad por parte del investigador al computador, ha generado una presión que obliga al desarrollo de nuevos métodos automáticos para resolver problemas en todos los campos de la Ciencia y la Técnica.

b. Creación de una nueva tecnología en dispositivos de salida. Desde la aparición del primer Delineador Digital en 1.957 hasta nuestro días, se han creado una serie de dispositivos de representación gráfica altamente sofísticados. Entre ellos podemos mencionar el Delineador Digital de Mesa ("FLATBEB"), la Pantalla de Rayos Catódicos ("CRT") (que permite el desarrollo de la Computación Gráfica Interactiva),

la Pantalla de Televisión, etc. Unido a lo anterior se encuentran los costos cada vez menores.

c. Precisión, rápidez y economía en la elaboración automática de imágenes. Un computador adecúadamente programado puede producir grandes cantidades de gráficos, mapas, dibujos, e tc. en un corto intervalo de tiempo. Si además se toma en cuenta la circunstancia de que el manejo, procesamiento y representación de millones de números es imposible de realizar por un dibujante, se puede concluir que eventualmente serân los computadores los encargados de efectuar los procesos de representación, sustituyendo o colocando en otro nivel al personal especializado.

d. Tendencia a la digitalización en la recolección de información. El gran avance tecnológico alcanzado en el campo de los sensores remotos y locales junto al perfeccionamiento y abaratamiento de los circuitos digitales, causan un incremento en la recolección de información capaz de ser leída directamente por un computador para su procesamiento e interpretación (tarjetas y cintas perforadas, cintas magnéticas).

e. Grandes avances logrados en los campos del procesamiento de imágenes y reconocimiento de formas. Como parte de una disciplina mas amplia llamada Inteligencia Artificial, se han logrado avances en el campo de la construcción automática de mapas físicos a partir del examen directo de aerofotos. Es posible anticipar que eventualmente serán los computadores los encargados de elaborar mapas topográficos usando lectoras ópticas como canales de entrada.

En el ámbito nacional, es aún muy poco lo logrado en esta área. Existen desarrollos en estado experimental como los parcialmente presentados en este trabajo. Para pasar a la implementación de la fase operativa se debe obtener la colaboración en actividades de experimentación, por parte de personas del campo académico y de investigación. Solo un continuo proceso cíclico de utilización - evaluación - crítica - corrección podrá crear el cuerpo coherente de conocimientos necesarios para poder usar con efectividad y con una visión realista el Software a ser creado y el Hardware ya disponible.

REFERENCIAS

1.- PARIS C., Bernardo A., "Algunas Aplicaciones del Computador Digital a la Cartografía Temática". II Jornadas de Computación Técnica. Trabajo No. 12. Caracas, Septiembre de 1.974.

2.- "Sistema de Delineación Digital". Manual del Usuario. Sección de Sistemas, Centro de Computación, Facultad de Ingeniería - U.C.V. (1972).

3.- KANEFF S., "Picture Language Machines". Academic Press. (1.970).

4.- JAMES, M.L. "Applied Numerical Methods for Digital Computation with Fortran". Pág. 463, International Textbook Company. (1.967).

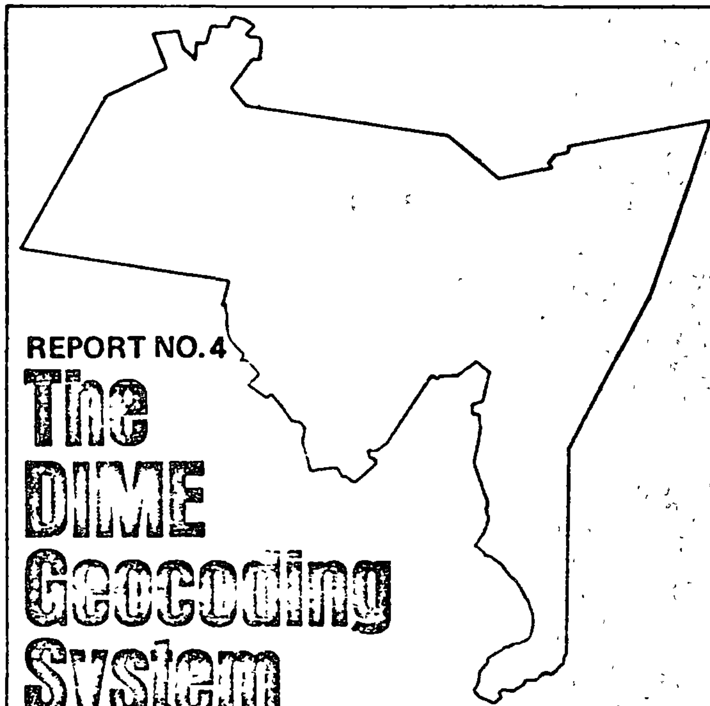
5.- ROBINSON A., "Elements of Cartography". (1.969). Pág. 150.

6.- PEUKER T., "Computer Cartography". Resource Paper No. 17. Association of the American Geographers. (1.972).

7.- MUEHRCKE P., "Thematic Cartography". Resource Paper No. 19. Association of the American Geographers. (1.972).

8.- "The Red Book". Laboratory for Computer Graphics and Spatial Analysis. Harvard University, Cambridge. Mass. (1.973).

Census Use Study



REPORT NO. 4
The DIME Geocoding System

U.S. DEPARTMENT OF COMMERCE
Maurice H. Stans, Secretary
Rocco C. Siciliano, Under Secretary
Harold C. Passer,
Assistant Secretary for Economic Affairs

Issued July 1970

BUREAU OF THE CENSUS
George Hay Brown, Director

CONTENTS

Introduction

This chapter describes the Census Bureau's development of a computerized geographic system for use in conducting the 1970 census.

This report has been prepared to introduce and explain the DIME system to a three-part audience: (1) Individuals interested in technical developments in geographic coding, (2) organizations within standard metropolitan statistical areas (SMSA's) participating in or eligible to use the products of the Census Bureau's address coding guide/Dual Independent Map Encoding (ACG/DIME) program, and (3) organizations interested in creating their own DIME file.

For individuals interested in technical developments in geographic coding, this report describes the experiences of the Census Bureau in developing a computer file for relating addresses on census questionnaires to census geographic areas for tabulating the 1970 census. It describes early efforts to create such a file—the address coding guide (ACG) and the development of the Dual Independent Map Encoding (DIME) technique for creating geographic base files. It also describes the census geographic base file (GBF),[1] in which the ACG and DIME concepts are combined. The report also describes uses of a DIME file, creation of a DIME file, and the conceptual basis of the DIME system.

For those organizations within SMSA's which currently possess or are in the initial stages of creating a GBF, or those that will be able to obtain an already created GBF for their own use, this report describes the uses of the file in computer mapping, address and area coding, network and node analysis, adjacency analysis, and other uses. It also discusses the question of file maintenance and expansion.

For those organizations contemplating the creation of a DIME file, this report provides an overview of the system. In conjunction with a clerical procedures manual, a computer procedures manual, and various computer programs, this report provides the complete documentation needed to create a DIME geographic base file.[2]

1970 Census Procedures

In 1970 the Bureau of the Census conducted the Nineteenth Decennial Census of Population and Housing by a combination of two methods: a mail canvass in the large urban areas of the country and a house-to-house enumeration in the remainder of the country.

In 145 of the 233 SMSA's and in certain adjoining areas the mail canvass procedure was used to enumerate approximately 60 percent of the Nation's population. Instead of a visit to each household by an enumerator, census questionnaires were mailed to all residents. These questionnaires were addressed using a computerized mailing list derived from commercial

[1] The DIME file described in this report is essentially the same as the Census Bureau's GBF. Minor differences occur in formats and in lengths of particular fields, but the theoretical framework is identical.

[2] This report, the two procedures manuals, and the computer programs constitute the New Haven Census Use Study computer program package titled *DIME: A Geographic Base File System.*

sources and corrected and updated by the Post Office Department.

Before the census, addresses were checked by the Post Office Department. Addresses on cards were given to the postal carriers on each route with instructions to add, delete, or correct the addresses as required. Once corrected, the computerized list of addresses was used to generate mailing labels for the mail-out questionnaires. A second check was made at the time of delivery, with the necessary corrections in the listings and mailing pieces made manually.

For those portions of the mail canvass areas which were not covered by city delivery postal service, address lists were prepared by Census Bureau field personnel. These addresses were also corrected and updated by the Post Office Department at the time of delivery. This file of addresses was not computerized and the mailing pieces were addressed manually prior to the mail-out.

Immediately prior to census day—April 1, 1970—the questionnaires were delivered to all households. Householders were asked to complete the questionnaire and mail it to a local office. For those householders not mailing back the questionnaire or erroneously completing any of the items on the questionnaire, a followup was conducted by telephone or by personal visit by an enumerator.

The remainder of the country was enumerated by the traditional house-to-house canvassing procedure. In the urbanized portion of the 88 nonmail SMSA's and in those smaller cities which have contracted with the Census Bureau for block statistics, each questionnaire will be coded to its census block. In the remainder of the nonmail areas, the enumeration district (ED) will be the smallest geographic area to which a questionnaire is coded.

Geographic Coding for the Census

As in previous censuses, geographic codes for each household in the 1970 nonmail enumeration areas were determined and coded by the enumerator during the house-to-house enumeration process. In areas scheduled for block tabulations, the enumerator entered the census block number from a census map on each questionnaire during the enumeration process. For the balance of the nonmail areas, the enumerator identified the ED of each household and coded the ED number on the household's questionnaire. From a combination of codes for these "lowest common denominator" areas (ED's and/or blocks), codes for larger areas such as census tracts, townships, counties, and States were developed.

For the 145 mail census SMSA's, a method was needed to code individual addresses on the mailing list to specific geographic areas for tabulation purposes. A master coding file has been developed to code approximately 40 million mail addresses by computer to the appropriate geographic areas. The coding file developed is called the address coding guide (ACG).

The Bureau decided that the best way to create this massive file (or series of 145 files) was to—

1. Prepare a preliminary file in a standard format for each mail SMSA.
2. Send the individual SMSA files on standard coding forms to a local agency in the SMSA for review and correction, or review the file at the Census Bureau using available reference sources.
3. Edit the files by computer to obtain the most accurate files possible.

To create the preliminary files, the Bureau contracted with commercial firms for mailing lists and city directories. From these lists and directories and Bureau source files, computer records containing street names, address ranges for each block side, intersecting street names, and various geographic codes (such as State, county, congressional district, municipality, ZIP code, and census tract) were created.

To prepare the files for review and correction by either local cooperating agencies or by a centralized Census Bureau coding staff this information was computer-printed on FOSDIC worksheets.[3] Once prepared, the FOSDIC worksheets were sent to the

[3]FOSDIC (Film Optical Sensing Device for Input to Computers) is a coding method whereby prearranged small circles or spaces on a form or worksheet are coded using a "fill-in" coding method. FOSDIC worksheets, once coded, are microfilmed and the worksheet codes converted directly to codes on computer tape. In a loose sense the filled-in spaces on the worksheet serve the same purpose as punched holes in a punchcard.

cooperating agencies for review, correction, updating, and coding. The Department of Housing and Urban Development provided "701" funds to qualified local agencies to defray up to two-thirds of the local costs of this operation. Highway planning and research funds, administered by the U.S. Department of Transportation, were also available for this purpose.

For each block side within the coding area, the following codes were reviewed, and corrected or updated: Census tract, census block, beginning and ending address numbers,[4] area code (a municipality code), congressional district, voting wards (optional), and an optional or local code field. When review was completed, the worksheets were clerically edited by the local agency and a sample of the work was independently verified as a quality control measure. Units of work not passing quality control were reviewed and corrected. The completed worksheets were then returned to the Bureau for processing.

At the Bureau, the worksheets were microfilmed and the information on the worksheets was transferred to magnetic tape by the FOSDIC equipment. Using the information from the FOSDIC worksheets, the original file was updated. Computer edit routines were applied to the file and clerical corrections were inserted, as necessary. Coding errors were corrected by the Bureau staff when adequate reference documents were available. Other errors were referred to the local cooperating agency for correction. After edit and correction routines were completed for an SMSA, the file was made available to non-Bureau users at reproduction cost. Immediately prior to the census, the ACG files were updated from information supplied by the Post Office Department and by the commercial mailing list sources. The ACG files were then considered final and were used to code the census questionnaires geographically. They will also be available at the cost of reproduction to non-Bureau users. Figure 1 illustrates the basic steps necessary to create an ACG.

ACG files consist of block side records for all streets within the coded area. Nonstreet features, such as municipal boundaries, rivers, and railroad tracks are not included in the file. A block side is one side of a street between two intersections on that side of the street. A block side for a dead-end street is one side of the street from its beginning intersection to the dead end. Figure 2 illustrates some typical block sides.

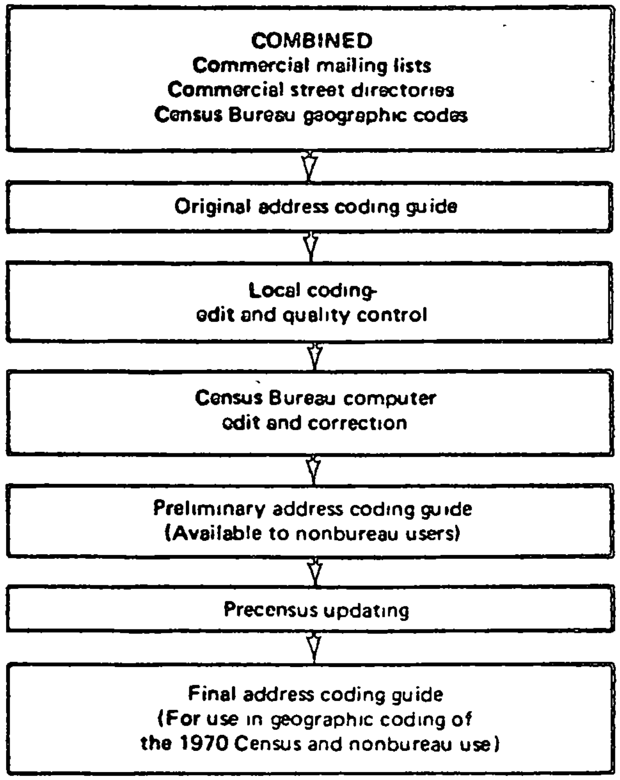


Figure 1. ACG creation process

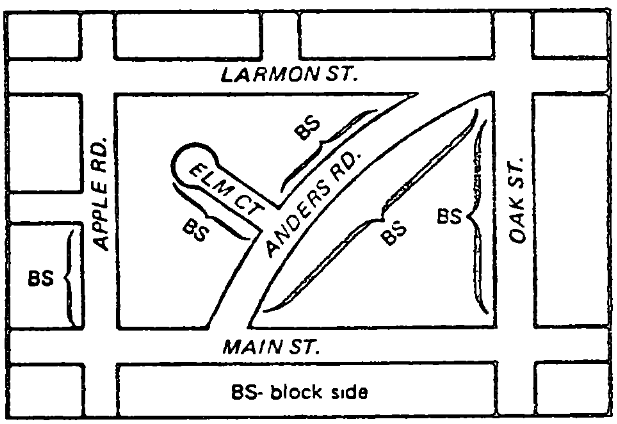


Figure 2. Typical block sides

Each ACG block side record contains the following codes: State, county, minor civil division (or census county division), place, ZIP code, 1970 census tract, street (includes street direction, name of street,

[4]For the purposes of this report, beginning and ending or lowest and highest address numbers refer to the lowest or highest *possible* address numbers, rather than actual address numbers, for each block side or street segment.

street type, and a street serial number), low and high address numbers, census block number, SMSA, district office (Bureau administrative code), area code (a code used in place of the minor civil division and place codes, above), optional code (local code or census serial number), ward (election district), annex (area annexed between the 1960 and 1970 censuses), congressional district, Post Office administrative codes, and serial number (unique number for each block side).

Development of DIME by the Census Use Study

The first ACG reviewed locally was the one for the New Haven, Conn. SMSA. This ACG was required to process the special census which was conducted in April 1967 to test the mail-out/mail-back procedures for the 1970 census. Local review and coding of the FOSDIC worksheets was completed in the fall of 1966.

The special census also provided a current data file for use in the research and development activities of the Census Use Study.[5] One of the primary research activities of the Use Study involved computer mapping.[6] A prerequisite for such research was the availability of a geographic base file with coordinates to permit the assigning of coordinates to data files for mapping. Since this type of file did not exist for New Haven, the Census Use Study attempted to create from the ACG a geographic file with coordinates.

Three alternative approaches were available for assigning coordinates to the ACG: (1) Assign coordinates for the approximate center point (centroid) of a block, (2) assign coordinates for the centroid of a block side, and (3) assign coordinates for both ends of each block side. The third alternative was selected because it provided the capability of plotting a street or network map that could be used for display or editing purposes. It also provided the ability to map by shading data values within the actual configuration of a block. The other two alternatives did not provide these capabilities. However, the selection of the third alternative made it possible to compute coordinates for the other two approaches, i.e., block centroid and block side centroid.

Portions of the New Haven ACG were tested using the third alternative. Coordinates for the beginning and ending of each ACG block side as shown on the Census Bureau's single-line metropolitan maps were measured and recorded using a semiautomatic device called a "Coordinate Locator" built by the Bureau. This process is called "digitizing." Coordinates digitized for each ACG block side were converted to state plane grid coordinates and added to the ACG file as part of each ACG block side record.

Substantial problems were encountered during the testing of the digitizing process. Since the ACG is a file of streets only, nonstreet features existing in the coded area such as railroad tracks, rivers, lake shores, and municipal boundaries were not included in the file and therefore could not be digitized. The absence of nonstreet features which formed block boundaries made it impossible to accurately plot block maps or calculate the area of blocks. This absence also limited the usefulness of a plotted map for display or analysis.

Another problem was that the digitizing process itself was inefficient: for the typical intersection (two streets crossing), eight readings had to be made on the coordinate locator, one for each of the block side ends forming the intersection. When such an intersection was plotted, the plot usually showed more than one of the plotted points at the intersection in slightly differing locations. This discrepancy was due to variability introduced manually in digitizing the point eight times and also to error introduced by mechanical slippage in the digitizer.

Still another problem was the inability to digitize curved streets accurately. In the ACG, each block side is coded from intersection to intersection without regard to curves that may occur along the block side. As a result, when a curved block side was digitized and plotted, it plotted out as a straight line. This occasionally led to plotted curved streets crossing other plotted streets incorrectly and the depicting of nonexistent intersections.

A substantial number of coding errors and omissions were detected through the New Haven ACG test as it was used by the Census Use Study. The discovery of these errors, not detected by the existing ACG clerical and computer edits, and the coordinate assignment problems discussed above, led to the

---

[5] See Census Use Study report, *General Description*.
[6] See Census Use Study report, *Computer Mapping*.

conclusion that alternative approaches should be explored. Although the ACG was sufficiently accurate for the geographic coding of census questionnaires, the Census Use Study concluded that improving it for other purposes, such as computer mapping, would not be practical.

The problem of developing a new system was referred to the Technical Steering Group (TSG) of the Census Bureau Advisory Committee on Small Area Data, the body overseeing Census Use Study activities. A subcommittee was established to explore other methods of automated geographic coding or to conduct original research, if necessary, to develop a geographic base file which could be used efficiently for computer mapping and geographic analysis.

The subcommittee first considered extant methods of creating geographic base files. Then it developed a proposal for constructing a geographic base file using graph theory as the conceptual framework. Since a linear graph can be described as a series of straight line segments in a plane connected at vertices, it is readily apparent that a single line map is a form of linear graph. Figure 3 illustrates this point.

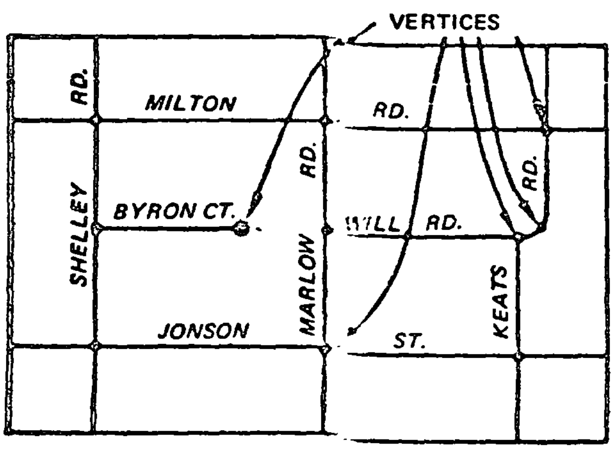


Figure 3. Graph elements of a single line map

Each street, river, railroad track, municipal boundary, or other map feature can be considered as one or more straight line segments. Curved lines can be divided into a series of straight line segments. When features intersect or when straight line segments change direction, vertices are formed. Using a concept derived from graph theory, a method was devised for creating a geographic base file where each vertex, intersection, or node,[7] as well as each line segment and intervening enclosed area, could be uniquely identified. Also each could be identified in terms of its place within the entire region encompassed by the file and in relation to adjacent nodes, line segments, and enclosed areas. The entire region encompassed by the geographic base file could be viewed as a series of interrelated nodes, lines, and enclosed areas. Each line segment is identified by its name as taken from the source map, e.g., Main Street or Green River, or by the node numbers at each end of the line segment. Nodes are numbered sequentially and uniquely. Enclosed areas are numbered using a systematic series of block numbers or other areal identifiers, or they can be described by the nodes or line segments which bound them.

The creation of such a file of line segments, nodes, and enclosed areas allows for the editing of the file by computer using an algorithm derived from graph theory. When coded, the three elements—segments, nodes, and enclosed areas—can be formed into separate incidence matrices, e.g., line segment-node, line segment-enclosed area, and enclosed area-node, by computer. The relationship established within these incidence matrices provided for the development of the computer editing system.

Such a file, as a representation of a map, can be digitized and plotted by computer to produce a complete replica of the source map. Often it is impractical or unnecessary to correct every edit reject or to reproduce faithfully all curved features. However, a perfect file can be produced at a reasonable additional cost.

The system that developed from this research activity was named Dual Independent Map Encoding (DIME) because the basic file is created by coding two independent incidence matrices from the source map.

From this theoretical base, the Census Use Study staff developed procedures and created a DIME file for New Haven. The same type of reference sources used to create the ACG—Census Bureau metropolitan

[7] The terms vertex, intersection, and node are considered to be essentially the same for the purposes of this paper, node is the term used because of its wide use in transportation planning and other geographic applications.

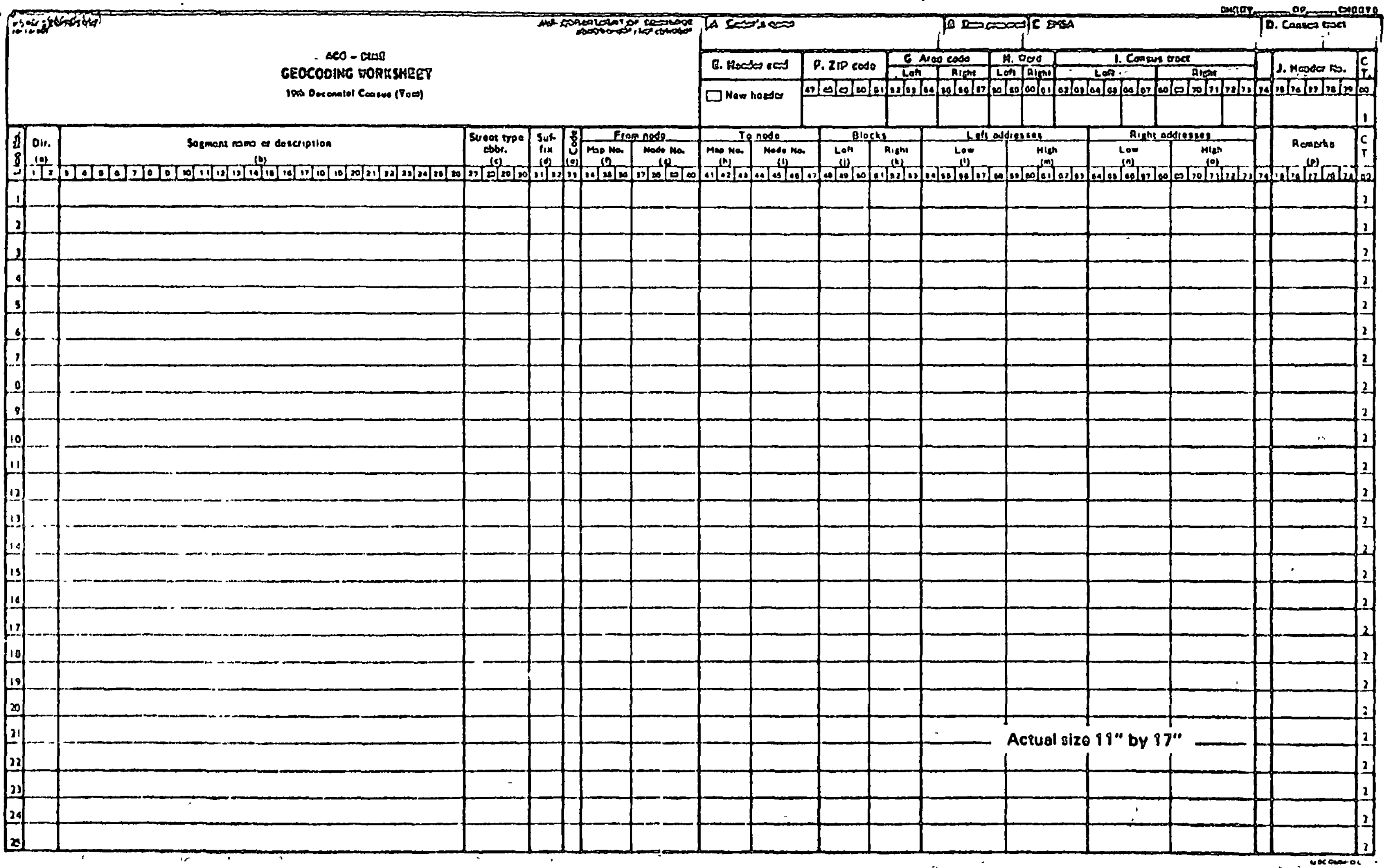


Figure 4. DIME coding worksheet

maps and local address maps or listings—was employed. Only one major change was made to the metropolitan maps prior to DIME coding—each node on the map was assigned a unique identification number. The coding form used was a standard punchcard transcription worksheet. A replica of the final coding form developed for the system is shown as figure 4.

While an ACG is constructed on a block-side basis, a DIME file is constructed on a street-segment basis. Each ACG record contains the appropriate codes for *one* side of a street between two intersections on that side of the street. Each DIME segment record contains the appropriate codes for *both* sides of a street between two nodes.

Essentially the same information is contained in both systems—street name, address number ranges, census block numbers, census tract numbers, and other geographic codes. The DIME system has two additional codes: node numbers and left/right orientation.[8] The DIME system also contains all meaningful nonstreet features such as rivers, municipal boundaries, shorelines, and railroad tracks shown on the metropolitan maps. These are not contained in the ACG system. Nodes are placed at sharp curves in streets or other features so that such curves can be adequately described by a series of straight line segments when plotted by computer.

Worksheets were coded and keypunched; the data were put on magnetic tape. The file was computer edited and edit rejects were corrected. After two or three edit cycles, the file was considered sufficiently accurate. The metropolitan maps used to code the DIME file were then digitized and the node numbers and coordinates were merged into the DIME file. The file was then ready for use in computer mapping.[9] The file was also used for address and area coding, as discussed in chapter II

[8]The left/right orientation code separates the codes for one side of a line segment from the other.

[9]The computer mapping experiments conducted are described fully in the Census Use Study report, *Computer Mapping*.

Further Experiments

To examine the applicability of DIME to ongoing programs of the Bureau, and to develop standard coding procedures and Bureau capabilities in the use of the system, further research was initiated.

A coding manual, forms, and procedures based on the New Haven experience were developed; and Roanoke, Va. was chosen as a test site for further DIME development activities. The Roanoke City Planning Commission agreed to supply coding clerks. The Use Study staff trained the clerks and supervised the coding operation.

After the local coding operation was completed, the coding worksheets were keypunched and the punchcards were converted to magnetic tape. The file was edited using DIME edit techniques. Rejects were corrected and inserted in the file. The coding maps were digitized with the Bureau's digitizing equipment. Nodes and coordinates from this process were inserted into the DIME file. Test maps were prepared on the Bureau's Calcomp plotter. Coordinate errors detected on the maps were corrected and the file updated. The file was then considered complete.

Further testing of coding procedures was conducted at Binghamton, N.Y., by the Broome County Planning Board. Areal coverage for this test was extended beyond the central city to the county boundary.

After training by Census Use Study staff, local personnel completed the coding operation with only nominal monitoring by Census Use Study staff. The Census Bureau's metropolitan maps were used for the urbanized portion of the county. County highway maps were used for the remainder. The manuals and worksheets used at Roanoke were revised and a coding manual was introduced for this test.

When coding and quality control operations were completed, the worksheets were keypunched and computer edited. The maps used for coding were digitized. When completed, the coordinate information was merged into the DIME file at the Census Bureau. Test maps were plotted and the file was corrected to complete the Broome County test.

The final test leading to the development of the Census Use Study's computer program package, *DIME: A Geographic Base File System*, was conducted in the Greenville, S.C. SMSA by the Greenville County Planning Commission. Prior to the test, the coding and supervisors manuals and the coding and administrative forms were refined according to findings of the previous tests. However, this test was designed not only to evaluate the DIME computer program package but also to develop materials which could be used to create geographic base files for the nonmail census SMSA's.[10] As a result, materials which could be adapted to both uses were developed. Geographic coverage was limited to the city postal delivery portion of the urbanized area of the SMSA, as in the ACG program.

As in Binghamton, supervisory training was conducted by Census Use staff members. The supervisor, in turn, trained the coding staff. Biweekly monitoring visits were made by Census staff. Upon completion of coding and local quality control measures, the worksheets were returned to the Census Bureau. Keypunching, editing, and corrections were completed, and the file became the prototype for the nonmail census ACG/DIME geographic base file program.

Concurrent with the Greenville test, an experiment was conducted to examine the feasibility of merging an independently created DIME file without addresses for Madison, Wis. with the ACG for Madison. The DIME file was created and edited by the Use Study staff; the ACG file was coded by a local agency and processed at the Bureau as part of the standard ACG program The metropolitan maps were used as the base for creating both files. Theoretically it should have been possible to merge these two files into an ACG/DIME file However, a substantial number of records could not be merged. For example, of the 2,345 block records which could be merged, 1,104 records had at least one difference from one file to the other file. An analysis of the differences revealed that 19 percent were street name spelling differences, 3 percent were DIME errors, 47 percent were ACG errors, 9 percent were errors in both files, and 22 percent were ambiguities in the metropolitan maps. Reconciling the various differences took a substantial amount of time. In fact, more time was spent on this process than would be required to create a DIME file with addresses for Madison using the Greenville approach. As a result, the merging of separate ACG and DIME files by computer was discarded.

However, the desirability of inserting DIME features into the ACG's increased as it became apparent that the ACG's could contain a considerable number of errors that were not detectable by existing ACG edits. The benefits of the DIME edits and the potential uses of DIME files for computer mapping and other uses were also becoming widely known. Inquiries as to the availability of DIME files in the mail census SMSA's were received by the Bureau from local, State, and Federal agencies.

After the close of the ACG/DIME merging experiment, further research into adding DIME features to the ACG's by computer was undertaken by both the Census Use Study and the Bureau's Geography Division. The Geography Division experimented with inserting unique, computer-generated node numbers into a specially prepared ACG. The attempt was only partially successful and was abandoned.

At approximately the same time, the Census Use Study investigated the possibility of adding DIME features, i.e., node number and left/right parity status, to a specially prepared ACG block side file by clerical procedures. The following procedure was developed for adding these features:

1. Metropolitan maps would be node numbered as in DIME; i.e., all intersections and points along curved features would be numbered uniquely.

2. A specially formatted worksheet printout version of the ACG allowing for the clerical insertion of node numbers and left/right parity status for each block side in the file would be prepared.

3. Nodes and parity status would be coded on the worksheet. Additional records would be coded for nonstreet features (nonstreet features were not coded in the ACG) and for block sides which had to be split into two or more pieces because of curves or other reasons.

4. After coding was completed, the added codes (and unique identification numbers for each record) would be keypunched and merged into the original ACG file. The resulting "segment side" records would then be merged to form segment records.

[10] Nonmail census geographic base files will be discussed in greater detail in the next section of this chapter.

5. The resulting file would be edited using the DIME edits.

6. Edit rejects would be corrected.

7. The metropolitan maps used in coding would be digitized to create a file of node records—each node record would contain the node number and its coordinates.

8. The node record file would be merged with the DIME file and would thus be in final form for release to appropriate agencies.

A generalized flow chart of this operation is shown in figure 5.

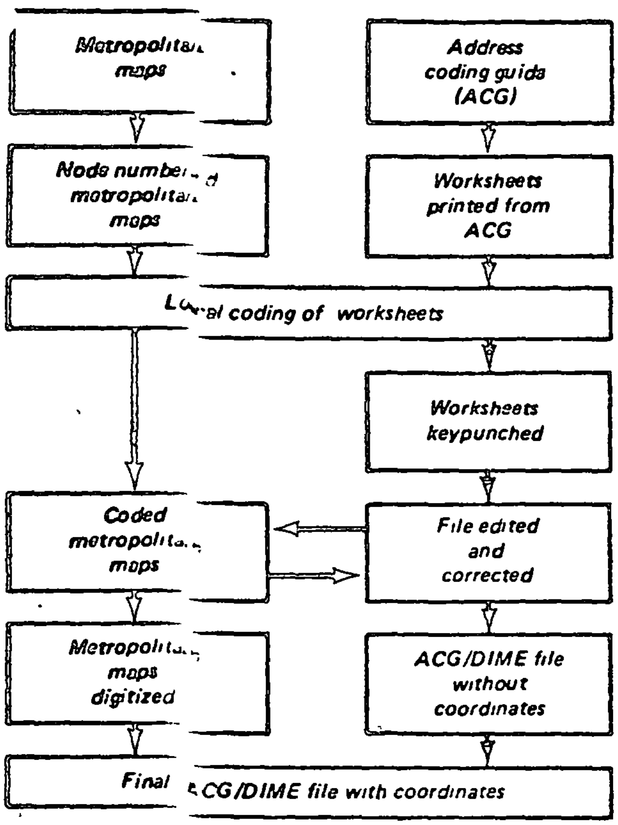


Figure 5. ACG/DIME creation process

This technique was tested using the ACG for a small number of census tracts in Madison, Wis. Testing showed the system was workable and sufficiently economical to warrant its use in a nationwide program. The Bureau accepted it for the program for inserting DIME features into the ACG's.

When the technique was accepted, a task force was created to conduct a large-scale test and prepare necessary clerical manuals and computer programs. This test was conducted successfully in Madison, Wis. under field conditions using draft manuals. The file was processed using the steps outlined above. The ACG/DIME program was then turned over to the Bureau's operating divisions for implementation.

The ACG/DIME Geographic Base File Program

As discussed earlier in this chapter, the 1970 census is being conducted using two different enumeration methods. In the major urban areas of the country (approximately 60 percent of the population), the census will be conducted using the mail-out/mail-back technique. The remainder of the country will be enumerated by traditional house-to-house canvassing methods. The Bureau has established one geographic base file program for the 145 mail census SMSA's and a separate program for the 88 nonmail census SMSA's. Different methods are being used to create ACG/DIME geographic base files in each program, although the completed files are essentially identical. In the mail census areas, the ACG's are being "improved" by adding DIME features. In the nonmail census areas, ACG/DIME files are being developed "from scratch." A listing of the mail and nonmail census SMSA's and the availability of ACG and ACG/DIME files is provided in appendix A.

Mail Census Geographic Base File Program

For the SMSA's included in the mail census, ACG's have been prepared for the central city (or cities) of the SMSA and the surrounding urbanized area, generally for the portion of the SMSA in which mail is delivered to a street address (such as 115 Main St.). The Bureau has required this areal coverage because each census questionnaire must be coded to its proper geographic area for tabulation purposes. Geographic coding of the questionnaires is accomplished by computer matching the address to which a questionnaire is mailed (115 Main St.) to its matching address range in the ACG (101-199 Main St.). Once this match is accomplished, the requisite geographic codes in the ACG are appended to the computer-generated record for the questionnaire.

The ACG/DIME geographic base file program in the mail census SMSA's (called the ACG Improvement Program) originated when it became obvious to the Bureau, and other Federal, State, and local agencies, that it would be desirable to add DIME features to the already existing ACG's. Federal agencies, including the Department of Housing and Urban Development and the Department of Transportation, became interested in supporting the program. They provided financial support when the Bureau decided to proceed with the program.

Time requirements for creating the machinery to conduct the 1970 census will not allow the use of the improved ACG's. The ACG/DIME files will, however, be used in future censuses and surveys, and will be periodically updated.[11] The files are being developed primarily for local and State agencies concerned with small area data analysis and planning. They will also be useful for a variety of planning and analytical studies in the private as well as the public sector. Uses are discussed in chapter II.

The local coding phase of the program is scheduled to conclude in late 1970. In the majority of SMSA's, coding should take from 3 to 6 weeks. However, the larger SMSA's may take from 3 to 4 months. Computer editing will begin when the coding worksheets are returned to the Bureau. The processed geographic base files (without coordinates) are scheduled to begin becoming available 6 to 9 months after the coding worksheets are returned to the Bureau. They should be available for use soon after the census summary computer tapes begin to become available. Although the census will be conducted using the presently existing ACG's, the geographic area codes—census tracts, census blocks, street names, address ranges, etc.—will be essentially the same in both the ACG and the improved ACG/DIME files. Therefore data from the 1970 census will be usable with either file.

Coordinates for the ACG/DIME geographic base files are scheduled to be digitized from the node-numbered coding maps beginning in calendar 1971. The coordinates are scheduled to be added to the geographic base files, and tested for accuracy.[12] The files will be available on a flow basis beginning in late 1971 and continuing through 1972.

Nonmail Census Geographic Base File Program

In nonmail census areas, place-of-work addresses reported on the sample census questionnaires will be coded to census tracts for tabulation purposes. To do this, an address coding guide that relates addresses to the census tract level is required. The Bureau intended to create only census tract address coding guides for this purpose. Since such guides would be of little use to local users or to the Bureau in future censuses because census tract boundaries often change between censuses, the Bureau decided to develop DIME-type geographic base files for the nonmail SMSA's when the Department of Housing and Urban Development and the Department of Transportation agreed to support a program to create such files.[13] Nearly all of the SMSA's eligible for the program have participated (see appendix A). Census tract coding guides will be prepared for those SMSA's not participating.

The process for creating an ACG/DIME geographic base file in this program is essentially the same as the method developed in New Haven and refined in subsequent tests. The local coding phase of the program was completed in early 1970. Computer processing and editing of the files is scheduled to be completed by mid-1970. Files without coordinates will become available for local use at that time. The coding maps are scheduled to be digitized during 1971 and files with coordinates will be released on a flow basis to local agencies starting in spring 1971 and continuing into 1972.

[11] The Southern California Regional Information Study (SCRIS), a joint research study sponsored by the Bureau and the Southern California Association of Governments, of which the Census Use Study forms the Bureau's contingent, has been charged with developing a maintenance and updating system for the ACG/DIME Geographic Base File Program. Development of the system is scheduled to be completed during 1971.

[12] Coordinates in the ACG/DIME files will be expressed as geographic coordinates (expressed in degrees of latitude and longitude carried to four decimal places), state plane coordinates and "map miles" from an arbitrary point.

[13] Funds have been made available for local agency coding work under the "701" program of the Department of Housing and Urban Development and through State highway planning programs whose funds are administered by the Department of Transportation.

Uses of the DIME File

This chapter describes the use of census and local data with the DIME file for planning and analysis in both the public and private sectors.

There are many potential uses for a DIME geographic base file or other geographic files incorporating DIME features. Several were explored by the Census Use Study and others await exploration. The description that follows is intended to provide a preliminary look at these uses.

Considerable research into uses of DIME files is being undertaken at the Southern California Regional Information Study (SCRIS), a joint project of the Census Bureau and the Southern California Association of Governments.[1]

Computer Mapping

Computer mapping using a DIME file can only be accomplished if the file contains coordinates.[2]

Node coordinates alone allow only one basic type of map to be produced—a network or outline map. These coordinates are used to plot street networks, to shade within the areas described by the nodes, or to print characters or symbols within the same area. Figure 6 illustrates an outline map with shaded node-defined areas.

To allow for full use of a DIME file for computer mapping, area centroid (center point) coordinates can be calculated and added to the DIME file. For example, block centroids are necessary to plot block data to a grid, as in figure 7; to a contour interval, as in figure 8; or to proximal or incidence map. Centroids are also necessary for mapping data to other areas, such as police precincts, census tracts, and transportation zones.

The Census Use Study tested several mapping techniques and systems using the DIME file, such as MAP 01, SYMAP, pen plotters, cathode ray tubes, and the Geospace plotter.[3]

In general there are six steps in computer mapping: Selecting and specifying the data to be mapped, selecting the mapping system to be used, attaching coordinates to the data by matching the data records to a DIME-type geographic base file with coordinates,[4] manipulating and organizing the data for input into the computer mapping system, selecting the appropriate cartographic features and data categories of the map to emphasize the salient characteristics of the data, and producing the map. Figure 9 illustrates this process.

---

[1] This research will be documented and reports released on a periodic basis.

[2] The Census Bureau presently plans to provide coordinate readings for all nodes in the geographic base files completed as part of the ACG/DIME Geographic Base File Program.

[3] For a full description of computer mapping research carried out at the Census Use Study, see Census Use Study report, *Computer Mapping*.

[4] This can be normally done by using an address matching system such as ADMATCH, described on page 18.

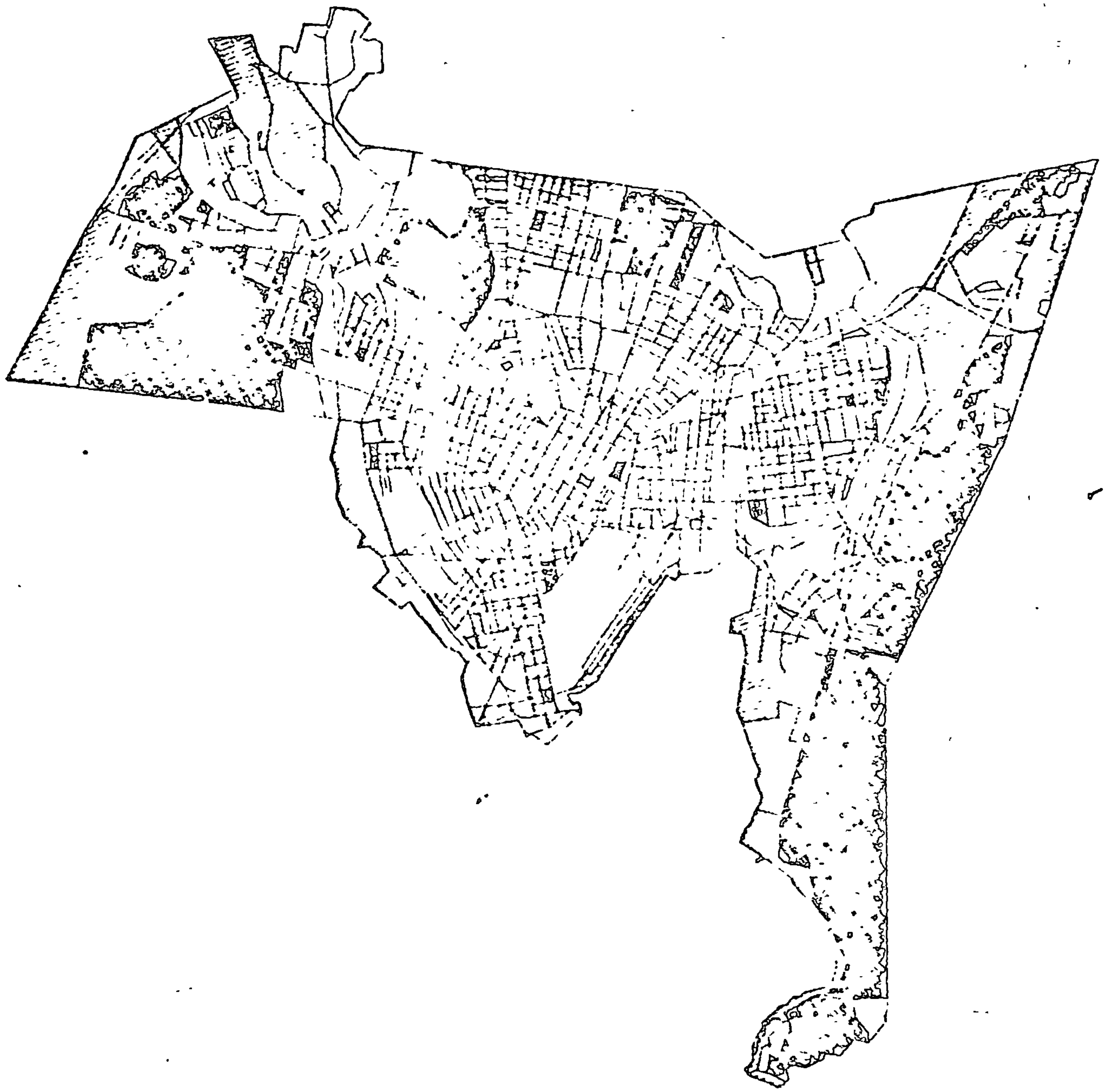


Figure 6. Outline Map—Geospace plotter shading map identifying blocks (percent of owner-occupied housing)

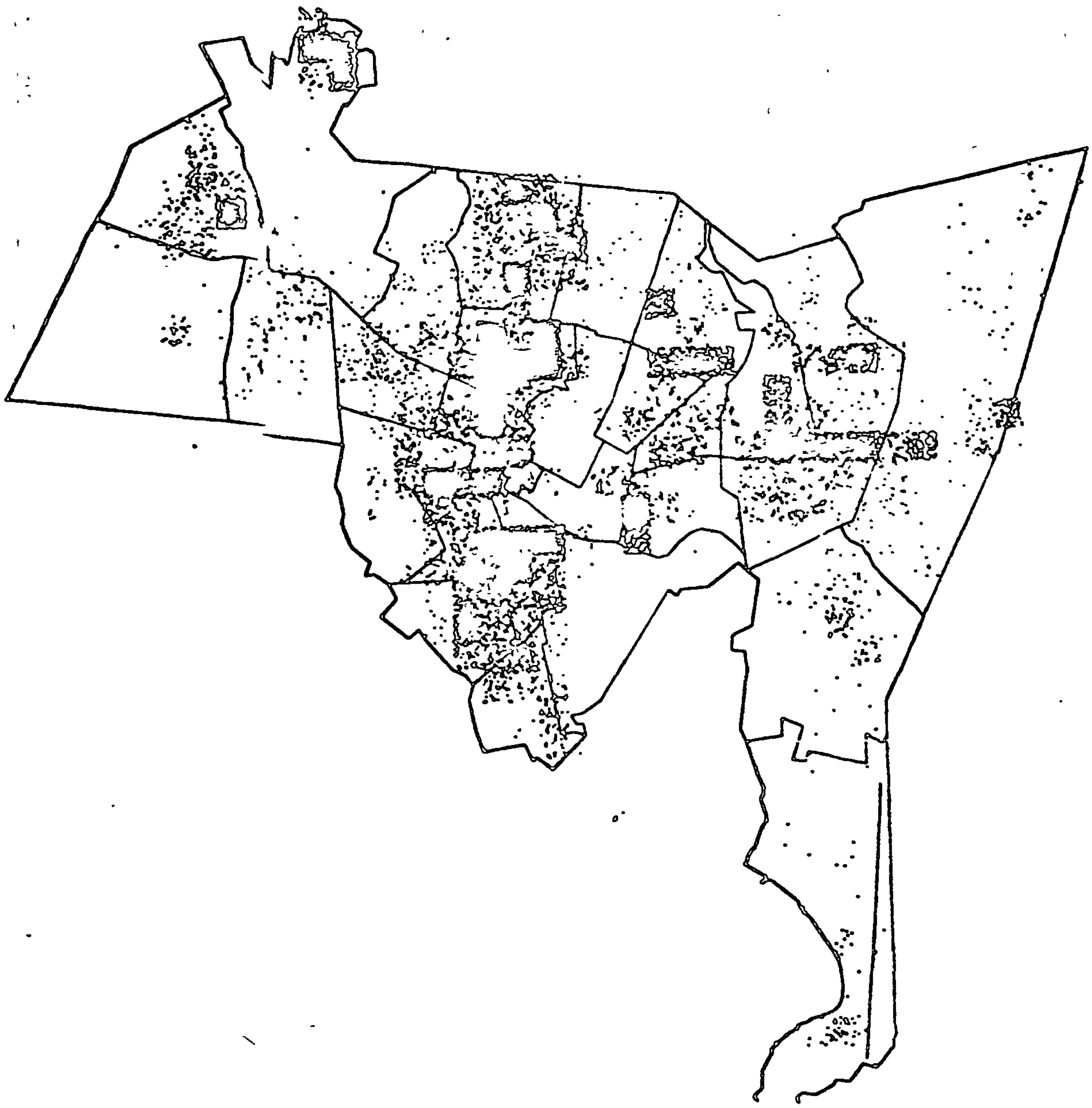


Figure 7. Grid Map - Cathode ray tube map showing grid cells (adjusted hours spent by visiting nurses)

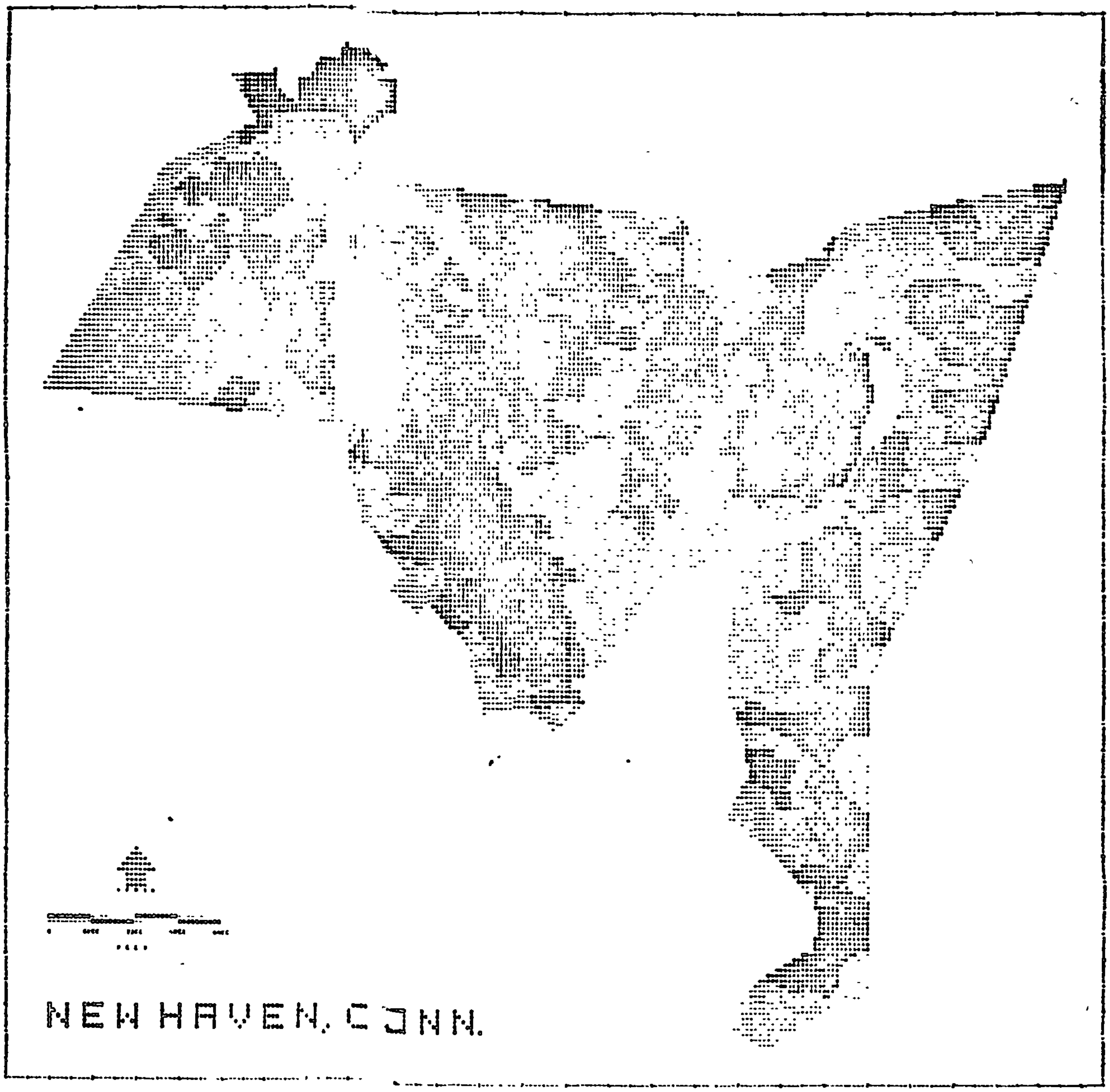


Figure 8. Contour Map—SYMAP showing contour levels (density of preschool children)

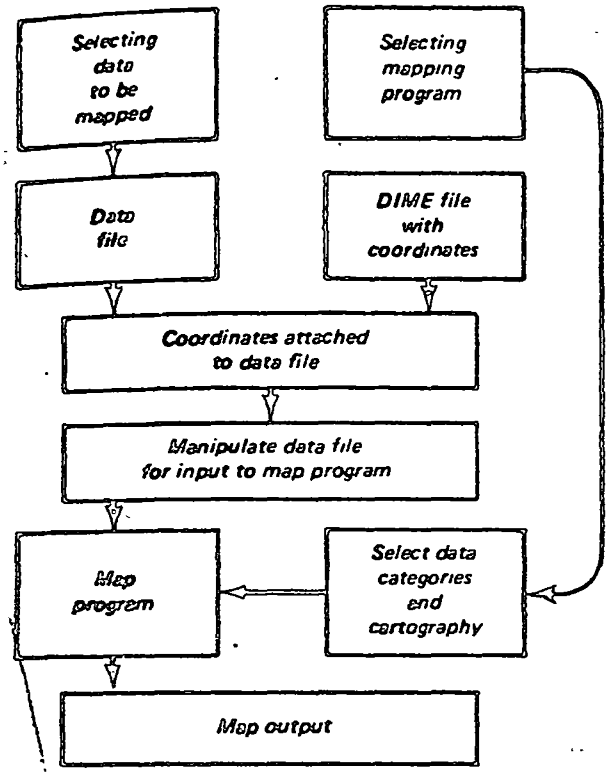


Figure 9. Computer mapping process

Adding Local Area Codes to the DIME File

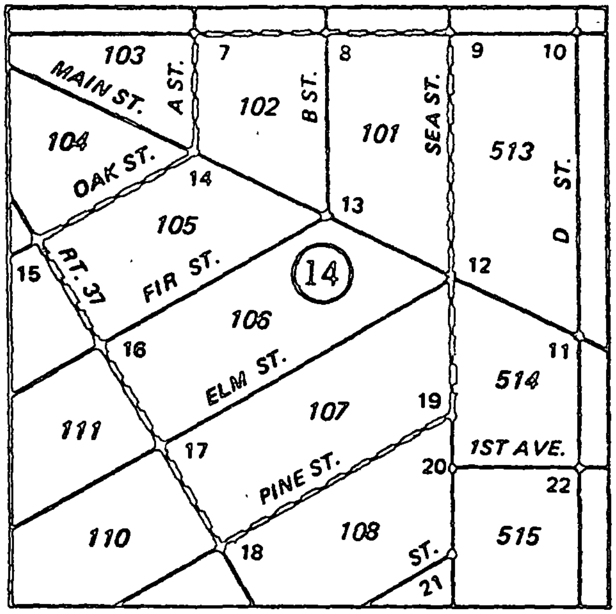
The addition of local area codes allows use of the DIME file in the analysis of local and census data as they relate to local areas. Examples of local areas for which codes can be added to geographic base files include traffic zones, planning districts, school districts, police precincts, poverty neighborhoods, and health districts. Local area codes in a DIME file facilitate the interrelating of census and local data either for the areas originally coded in the DIME file, such as census tracts, blocks, municipalities, and ZIP code areas, or for local areas such as land use files or assessor records.

Local areas are usually defined as agglomerations of city blocks, parcels, or other small areas. However, a local area may also be defined in a linear sense, i.e., a series of street or other segments, or it may be identified by a series of points such as street intersections. The 1970 census data will be available only for areal units. The census block will generally be the smallest unit.[5] However, local data may be available for any of the three types of local areas.

Local codes may be added to a DIME file by means of a dictionary or correspondence table of local codes for each census area. There are two relatively simple methods which can be employed. The primary method is to plot the local areas on a Census Bureau metropolitan map and then manually prepare a corresponding list of local area codes for each individual or series of census area codes comparable to the local areas. The census codes necessary for areal units are census tract and block, as illustrated in figure 10. For linear coding of one segment at a time (streets or other segments) the census codes necessary are a combination of the census tract number and the node numbers at each end of the segment, as illustrated in figure 11. For linear coding of several continuous segments along a street, the necessary codes are the segment name and census tract and node numbers for the beginning and ending points of the contiguous segments, as illustrated in figure 12. For node or intersection coding, the necessary codes are census tract and node number, as illustrated in figure 13. Once a dictionary or correspondence table is prepared, it is keypunched and the file processed to add the local codes to the DIME records.

The variant method of creating an area dictionary utilizes the DIME file as the source for the dictionary records. A computer program is written to process the DIME file and punch a card for each unique occurrence of the basic component of the dictionary: node, segment, block, tract, etc. For example, if traffic zones are the local areal units to be included in the area dictionary, the tape file is processed and a card punched for each unique tract/block number combination. The cards are processed automatically in a card interpreter machine which imprints the cards with the coded elements so they can be read by clerks. These cards are manually divided into traffic zones, using a metropolitan map marked with traffic zone boundaries as a guide. The traffic zone identifiers are "gang-punched" into blank fields in the cards

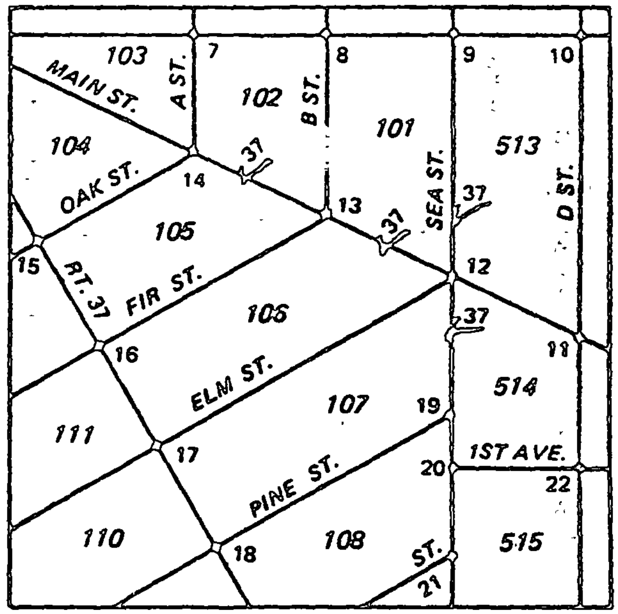
[5] Block data will only be available for Census Bureau-defined urbanized areas and other urban areas which contract with the Census Bureau for block statistics. An urbanized area contains a city (or twin cities) of 50,000 or more population plus the surrounding closely settled incorporated and unincorporated areas which meet certain criteria of population size or density. *Block side* data will be available for the urbanized area of mail census SMSA's on a contract basis only.



Local area boundary
Local area code - (14)
Entire map in census tract 21.01

Local code	Census tract	Census block
14	21.01	101
14	21.01	102
14	21.01	105
14	21.01	106
14	21 01	107

Figure 10. Areal unit coding - local area 14 outlined on map coded to census tract and block.



Local linear code - 37
Coding key -
Entire map in census tract 11

Local code	From node	To node	Census tract
37	14	13	11
37	13	12	11
37	9	12	11
37	12	19	11

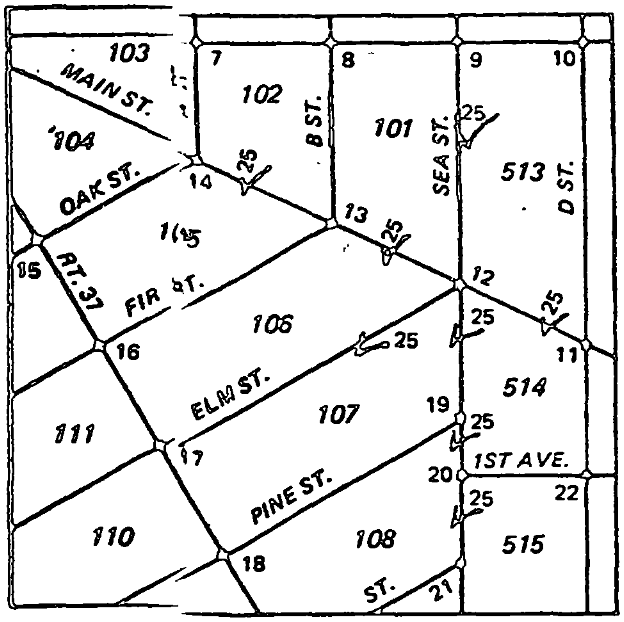
Figure 11. Linear coding by segment - local linear code /37 marked on the appropriate street segments of the map is coded to census tract, from node, and to node.

using a card reproducer machine. The card file is then usually converted to a tape file Depending on the use which will be made of the file, it may or may not be combined with the original DIME file. When it is necessary to assign the local codes to the DIME file, the dictionary and the DIME file should be sorted by tract and block, and a program written to accomplish the code assignment.

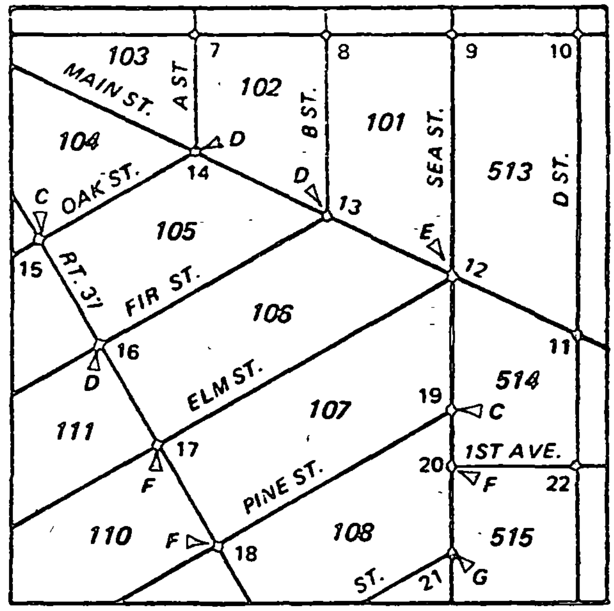
More sophisticated methods exist for adding local geographic codes to a geographic base file such as DIME. One example is the area bounding method, which describes the local geographic area by means of coordinates. For this method, a computer program could be written to determine the specific local area code for each record in the DIME file and insert the code into the record.

Using the dictionary method, the Census Use Study added local geographic area codes to the New Haven DIME file. These included codes for 1960 census enumeration districts, Community Action Program (CAP) areas and sub-CAP areas, elementary and secondary school districts, traffic zones, and telephone company central office and market areas.

Agencies requiring special local area tabulations were requested to provide a punchcard file consisting of one card per census block. The card contained the census tract and block and its corresponding local



Local linear code - 25
Coding key - ⌊—
Entire map in census tract 101



Local node code in italics - arrowed to node
Entire map in census tract 25

Local code	Street name	From node	To node	Census tract
25	*ELM ST*	17	12	*101*
25	*MAIN ST*	14	11	*101*
25	*SEA ST*	9	21	*101*

Figure 12. Linear coding by street - local linear code /25 marked on the appropriate street segments of the map is coded by street name, from node for the first segment along the street, to node for the last segment, and census tract.

Local code	Census tract	Node number
C	25	15
D	25	16
F	25	17
F	25	18
D	25	14

Figure 13 Node or intersection coding - local node codes (in italics) entered on the map and arrowed to the nodes they represent are coded by census tract and node number.

ea code. If a block was partially in one local area and partially in another, the agency assigned the block to only one local area. This situation occurred infrequently because the local areas tended to be much larger than the average city block and to be bounded by streets or other features used by the Census Bureau to define blocks. For the telephone company central office and market area codes, however, a large number of arbitrary decisions had to be made because the areas were linear in nature, i.e., they were defined as being both sides of a street, rather than as a series of blocks.

When the punchcard file containing the local area code dictionary was received by the Census Use Study, it was computer edited to ascertain that each block was assigned to only one local area. After all local area block dictionaries were submitted and checked, one master computer tape containing all local area codes for each tract/block combination was created.

As in any large-scale data processing operation, certain problems were encountered due to clerical errors and omissions. The census data tape for New Haven city contained 993 unique city block codes. The DIME file and the Census Bureau metropolitan maps showed 1,117 blocks in the city. The difference between these numbers typically indicates the number of blocks with no population. However, it

was discovered that 12 blocks appeared on the data tape with population figures; this error was due to clerical or processing errors in the construction of the original address coding guide for New Haven. These blocks did not appear in the local area block dictionaries. To add these blocks to the dictionaries it was necessary to refer to the address coding guide for street names and address ranges and then relate this information to the maps containing the local areas.

The most widespread use of local area codes was for creating tabulations of basic census data for the local areas. All special tabulations were checked before release to be sure of adherence to Census Bureau confidentiality rules. Special tabulations for any local area with fewer than five households were suppressed.[6]

Adding DIME File Codes to Local Records

Another valuable feature of a DIME-type file is its capability for address matching, the matching of records in two files on the basis of address numbers. When using a DIME file, address matching is the matching of local record files with address numbers to a DIME file with address number ranges to assign geographic codes to the local record files.

Address matching has many uses. For example, files containing records of building and demolition permits can be address matched to insert census tract numbers so that tabulations of new construction and demolitions can be prepared by census tract. These can be used with decennial census population data in estimating current population by tract. Address matching crime incidence records to census tracts or block groups can be used to indicate possible correlations between crime and census socioeconomic data. Address matching assessor records containing land-use data to planning district codes contained in the DIME file would show land-use patterns within the planning districts. Using the same assessor records matched to geographic coordinates in DIME file, computer maps of the land-use patterns can be prepared.

To facilitate address matching, the Census Use Study has developed an address matching computer package named ADMATCH.[7] The ADMATCH system can be used with DIME files, address coding guides, ACG/DIME geographic base files, census tract street indexes, or any similar computer file containing addresses or address ranges and geographic area codes. The ADMATCH system requires that data and reference files (DIME, ACG, etc.) be processed in a three-phase operation:

1. Preprocessing—The ADMATCH Preprocessor program deciphers the address and creates a 75-character match key. Special features of the Preprocessor include the ability to create city and State codes, standardize street types (e.g. ST, AVE), correct misspelled street names on data records, and add records to the reference file to compensate for street names that are misspelled or truncated.

2. Sorting—The data and reference files must be sorted on certain items in the match key, including ZIP code or State and place (city) codes, street name, and house number or address range.

3. Matching—The ADMATCH Matcher program compares items in the match keys of the data file and reference file. When a match occurs, user-specified geographic codes from the reference file are attached to the data records. Since the quality of data files varies, the Matcher program has an option to accept mismatches on the street type and direction to improve the chances for matching. For example, the Matcher can be instructed to accept a record such as 151 Elm St. when matched to 151 Elm, or 76 W. Washington St. N.W. when matched to 76 Washington St. N.W. Individual street addresses can be matched against street addresses as well as address ranges.

The concept of the matching process is illustrated in figure 14.

In New Haven, several local files were matched to the DIME file to produce tabulations and input to computer mapping programs. Local files matched included birth records, police complaints, police arrests, building code violations, fires, school attendance records, school census records, visiting nurses hours, and health records. The match rate averaged 85 to 90 percent for a perfect match (all items match perfectly) and it approached 100 percent for less than a perfect match (some items allowed to mismatch).

In one experiment, data on visiting time spent by nurses at various locations throughout New Haven

[6] For details, see Census Use Study report, *Data Tabulation Activities.*

[7] See Census Use Study computer program package, *ADMATCH: An Address Matching System.*

were provided by the Visiting Nurses Association of New Haven. This file, consisting of 29,000 records, was matched to census tract and block and coordinates in the New Haven DIME file. A 98.6 percent match rate resulted. Once matched, the local data record file with coordinates became input to a computer mapping program. Maps displaying the visit locations were prepared. Tabulations of the data were also prepared for census tracts and blocks.

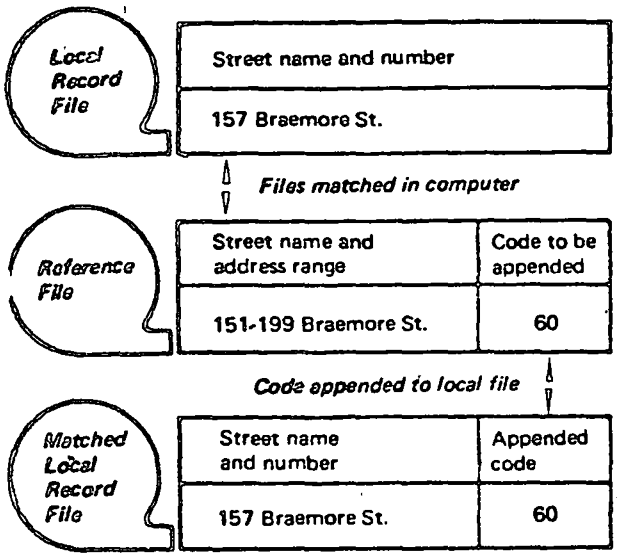


Figure 14. Matching process.

ADMATCH was also used extensively in matching local records for input to the Health Information System.[8] Among the items matched were birth records, infant and fetal death records, and hospital obstetrical records. These local record files were matched to the DIME file, and census tract and block group codes appended. Once matched, the local records were tabulated, analyzed, and input into the Health Information System.

Network and Node Analysis

Nodes or link-node combinations are used in node analyses; calculation of area, density, and network analyses; and calculation of centroids for blocks, segments, or block sides.

Node Analysis

A DIME file or other geographic base file with node coordinates can be used to geographically order data by street intersections, such as data on traffic accidents at intersections. These data can then be compared to information on traffic flow and the existence of traffic signals and signs and pedestrian crosswalks at intersections with high accident rates. Analyses can be made concerning changes needed in speed limits, the timing of traffic lights, the need for warning signs, additional pedestrian crosswalks, stop signs, etc.

Processing is accomplished in a number of ways. Local data can be coded directly to node numbers in a DIME file by clerical transcription and keypunching methods. Or a separate intersection file can be created by sorting the DIME file on node number and creating separate records for each intersection (containing geographic identifiers, node numbers, and intersecting street names—listed once for each node). This file can be sorted on intersecting pairs of streets in standard alphabetic order and local data records for intersections can be added.

A special matching program that can correct and match on misspelled street names or add supplementary records could also be created to match local intersection data to the DIME intersection file and transcribe geographic or other codes from one file to the other. Figure 15 illustrates the concept of matching a DIME intersection file and a local intersection file and appending a code from the DIME file to the local file.

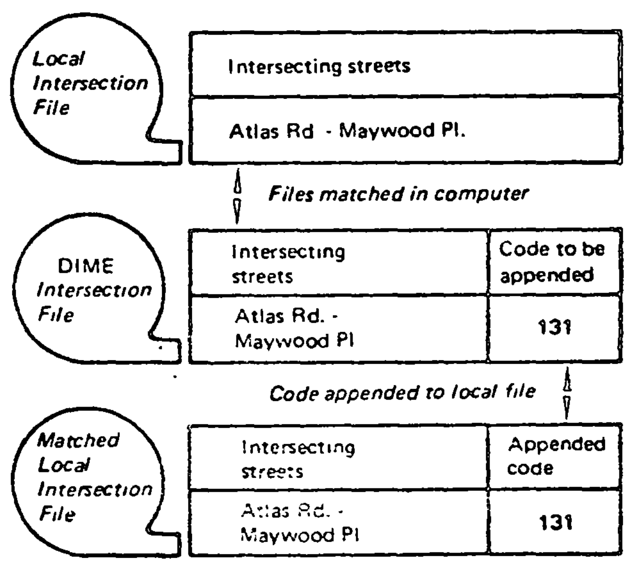


Figure 15. Matching intersection files.

[8]For a full explanation, see Census Use Study report, *Health Information System.*

Area Calculations and Arrays

Polygonal areas can be computed given coordinate information describing the polygon. Polygons can represent agglomerations of blocks, census tracts, police precincts, or arbitrary areas (centered on a set of nodes or one node) of fixed or variable width. For example, an area calculation may be required for a strip 100 feet wide on both sides of a series of street segments, as one of the data elements needed in a study of land requirements for highway construction. Figure 16 illustrates this. Or an area calculation may be required for an area, radiating out ¼ mile from a set of nodes, describing a particular set of conditions such as a parcel of land proposal for rezoning (figure 17). Or again, area calculations may be required for rings and sectors at fixed distances from a central business district.

Also, given data by polygons described in terms of coordinates, it is possible to determine by "point-in-polygon" or "polygon-in-polygon" methods which polygon a given data item belongs to. This permits arraying data by irregular polygons.

Similar but simpler methods permit arraying data by regular polygons such as grid squares, triangles, circles, rings, or sectors.

Network Analysis

A DIME file forms a substantially complete link-node system or network. Networks are used in transportation planning and resource allocation studies in which movements take place from one point to another over a network.

Transportation Studies—An important device in transportation studies is the simulation of traffic flow over a network using minimum path algorithms This device can be used for processes such as assignment of delivery trucks and postmen to least time/distance paths to make routing more efficient, assignment on a real-time basis of police cars, ambulances, fire trucks, etc., from one point in a network to another over a least time/distance path; estimation of the effects on traffic patterns when elements of the network change; analysis of effects of changes in traffic generators, etc.

Planners have used network files representing major traffic arteries for many years in conducting

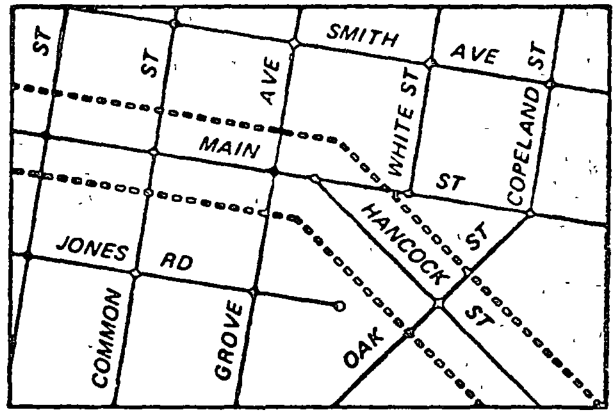


Figure 16. Area calculation for a strip centered on a series of linear segments.

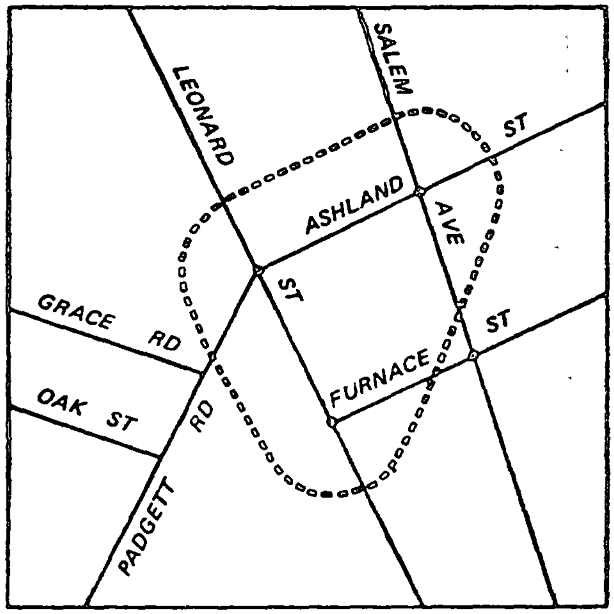


Figure 17. Area calculation for an area radiating out a certain distance from a set of nodes.

transportation studies. A DIME file contains all streets and many nonstreet features and boundaries for an area. In some cases, it may contain too many records to be handled by computer programs currently used in transportation planning. A DIME file may also lack data important to transportation system simulations, such as direction for one-way streets, number of vehicle lanes in a street, speeds, and travel time.

A DIME file can be readily adjusted for transportation applications. Data needed can be selected by developing a dictionary of required streets, comparing this to the DIME file, and creating a new network file

for only the required streets. It is also possible to add data on direction, capacity, etc., to a DIME file. Figure 18 illustrates this process.

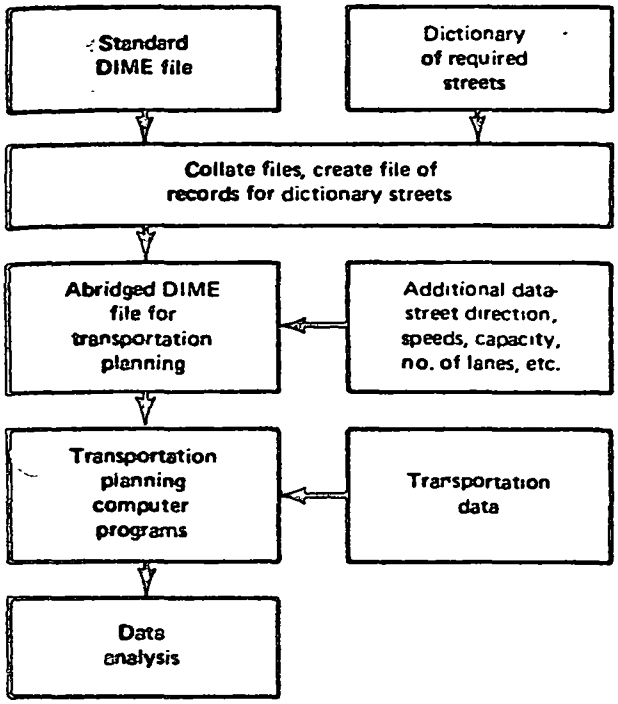


Figure 18. Adapting the DIME file for transportation planning.

Transportation network data currently being used vary considerably in format, content, coverage, and level of detail from one area to another. Standard DIME files could be used as the basis for development of standard transportation network data throughout the country. Standard computer programs, technical specifications, and analytical documents would result in considerable savings and increased efficiency. Standard DIME files for all major metropolitan areas would make it worthwhile to develop standard computer programs for solving allocation problems involving delivery truck routes, trash removal routes, bus routes, computer routing of vehicles, "dial-a-bus" systems, analysis of potential loading on existing and proposed mass transportation routes, etc.

Allocation of Resources to Facilities—With the advent of the DIME file, the Office of Civil Defense realized that substantial benefits could be gained from standardizing the Community Shelter Planning (CSP) process throughout the country. The Census Use Study is designing and developing a computerized model which will allocate people to community fallout shelters.

In the past, the CSP process relied on local information and manual map-oriented data manipulation to locate shelters, identify population distributions by small areas and by time of day, and assign people to shelters to minimize the distance traveled and use available shelter spaces most efficiently.

The 1970 census block data will minimize the need for detailed local data. Uniform DIME files for all the links in the total transportation system of a metropolitan area will make it feasible to assign people to shelters by computer. To modernize the CSP process, two objectives are being pursued in the Census Use Study project:

1. Development of a computerized model to automatically assign people to fallout shelters along a DIME network, and

2. Preparation of a data management system capable of assembling and organizing census and local data into the form required by the model.

A preliminary computer model has been developed—CRAM (Computer Resource Allocation Model). It represents the first use of the DIME file as a network. The spatial relationships between shelter facilities and people using the shelters (population data by census blocks) are determined through the use of the DIME file. The model can perform an allocation based on pedestrian and/or vehicular modes of travel. Figure 19 illustrates the CRAM model.

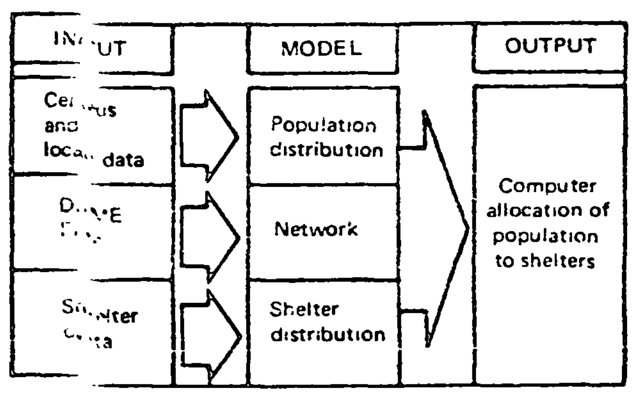


Figure 19 Network allocation of population to shelters (CRAM) model.

Although the model has been developed to assign people to shelters, it can be used for a wide range of allocation problems—for example, to allocate children to schools, determine logical service areas for community health facilities, evaluate alternative sites for new plants, evaluate existing facilities, and plan more effective utilization of existing facilities. Since the model allocates resources to facilities along a transportation network, it may be applied to any situation which embodies these three components.

Adjacency Analysis

The ability to group contiguous blocks (or other small geographic areas) by certain homogeneous characteristics of the blocks has interested urban planners and analysts for some time. A DIME file and appropriate software make this process faster and more efficient by relating all nodes, segments, and blocks to each other. Just as network analysis uses the link-node relational characteristics contained in a DIME file, adjacency analysis uses block-node relational characteristics.

Aggregation of Areas

Contiguous areas such as blocks, tracts, and police beats can be combined to create an area of predetermined size in terms of land area, number of households, number of street segments, etc. If the areas are to be defined in terms such as square feet or acres, a DIME file with coordinates would be the only required input to the computer process for aggregation. If the areas are defined in terms of other data, a data file containing the necessary data elements must also be used.

Computer programs for aggregation can be designed to provide compact aggregations so that fingers, holes, or serpentine aggregations are avoided. This can be done by specifying in the computer program that each successive area added to the aggregation should be the area closest to the initial area. Divergencies from this rule can be allowed. For example, a 40-percent divergency would allow one dimension of the aggregation to be 40 percent greater than the other. Other constraints such as a feature to stay within block groupings or census tracts as much as possible or to stay within municipal boundaries, can be added.

A prime use for aggregation techniques would be the delineation of legislative districts, based on the one man-one vote ruling of the Supreme Court. Although the results of an aggregation program may not be completely acceptable to lawmakers responsible for drawing such districts, they would provide a solid base or starting point from which equal population districts could be delineated. Other uses would be the devising of police patrol beats based on the accumulated length of the segments within a compact aggregated area, or the determination of school or health districts based on existing and/or proposed schools or clinics.

Clustering of Similar Areas

Contiguous areas, all within the same range of data values, can be accumulated into larger areas of undetermined size using a DIME file, data files, and appropriate computer software. Data values are the basis for the clustering.

The computer software may be designed so data value ranges can be stipulated prior to operating the computer program or so that they can be determined mathematically during computer processing in terms of quartiles, quintiles, etc. Also, the software may provide for stipulating size criteria for clusters (e.g., maximum or minimum number of areas) as well as criteria on data value discontinuities.

The concept of clustering allows for the determination of homogeneous areas for community action programs, model cities development, poverty areas, high health risk areas, ethnic enclaves, areas where the elderly predominate, or other types of homogeneous areas.

Other Uses

Other potential uses for a DIME file include—

Monitoring of Programs by Geographic Areas

A DIME file framework can be used for reporting the detailed analysis data on the progress of a program for a specific geographic area, ranging from the block level upwards. The Model Cities Administration of the Department of Housing and Urban Development is currently exploring the possibility of using DIME files for this purpose.

Development of Parcel Files

The DIME techniques for creating a geographic base file can be used to create a parcel file. Coding and processing would probably be expensive. The limited amount of research carried out by the Census Use Study to date indicates that a DIME-type parcel file would cost at least 20 times more than a typical DIME file at the block level. The methodology is identical: blocks become parcels; nodes are placed at parcel corners; and each lot or parcel receives a unique number, like a block number. Address ranges are coded for each parcel, if available. Since a geographic base file must contain records for all the geography within an area, sidewalks, streets, intersections, and other elements such as traffic islands, cul-de-sacs bodies of water, and bridges, should be considered as parcels and so coded.

For some applications, it would not be necessary to create a parcel file in the detail outlined above. For each segment record in a regular DIME file, subsidiary records for each parcel along the segment could be created. Using this method, parcel data could be assigned through the segment record to the parcel subsidiary record. Parcel data could then be retrieved at the parcel, segment, block, or any higher geographic level.

Development of Larger-Grained Files

Just as it is possible to create DIME files for parcels, it is also possible to create files for areas larger than blocks such as block groups (or enumeration districts),[9] census tracts, planning districts, police beats, municipalities, and counties. For example, a researcher concerned with rural problems might wish to create a DIME file for a large depressed rural area. Such a file could be based on enumeration districts, these being the smallest tabulation units available from a decennial census for rural areas. The researcher would then be able to manipulate census data, and local data coded to the ED, using the techniques outlined previously in this chapter.

Voting ward or precinct DIME files could be used by political scientists in analyzing election returns or by politicians designing new legislative district boundaries. County DIME files could be used by economic planners at the State or Federal level to study and analyze economic trends and projections.

Information Systems

The DIME file can be readily used as the geographic base for information systems. For example, in the Health Information System[10] developed in New Haven by the Census Use Study for the State of Connecticut, block groups derived from the DIME file were the basic geographic units used in the system. Local data files were matched to block groups by means of the ADMATCH system. Local and census data were combined at the block group level and various statistical analyses were performed to generate socioeconomic and health indicators.

Information systems for municipalities and other political or planning jurisdictions should be based on a geographic framework such as DIME. By using a DIME file, the data elements in an information system could be manipulated, analyzed, and displayed using network and node analysis, adjacency analysis, computer mapping techniques, etc.

Area Sampling

Random or stratified geographic sampling on a segment side, segment, block, or higher geographic level could be accomplished using a standard DIME file. Sampling at the parcel or large-grained level could also be done if such files were created.

Geographic Grouping of Census Summary Tape Data

The use and manipulation of data provided on a block and block grouping basis by the Census Bureau after the 1970 census will be greatly facilitated by the use of DIME files.

---

[9] The areal unit called the block group corresponds to the enumeration district in previous censuses There are approximately 10 blocks per block group The number of block groups per census tract varies from one to eight, the average is four The block group is always a subdivision of a census tract.

[10] For further information see Census Use Study report, *Health Information System*. While this report details a system primarily concerned with health, its methodology is widely applicable to other functional areas.

Time Series Studies

With a geographic base file such as DIME, data can be coded to the block, segment, or other level to reflect historical developments for time series analyses.

Spatial Comparisons

Researchers may wish to compare and study differences in the characteristics of certain areas from one city to another. Using some of the techniques described in this chapter, such study areas could be delimited and various data items retrieved and studied for the areas. As an example, socioeconomic characteristics could be obtained for a series of hard-core unemployment areas in various cities throughout the country from 1970 census summary tapes and analyzed using the appropriate DIME files and analytical techniques.

Market Analysis and Other Business Uses

Many of the techniques described above have applications in market analysis. For example, market analysis areas can be inserted in a DIME file, resulting in the ability to retrieve census or company data by market areas. Network and node analysis has value in the determination of delivery routes. Use of a DIME file with lists of prospective customers gives a firm the ability to map out a rational scheme for following up on such prospects. Placement of new facilities or relocation of existing facilities because of changes in the nature of business or customer characteristics can be planned by using the appropriate software, data, and DIME files.

Technical Description of the DIME System

This chapter describes the conceptual origins and technical aspects of the DIME system.

Conceptual Origins

The most significant technical contribution of the DIME geocoding system is the topological edit. It provides for accurate computer editing and correcting of the structural elements of the coded geographic file; i.e., streets and other linear elements, points where the linear elements intersect, and area identifiers. The edit can theoretically be done manually, although practically it is only done by computer, except in a demonstration or test situation. Depending on the number of times the file is cycled through the edit process, it is possible to correct the clerically coded file so that it becomes a perfect replica of the elements on the map coded. However, in practice, cost limitations usually do not permit a sufficient number of edit cycles to correct the file completely. A residue of uncorrected errors may range from less than 1 percent to 4 or 5 percent. At least two or three edit cycles should be completed to assure sufficient accuracy in the file.

As stated earlier, the edit system is based on concepts derived from graph theory. Since any street map is basically a type of linear graph, concepts derived from graph theory can be used as a means for creating, correcting, and maintaining a high quality geographic base file. Single-line maps can be categorized as simple geometric figures consisting of three basic uniquely identifiable elements. Points or vertices, lines connecting these points, and areas enclosed within a series of lines forming an enclosed space. Figure 20 illustrates these elements of a typical single-line map.

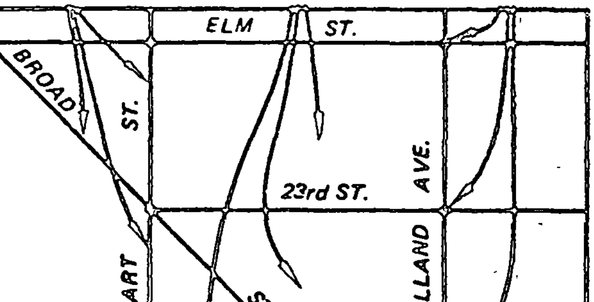


Figure 20. Linear graph elements of a map.

The Census Bureau's series of metropolitan maps is ideally suited for this purpose because all three elements can be identified uniquely. Two of these elements are identified on the maps when produced: lines are identified as Grant St., Muddy River, etc., and areas are identified by block numbers, census tract numbers, etc. The third element can be identified by uniquely numbering all vertices or points where lines begin or end. Curved streets or other linear features on the map, such as rivers, railroad tracks, or boundaries, can be represented as a series of straight line segments by defining a sufficient number of vertices along the curved line.

Composition of DIME File

A DIME file is composed of segment records. A segment is defined as a length of a street or other feature between two distinct vertices or nodes. Other

features are imaginary lines defining political or other boundaries; topological features such as rivers, shorelines, and canals; other map features such as railroad tracks, airport runways, and piers; and any other feature defining a block boundary. Nodes are points where features begin, end, intersect, or curve sharply. Unnamed features, when coded, are described uniquely.

Each segment is coded separately with the three basic codes needed to edit the file—segment name describing the linear element, "from" and "to" node numbers describing the nodes at either end of the segment, and left and right block numbers describing the areal identifiers on either side of the segment. If the segment is a street, address ranges for both sides of the segment are usually coded. Address ranges are coded because DIME files will frequently be used to assign geographic codes to local data files containing street addresses as the only locational code. There is also a separate coding field which can be used to describe nonstreet features and other unique segments such as proposed streets, pedestrian walkways, lake shores, rivers, and boundaries.

Additional code fields which are used for a series of segments are ZIP code, area code (a municipality code), election ward numbers, and census tract. It is usually not necessary to code ZIP codes or election ward numbers if they will not be used. If more than one municipality is coded, area codes should identify each segment to its proper municipality. If the area has census tracts, they should be coded. If census tracts do not exist for an area and the size of the area is small enough so that each block within the area can be uniquely identified with a three-digit code, then census tract codes need not be used. "Pseudo census tract" areas can be created and numbered for nontract areas where an areal identifier larger than a block but smaller than a municipality is needed.[1]

In summary, the elements in the DIME file are:

Mandatory elements
- Segment name and/or description
- Node numbers (two for each segment)
- Block codes (two for each segment)

Elements mandatory under certain conditions
- Address ranges
- Area codes (municipality code)
- Census tract codes

Optional elements
- Code for nonstreet features
- ZIP code
- Ward or other election district codes

Figures 21 and 22 illustrate a listing of a typical DIME file.

DIME Topological Edits

There are two types of DIME edits. One "chains" the string of segments that bound a block and is therefore known as the block chaining edit. The other "chains" the blocks surrounding a node and is known as the node chaining edit. The block chaining edit is performed first and is the most important edit as it detects the great majority of structural defects in the coded file. Because of this, the node chaining edit is frequently omitted.

Block Chaining Edit

The block chaining edit operates on the three mandatory coded elements for each segment record: Segment name or description, node numbers, and block codes. It also serves as a check on the accuracy of the census tract code as the records are sorted by tract prior to the computer processing. As a byproduct of the topological edits, erroneous tract codes are detected. An elementary illustration of the method used in the block chaining edit is shown below. Block 105, the block to be edited, is shown in figure 23. The basic elements needed for the edit of block 105 are shown in figure 24.

The basic steps followed by the computer are:

1. All segments coded to block 105 (either block-left or block-right) for the census tract being edited are selected from the file.

2. As each segment record for block 105 is selected, the computer checks the position of the block number of the block being edited.

[1] A full explanation of the coding methods, supervision, problems, and alternatives is provided in the clerical instructions manual of the Census Use Study computer program package, *DIME: A Geographic Base File System.*

	Area Code		Ward		Census tract		Header No.
	Left	Right	Left	Right	Left	Right	
11001	35	35	7	7	14	15	30151

Figure 21. Header items (codes for a series of segments)

Segment name or description	Code	From node	To node	Block No.		Left Addresses		Right Addresses		Header No
				Left	Right	Low	High	Low	High	
ANDERSON RD.		75	76	111	120	900	998	901	999	30151
ANDERSON RD.		76	77	112	119	1000	1098	1001	1099	30151
ANDERSON RD.		77	78	113	118	1100	1198	1101	1199	30151
ANDERSON RD.		78	79	114	117	1200	1248	1201	1249	30151
ARGONNE ST.		34	36	271	279	400	488	401	449	30151
ARGONNE ST.		36	35	270	283	450	498	451	499	30151
ARGONNE ST.		35	39	270	282	500	598	501	599	30151
BADGER RIVER	2	107	108	137	137					30151
BADGER RIVER	2	108	112	137	137					30151
BADGER RIVER	2	112	113	138	137					30151

Figure 22. Segment items (codes for each segment)

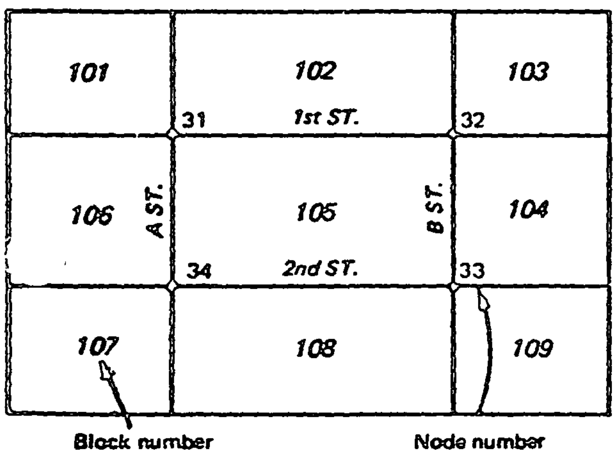


Figure 23.

Segment name	From node	To node	Block left	Block right
1st St.	31	32	*102*	*105*
2nd St.	34	33	*105*	*108*
A St.	34	31	*106*	*105*
B St.	33	32	*105*	*104*

Figure 24.

a. If the block number is in the block-left position, it is transferred to the block-right position and the other block number is transferred to the block-left position. The node numbers are also exchanged; the "from" node replaces the "to" node and vice versa.

b. If the block number is in the block-right position, no changes are made.

3. When all the block numbers for the block being edited are in the block-right position, the computer attempts to link or chain the nodes from one record to another, rearranging the sequence of segments as necessary. Notice that it was necessary to move the last segment record in figure 25 to a position between the first and second records. Figure 26 illustrates the final arrangement of the segments and the dotted lines indicate how the computer chains the segment records.

If the nodes chain and the first "from" node is the same as the last "to" node the block is considered topologically correct. Note the parallel of the computer operation in the hypothetical chaining of block 105 in figure 27.

Segment name	From node	To node	Block left	Block right
1st St.	31	32	102	105
2nd St.	33	34	108	105
A St.	34	31	106	105
B St.	32	33	104	105

Figuro 25.

Segment name	From node	To node	Block left	Block right
1st St.	31	32	102	105
B St.	32	33	104	105
2nd St.	33	34	108	105
A St.	34	31	106	105

Figure 26.

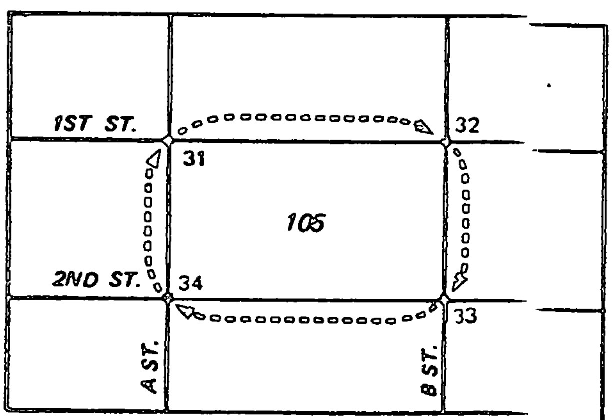


Figure 27.

If any segments remain, or if the block cannot be chained, the block records are rejected as a potential error. For instance, if any of the records in the above example were missing (i.e., not coded) the block would not chain and would therefore be rejected.

If the node numbers or block numbers were reversed, the block would not chain properly and would be rejected. As an example, if the left and right block numbers for 1st Street in figure 24, were coded 105 to the left rather than to the right and 102 to the right rather than the left, the block would contain a "reversal" and would be rejected. Figure 28 illustrates this point.

Segment name	From node	To node	Block left	Block right
B St.	32	33	104	105
2nd St.	33	34	108	105
A St.	34	31	106	105
1st St.	32	31	102	105

Figure 28.

Segment records for the blocks rejected are printed out on a reject listing for review. When reviewed, and corrected or recoded, the segment records are keypunched, inserted in the computer file, and reedited.

Node Chaining Edit

The node chaining edit operates with the same elements, and chains blocks around a node rather than nodes around a block. An elementary example of the method used is illustrated below. Node 29, the node to be edited, is shown in the center of figure 29. The basic coded elements needed for the edit of node 29 are shown in figure 30.

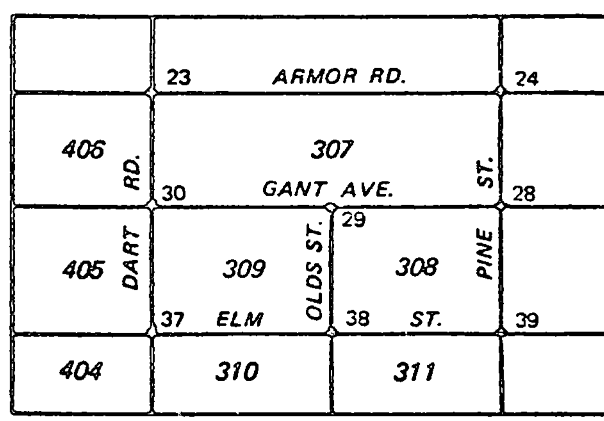


Figure 29.

Segment name	From node	To node	Block left	Block right
Gant Ave	30	29	307	309
Gant Ave.	29	28	307	308
Olds St.	38	29	309	308

Figure 30.

Essentially the same process is used in the node chaining edit as is used in the block chaining edit. The basic steps are:

1. All segments coded to node 29 (either in the "from" or "to" node position) for census tract being edited are selected from the coded file.

2. As each record is selected the node number for the node being edited is automatically transferred (if necessary) to the "to" node position as illustrated in figure 31. If the node number is transferred the block numbers are also exchanged.

Segment name	From node	To node	Block left	Block right
t Ave.	30	29	*307*	*309*
Gant Ave.	28	29	*306*	*307*
Olds St.	38	29	*309*	*308*

Figure 31.

3. When the segment records are structured so that the node being edited is in the "to" node position, an attempt is made to chain the blocks around the node, rearranging the sequence of segments as necessary. Notice that it was necessary to move the last segment record in figure 31 to a position between the first and second records. Figure 32 illustrates the final arrangement and the dotted lines indicate the chain around the node.

Segment name	From node	To node	Block left	Block right
Gant Ave.	30	29	*307*	*309*
Olds St.	38	29	*309*	*308*
Gant Ave.	28	29	*308*	*307*

Figure 32.

If the blocks chain and the first left block is the same as the last right block, the node is considered to be topologically correct.

If any segments remain, or if the the node cannot be chained, the node records are rejected as a potential error. For example, if any of the records in the above example were missing, i.e., not coded, the node would not chain and would, therefore, be rejected. If the node numbers or block numbers were erroneously reversed, the node would also be rejected.

All segment records for the nodes rejected are printed out on a standard reject listing for review. When reviewed, and corrected or recoded, the segment records are keypunched, inserted in the computer file and reedited.

Thus, two basic structural elements in the file can be made 100-percent accurate by processing the file until all rejects are corrected. The other structural element—segment name (and address numbers)—is edited during the address edit, discussed below

Interior Segment Edit

There is a limitation to the topological edits. Segments interior to a block such as dead-end streets cannot be edited. For example, figure 33 illustrates a dead-end street--Pine Place, which does not form part of the boundary of block 305. When edited, the segment for Pine Place would be rejected as a potential error because it does not form part of the chain around block 305. When investigated, it would be found that Pine Place is a legitimate segment record and as a result would be retained in the file as correct. However, if Pine Place were missed in coding, it would not have been detected by the topological edit because the block would have been chained and found acceptable.

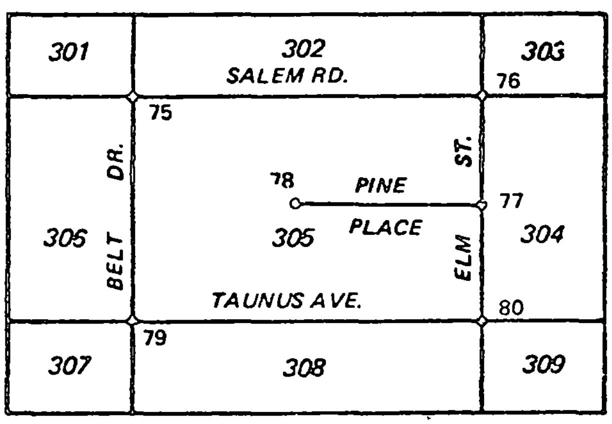


Figure 33.

Segment name	From node	To node	Block left	Block right
Salem Rd.	75	76	*302*	*305*
Elm St.	76	77	*304*	*305*
Elm St.	77	80	*304*	*305*
Taunus Ave.	80	79	*303*	*305*
Belt Dr.	79	75	*306*	*305*
Pine Place	77	78	*302*	*305*

Figure 34.

There is a relatively simple method of assuring the inclusion in the file of all interior segments. The method entails matching node numbers in the file against a listing of nodes developed clerically when the coding maps are node numbered. The development of this listing on a node control form is discussed further in the clerical manual of the DIME computer program package. The node control listing includes inclusive ranges of all node numbers used in the area coded. Therefore, each of the node numbers listed on the node control form should match at least once to the node numbers in the DIME file.

Matching can be done either manually or by computer. To accomplish a manual match, each node in the DIME file should be computer listed once—in map number, census tract number, node number order. The resulting listing should then be matched to the node control list, which is prepared in the same order. All unmatched node numbers should be investigated for possible errors. The computer matching process requires the development of a computer program to compare the two lists in the order indicated above and print out any unmatched cases. The node control form should be keypunched for insertion into the program. Unmatched cases should be investigated for possible errors.

DIME Address Edit

The DIME address edit was developed primarily to check the completeness and consistency of address ranges and street names in a DIME file. The edit checks one entire street at a time by stringing together all segments for the street by linking node numbers. Once linked, the address ranges on each side of the street must be in ascending order without overlaps. One side of the street must have even address numbers and the other, odd address numbers. If a street is broken into pieces or passes from one jurisdiction to another, appropriate exceptions are allowed. Nonstreet records are not edited.

Coordinates

The node numbered maps used in coding can be digitized; i.e., coordinates can be determined for each node, at any time after coding has been completed. However, the insertion of digitized coordinates into a DIME file is not usually performed until the file is computer edited as it is desirable to have a "clean" file for coordinate insertion.

The primary reason for adding coordinates to a DIME file is to prepare the file for computer mapping, distance calculations, and other applications involving spatial relationships.

There are a number of different coordinate systems and each has its own strengths and weaknesses. The Census Use Study used only the state plane coordinate system. As mentioned earlier in this report, the Census Bureau plans to make the ACG/DIME geographic base files available with geographic coordinates (latitude-longitude), in state plane coordinates and also with "map miles" north and east from an arbitrary point.

Several methods can be used to digitize a map and insert coordinates in a DIME file. The method used by the Census Bureau entails the use of a semiautomatic coordinate locator built by Bureau technicians. Generally there are five steps in the process. The first step is to set up the map on the coordinate locator (digitizer) table and prepare the map for digitizing. The second step is to read or digitize the map coordinates for each node. The third step is to convert the map coordinates, which are read in hundredths of an inch by the digitizer, to actual ground coordinates—state plane, geographic (latitude-longitude) coordinates, etc. Some digitizing equipment combines steps two and three. The fourth step is to attach the coordinate records to the DIME file records to which they apply. The fifth step is to plot the DIME file with coordinates at the same scale as the original map, compare the two maps, and correct any obvious errors. These five steps are explained in more detail below.

Map Setup

The map sheet to be digitized is placed on the digitizer table and positioned so that the state plane coordinate registration marks are parallel to the table edges. The registration marks, located on the edges of the maps, are aligned parallel to the table edge because in some cases map edges may not be parallel to the coordinate system indicated along those edges.

After each map is set up, an origin point is determined. This point should be to the left (west) of and lower than (south of) the most southwesternly node to be digitized. This point should be located in the border region of the map and clearly marked and identified on the map. The origin point can be any arbitrary point meeting the above requirements. The digitizer should then be adjusted to measure from the origin point.

Locations of at least two registration marks on each border are read and the coordinates and digitizer readings for these registration marks recorded for later use by the conversion programs.

Digitizing

Digitizing is the process of reading map coordinates for each node from the node-numbered coding map. In using the Census Bureau's semi-automatic coordinate locator, the operator of the machine places a cursor over each numbered node and presses a button which records or reads the location of the node. The node is read in hundredths of an inch up and to the right of an origin point located in the lower left corner of the digitizing machine. Each time a node is read, the operator must also keypunch the map sheet number, census tract number, and node number on a punchcard, thus creating a complete record for each digitized node.

Some digitizing equipment have the ability to automatically multiply each coordinate reading (in hundredths of an inch) by the appropriate scale factor and add the resulting reading to the geographic (latitude-longitude) or state plane coordinate reading of the origin point.

For small areas it is often practical to digitize the map manually on a drafting table. This is done by selecting an origin point on the map to the left and below the lowest left node to be digitized, and then measuring carefully the location of each node relative to the origin point. As each node is read, a record of its X and Y location, node number, and any other necessary map code should be prepared for keypunching. A variant on this method would be to use a light-table, and place graph paper under the map to be digitized. The grid lines on the graph paper should be systematically numbered according to some predefined scaling method. The map is placed over the graph paper and the nodes read by relating the node to a pair of grid lines, scanning the grid lines and recording the appropriate X-Y readings. Another method would be to transcribe the map onto graph paper and proceed as outlined above.

However the digitizing is done, certain quality control measures should be followed. It is less costly to catch errors at the initial digitizing stage than after coordinates are inserted in the DIME file. A suggested procedure is as follows:

1. After all nodes on a map have been read in the normal manner by one operator, a second operator rereads every 25th node starting at a randomly selected node in the first 25. The rereading must be independent, without reference or comparison to the original readings.

2. A third person compares the readings of the original operator to the readings of the second operator and notes all differences. Differences of less than .05 inches can be ignored. Each difference is checked to ascertain which operator made the error. If the original operator had more than a certain prespecified percentage of node readings in error (5 to 10 percent, depending on accuracy desired), the original digitizing is rejected. The map is then completely redigitized. Then, a new quality control sample is used. If less than the prespecified percentage of node readings are in error, the work is accepted and the node readings in error are corrected.

After all incorrect readings are corrected, the digitizing process is complete. During digitizing and quality control, the map should not be demounted from the digitizer because reading errors may result if the map is remounted and digitized. Quality control procedures should always be used, regardless of the reputation of the operator or the organization.

For digitizing at the Census Bureau, the node numbers to be digitized on each map are first selected from the ACG/DIME file and punched out on cards. This is done to eliminate manual entry of node numbers and is made possible by the fact that the Census Bureau Coordinate Locator can display prepunched information from the cards on the operator's console.

Conversion

This is the process of transforming the original digitizer readings (in hundredths of an inch) to a coordinate system. Conversion is usually done by multiplying each node reading in hundredths of an inch by the appropriate map scale and adding the result to the coordinate reading of the origin point. A simple computer program can be written to perform this conversion.

Scale factors and the state plane coordinate reading of the origin point are determined as follows:

1. Divide the difference between all pairs of state plane coordinate readings along each border of the map by the same difference in inches. The scales should be nearly identical for all borders. Significant differences should be checked, small differences ignored. Paper maps tend to have greater differences than more stable material such as Mylar.

2. Record in inches the readings of the registration marks around the map border. Multiply each registration mark reading by the scale factor and subtract the product from the state plane coordinate reading for the registration mark. The average of the numbers resulting from this subtraction yields the state plane coordinate reading for the origin point.

Similar processing is applied to convert digitizer readings to other coordinate systems.

Coordinate Insertion

Coordinate insertion requires that another computer program be written. The program reads and stores the coordinate file in a directly accessible medium, either in its entirety or in parts such as in map sheet or tract number parts. The DIME file (either in its entirety or in map sheet or tract number parts) is then processed, one segment record at a time. The coordinate file is searched for the coordinate reading for each end (node) of the DIME segment record. Node coordinates are then appended to each DIME record.

This program can also be written to perform a rudimentary edit by printing out a listing of all segment records for which coordinates are not available and all coordinate records which were not appended to a DIME record at least once. The program can also check segment records which seem to be excessively long for the nature of the area being digitized.

A subsequent program may be used to perform certain mathematical checks for errors of closure and the existence of intersections between segments which do not share a node. Such intersections would be evidence of digitizing errors since all legitimate intersections should be nodes.

Plotting Test Map

Once the final DIME file with coordinates is prepared, it is plotted with line plotter equipment at the same scale as the original map. The resulting map should be carefully compared to the original map and any serious divergencies noted. If there are sufficient numbers of serious reading errors to warrant correction, an interpreted punchcard file of the segment records containing coordinates is prepared. Each reading error is corrected by measuring the X and Y distances between the incorrect reading and its correct location, converting the distances to map scale, and entering the correct readings on the appropriate punchcards. These punchcards are then repunched and merged with the original file to create a final corrected DIME file with coordinates.

Development of a DIME File

This chapter describes in general terms the process of creating a DIME file from clerical coding through computer processing. It also describes coordinate insertion, and updating and maintenance research.

DIME or ACG/DIME geographic base files will be available beginning in 1970 for the urbanized areas in most of the 233 standard metropolitan statistical areas of the country. Regional planning agencies or councils of government will often be the agencies responsible for development of the file. In other areas, county planning agencies or transportation studies may have this responsibility.

For the remaining urban areas of the country, including those SMSA's which did not opt for the Census Bureau's ACG/DIME program and smaller cities not eligible for the program, local agencies, with adequate technical staff, can create a DIME file. Otherwise, a computer consultant or software firm can be hired to assist in creating a DIME file using the clerical instructions and computer procedures and programs in the Census Use Study computer program package, *DIME: A Geographic Base File System.*

Other research and planning organizations who wish to create a DIME file for experimentation may also obtain copies of the computer program package.

Figures 35, 36, 37, and 38 outline the preparation of maps for coding, clerical coding, computer processing, and insertion of coordinates.

Prerequisites

Creation of a DIME file, requires consideration of costs, technical and coding staff, computer requirements, coding maps, and address reference materials.

Costs

Preparation of a DIME file without coordinates by local agency personnel alone should cost approximately $2.00 to $2.50 per block for clerical coding and computer processing, assuming accurate and suitable scale coding maps, address reference materials, and the proper computer equipment are available. The digitizing and insertion of coordinates into the file should add $.50 to $1.00 per block, making a total of $2.50 to $3.50 per block.

Technical and Coding Staff

For most coding operations five coders and one supervisor will probably be sufficient. With five coders a 1,000 block area would take approximately 1 month to complete—1 to 2 days training, 1 to 2 days node numbering, 5 to 7 days coding, 1 to 2 days clerical edit and quality control, and 3 to 4 days each for clerical correction of the topological and address edit rejects. If the coding area is considerably larger than 1,000 blocks, it might be desirable to increase the coding staff to 10 or 15, keeping constant the 1 to 5 supervisor-coder ratio.

One professional should be responsible for the entire operation and the person should be familiar with the geography of the area and with computer processing capabilities and techniques. Specific suggestions concerning the coding staff are contained in the clerical instructions of the DIME computer package.

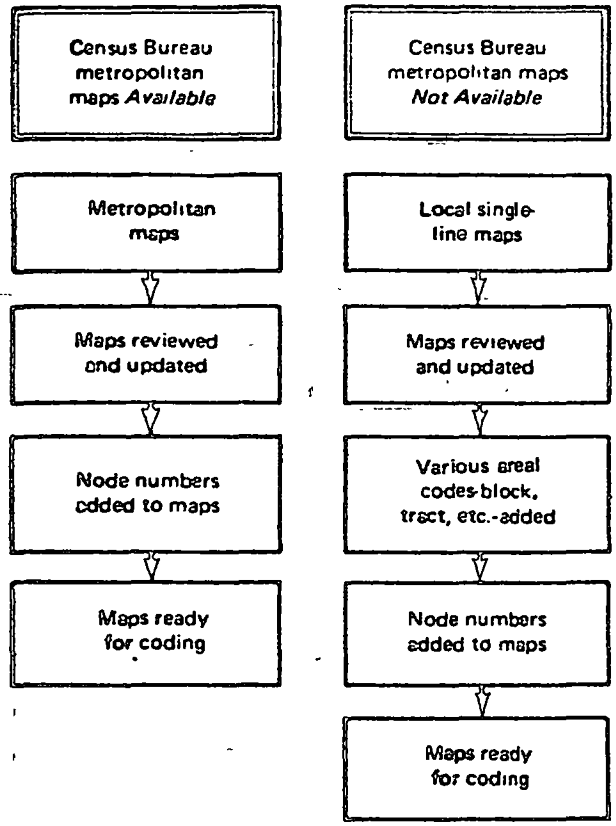


Figure 35. Preparation of maps for coding.

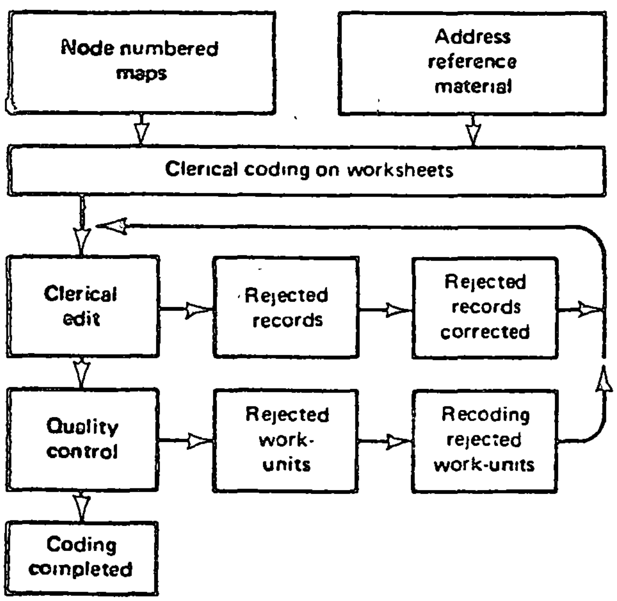


Figure 36. Clerical coding of DIME file.

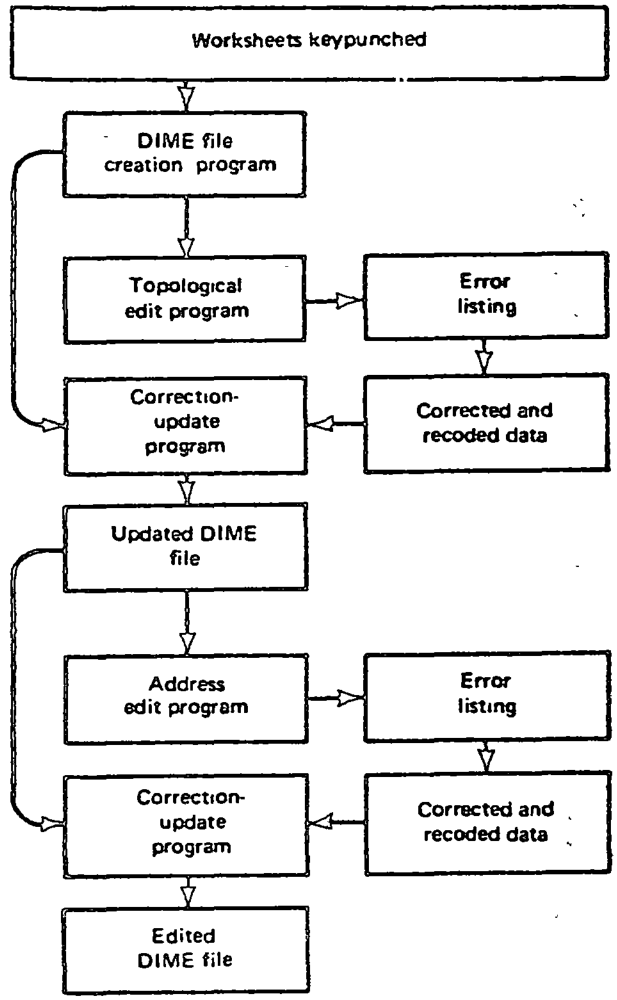


Figure 37. Computer processing of a DIME file.

Computer Requirements

DIME computer programs were written to run on IBM System 360/40 disk operating system with core capacity of 65K. The following devices are also required: A standard keypunch machine for key-punching the coding worksheets, a card reader for transmitting punched card data input, and a tape drive for the DIME master file. The DIME file may be held on direct-access (disk) storage, in which case sufficient space must be available. Direct-access storage, preferably on disk, is also needed to hold intermediate files. All programs are written in G-level FORTRAN IV.

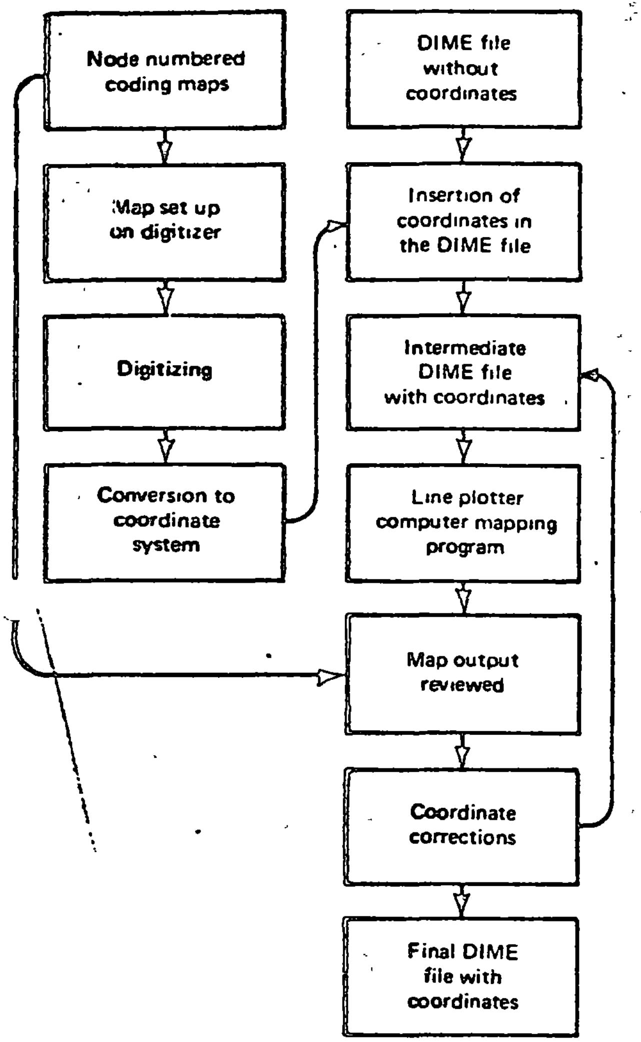


Figure 38. Coordinate digitization and insertion process.

The DIME computer package tapes are available in industry compatible seven- and nine-channel tape format. Documentation and technical specifications for the DIME computer edits and processing can be found in the computer manual of the DIME computer package.

Coding Maps

The requirements for coding maps are discussed fully in the clerical instructions of the DIME computer package.

Generally the scale of the coding maps should be within the range of 1 inch = 400 feet to 1 inch = 1,000 feet. Single-line maps are strongly recommended. They should contain all existing streets and street names, municipal boundaries, railroad tracks, and drainage features such as lakes and rivers.

Address Reference Materials

The address reference materials used in coding must contain the following information: Street names, intersecting streets (and, if possible, other intersecting features), and address numbers or ranges between intersects. The even-odd address number dichotomy should be apparent. Usually address reference materials will be maps, but they may also be street address listings, street or city directories, or some other type of index. All address reference material should be field checked for accuracy before using. If no accurate reference material can be located, the field listing form and instructions provided in the clerical instructions should be followed.

Clerical Coding

The clerical instruction manual contained in the DIME package describes the various steps and processes necessary to complete the coding phase of DIME file creation. It contains chapters on personnel and space requirements, materials to be used in coding, address reference materials, coding maps, the preparation of coding maps for coding, the preparation of special instructions for local problems, assignment preparation, training, and supervision of coding and postcoding operations. It also contains appendixes with sample forms, field listing instructions, and a coder's manual containing complete coding instructions.

The chapter on supervision of coding operations contains sections dealing with making assignments, controlling the operational flow, checking coders' work, and recordkeeping. It also provides technical procedures for coding regular and irregular or arbitrary address ranges and systems, problems such as unnamed streets, unknown street names, proposed streets, nonvehicular streets, addressable features other than streets, and the adding and deleting of node numbers.

Computer Processing

The computer procedures manual of the DIME package describes the computer processing steps required in a DIME operation. It also describes the overall system design and the hardware and software environment.

Each of the four computer programs comprising the system are described in full. These four programs are (1) the master file creation program, (2) the topological edit program, (3) the address edit program, and (4) the correction/update program. The section on master file creation discusses input requirements, program operation diagnostics, and output. The topological edit section includes a description of the input, editing procedures, output, and recoding. The address edit section includes input requirements, editing procedures, and output. The correction/update program section includes input requirements, diagnostics, and output.

Also included are descriptions of the sort/merge program, job control language, and sample listings.

Coordinate Insertion

Coordinates can be measured and recorded from the node-numbered coding maps by using a digitizer or other coordinate reader and they can then be converted to geographic coordinates and inserted into the DIME file.

For those areas of the country in the Census Bureau's ACG/DIME program, coordinates will be added to the file by the Bureau as an integral part of the program. Coordinates in this program will be latitude longitude carried to four decimal places for all nodes in the file. There are also plans to convert the latitude-longitude coordinates to state plane coordinates as an option.

Agencies that want to create their own DIME file will have to contract for digitizing services. Information on the technical aspects of coordinates is provided in chapter III.

The Census Use Study has had limited experience with organizations that perform this service. However, it is estimated that the cost of digitizing will run from 10 to 15 cents per node; or for an area of 1,000 blocks at 2.5 nodes per block, S250 to S375. These costs do not include conversion of the coordinates or insertion of the coordinates into the DIME file, a cost which could equal the digitizing cost.

File Maintenance

The Bureau plans to maintain the ACG/DIME geographic base files, probably updating them yearly. However, no operating system for updating and maintaining the files has yet been devised. One of the primary responsibilities of the Southern California Regional Information Study (SCRIS) is research into and development of operational methods for the periodic updating and maintenance of the interrelated ACG/DIME geographic base files and metropolitan mapping series.

If the Census Bureau receives approval from the Congress for a quinquennial census in 1975, updating of the geographic base files and the maps must begin in 1972 or 1973. Greater mail coverage in the census will require expansion of the coverage of the files and maps. If the ACG/DIME geographic base files are used for area sampling after 1970, it will perhaps be necessary for the Bureau to update the files on a more frequent basis. In any case, a technique and procedural requirements will be developed not only for use by the Bureau but also for use by local agencies who desire to update the files and maps periodically. The Bureau may provide a complete package for updating, maintenance, and expansion of the files. As research at SCRIS and the Census Bureau progresses, reports will be prepared describing the techniques and systems for file maintenance and updating.

Availability of Census Bureau Geographic Files

The alphabetical listing below includes all 233 standard metropolitan statistical areas. Listing anomalies are explained in the footnotes. There may be future additions to or deletions from this listing.

Address coding guides (ACG's) are now available, at least in preliminary form, for all mail census SMSA's. Final ACG's for all SMSA's will become available during the first quarter of 1970. They will be available on either seven- or nine-track magnetic tape or on high-speed printer output. As an example of costs, the preliminary ACG for the Fort Wayne, Ind. SMSA, with an estimated 1966 population of 264,000, costs $36 for a seven-track and $42 for a nine-track tape. Another example, the Milwaukee Wis. SMSA, with an estimated 1966 population of 1,335,000, costs $52 for a seven-track and $74 for a nine-track tape. Further information and order forms can be obtained from Central Users Service, Bureau of the Census, Washington, D.C. 20233.

The ACG/DIME or ACG Improvement Program geographic base files for mail census SMSA's will not begin to become available until late 1970. The SMSA's listed below are participating in the program or have expressed their intent to participate in the program.

ACG/DIME files for nonmail census SMSA's will begin to become available in mid-1970. Further information can be obtained from the Central Users Service, Bureau of the Census.

SMSA	Mail census ACG	Mail census ACG/DIME	Nonmail census ACG/DIME
Abilene, Tex.			X
Akron, Ohio	X	X	
Albany, Ga.			X
Albany-Schenectady-Troy, N.Y.	X	(1)	
Albuquerque, N. Mex.			X
Allentown-Bethlehem-Easton, Pa.-N.J.	X	X	
Altoona, Pa.	X	X	
Amarillo, Tex.			X
Anaheim-Santa Ana-Garden Grove, Calif.	X	X	
Anderson, Ind.	X	(1)	
Ann Arbor, Mich.	X	X	
Asheville, N.C.			X
Atlanta, Ga.	X	X	
Atlantic City, N.J.	X	(1)	
Augusta, Ga.-S.C.			X
Austin, Tex.			X
Bakersfield, Calif.			X
Baltimore, Md.	X	X	
Baton Rouge, La.			X
Bay City, Mich.	X	X	
Beaumont-Port Arthur-Orange, Tex.	X	X	
Billings, Mont.			X
Biloxi-Gulfport, Miss.			(2)
Binghamton, N.Y.-Pa.			X
Birmingham, Ala.	X	X	

See footnotes at end of list.

SMSA	Mail census ACG	Mail census ACG/DIME	Nonmail census ACG/DIME
Bloomington-Normal, Ill.	X	(1)	
Boise City, Idaho			X
Boston, Mass.	X	X	
Bridgeport, Conn.	X	(1)	
Brockton, Mass.	X	X	
Brownsville-Harlingen-San Benito, Tex.			X
Buffalo, N.Y.	X	X	
Canton, Ohio	X	X	
Cedar Rapids, Iowa			X
Champaign-Urbana, Ill.	X	(1)	
Charleston, S.C.			X
Charleston, W. Va.			X
Charlotte, N.C.	X	X	
Chattanooga, Tenn.-Ga.			X
Chicago, Ill.	X	X	
Cincinnati, Ohio-Ky.-Ind.	X	X	
Cleveland, Ohio	X	X	
Colorado Springs, Colo.			X
Columbia, S.C.			X
Columbus, Ga.-Ala.			X
Columbus, Ohio	X	X	
Corpus Christi, Tex.			X
Dallas, Tex.	X	X	
Davenport-Rock Island-Moline, Iowa-Ill.	X	X	
Dayton, Ohio	X	X	
Decatur, Ill.	X	(1)	
Denver, Colo.	X	X	
Des Moines, Iowa			X
Detroit, Mich.	X	X	
Dubuque, Iowa			X
Duluth-Superior, Minn.-Wis.			X
Durham, N.C.	X	X	
El Paso, Tex.[3]	X	X	
Erie, Pa.			X
Eugene, Oreg.	X	X	
Evansville, Ind.-Ky.			X
Fall River, Mass.-R I.	X	X	
Fargo-Moorhead, N. Dak.-Minn.			X
Fayetteville, N.C.			X
Fitchburg-Leominster, Mass.	X	X	
Flint, Mich.	X	X	
Fort Lauderdale-Hollywood, Fla.	X	X	
Fort Smith, Ark -Okla.			X
Fort Wayne, Ind.	X	X	
Fort Worth, Tex.	X	X	
Fresno, Calif.			X
Gadsden, Ala.			X
Galveston-Texas City, Tex.	X	X	
Gary-Hammond-East Chicago, Ind.	X	X	
Grand Rapids, Mich.	X	(1)	
Great Falls, Mont.			X
Green Bay, Wis.	X	X	
Greensboro-Winston-Salem-High Point, N.C.	X	(1)	
Greenville, S.C.			X
Hamilton-Middletown, Ohio	X	X	

See footnotes at end of list.

SMSA	Mail census ACG	Mail census ACG/DIME	Nonmail census ACG/DIME
Harrisburg, Pa.	X	(1)	
Hartford, Conn.	X	X	
Honolulu, Hawaii			X
Houston, Tex.	X	X	
Huntington-Ashland, W. Va.-Ky.-Ohio			X
Huntsville, Ala.			X
Indianapolis, Ind.	X	X	
Jackson, Mich.	X	(1)	
Jackson, Miss.			X
Jacksonville, Fla.	X	X	
Jersey City, N.J.	X	X	
Johnstown, Pa.	X	X	
Kalamazoo, Mich.	X	X	
Kansas City, Mo.-Kans.	X	X	
Kenosha, Wis.	X	X	
Knoxville, Tenn.			X
Lafayette, La.			(2)
fayette-West Lafayette, Ind.			(2)
e Charles, La.			(2)
Lancaster, Pa.	X	(1)	
Lansing, Mich.	X	X	
Laredo, Tex.			X
Las Vegas, Nev.			X
Lawrence-Haverhill, Mass.-N.H.	X	X	
Lawton, Okla.			X
Lewiston-Auburn, Maine			X
Lexington, Ky.			X
Lima, Ohio	X	X	
Lincoln, Nebr.			X
Little Rock-North Little Rock, Ark.			X
Lorain-Elyria, Ohio	X	X	
Los Angeles-Long Beach, Calif.	X	X	
Louisville, Ky.-Ind.	X	X	
Lowell, Mass.	X	X	
Lubbock, Tex.			X
Lynchburg, Va.			X
Macon, Ga.			X
Madison, Wis.	X	X	
chester, N.H.			X
Mansfield, Ohio	X	X	
Mayaguez, P.R.			(4)
McAllen-Pharr-Edinburg, Tex.			X
Memphis, Tenn.-Ark.	X	X	
Meriden, Conn.	X	(1)	
Miami, Fla.	X	X	
Midland, Tex.			X
Milwaukee, Wis.	X	X	
Minneapolis-St. Paul, Minn.	X	X	
Mobile, Ala.	X	X	
Monroe, La.			X
Montgomery, Ala.			X
Muncie, Ind.	X	(1)	
Muskegon-Muskegon Heights, Mich.	X	(1)	
Nashville, Tenn.	X	(1)	
New Bedford, Mass.	X	X	

See footnotes at end of list.

SMSA	Mail census ACG	Mail census ACG/DIME	Nonmail census ACG/DIME
New Britain, Conn.	X	X	
New Haven, Conn.	X	X	
New London-Groton-Norwich, Conn.	X	(1)	
New Orleans, La.	X	X	
New York, N.Y.	X	(5)	
Newark, N.J.	X	(1)	
Newport News-Hampton, Va.	X	X	
Norfolk-Portsmouth, Va.	X	X	
Norwalk, Conn.	X	(1)	
Odessa, Tex.			X
Ogden, Utah	X	X	
Oklahoma City, Okla.	X	X	
Omaha, Nebr.-Iowa	X	X	
Orlando, Fla.	X	(1)	
Oxnard-Ventura, Calif.	X	X	
Paterson-Clifton-Passaic, N.J.	X	(1)	
Pensacola, Fla.	X	(1)	
Peoria, Ill.	X	X	
Philadelphia, Pa.-N.J.	X	X	
Phoenix, Ariz.	X	(1)	
Pine Bluff, Ark.			X
Pittsburgh, Pa.	X	X	
Pittsfield, Mass.	X	X	
Ponce, P.R.			(4)
Portland, Maine			X
Portland, Oreg.-Wash.	X	X	
Providence-Pawtucket-Warwick, R.I.-Mass.	X	X	
Provo-Orem, Utah			X
Pueblo, Colo.			X
Racine, Wis.	X	X	
Raleigh, N.C.	X	X	
Reading, Pa.	X	X	
Reno, Nev.			X
Richmond, Va.	X	X	
Roanoke, Va.			X
Rochester, N.Y.	X	X	
Rockford, Ill.	X	X	
Sacramento, Calif.	X	X	
Saginaw, Mich.	X	X	
Salem, Oreg.			X
Salinas-Monterey, Calif.			X
St. Joseph, Mo.			X
St. Louis, Mo.-Ill.	X	X	
Salt Lake City, Utah	X	X	
San Angelo, Tex.			X
San Antonio, Tex.	X	X	
San Bernardino-Riverside-Ontario, Calif.	X	(6)	
San Diego, Calif.	X	X	
San Francisco-Oakland, Calif.	X	X	
San Jose, Calif.	X	X	
San Juan, P.R.			(4)
Santa Barbara, Calif.			X
Savannah, Ga.[3]	X	X	
Scranton, Pa.	X	X	
Seattle-Everett, Wash.	X	X	

See footnotes at end of list.

SMSA	Mail census ACG	Mail census ACG/DIME	Nonmail census ACG/DIME
Sherman-Denison, Tex.			X
Shreveport, La.			X
Sioux City, Iowa-Nebr.			X
Sioux Falls, S. Dak.			X
South Bend, Ind.	X	X	
Spokane, Wash.			X
Springfield, Ill.	X	(1)	
Springfield, Mo.			X
Springfield, Ohio	X	X	
Springfield-Chicopee-Holyoke, Mass.-Conn.	X	X	
Stamford, Conn.	X	(1)	
Steubenville-Weirton, Ohio-W. Va.	X	X	
Stockton, Calif.	X	X	
Syracuse, N.Y.	X	X	
Tacoma, Wash.	X	X	
Tallahassee, Fla.	X	X	
Tampa-St. Petersburg, Fla.	X	X	
Terre Haute, Ind.	X	X	
Texarkana, Tex.-Ark.			X
Toledo, Ohio-Mich.	X	X	
Topeka, Kans.	X	X	
Trenton, N.J.	X	X	
Tucson, Ariz.			X
Tulsa, Okla.	X	X	
Tuscaloosa, Ala.			X
Tyler, Tex.			X
Utica-Rome, N.Y.	X	X	
Vallejo-Napa, Calif.	X	X	
Vineland-Millville-Bridgeton, N.J.	X	(1)	
Waco, Tex.			X
Washington, D.C.-Md.-Va.	X	X	
Waterbury, Conn.	X	(1)	
Waterloo, Iowa			X
West Palm Beach, Fla.	X	X	
Wheeling, W. Va.-Ohio	X	X	
Wichita, Kans.	X	X	
Wichita Falls, Tex.			X
Wilkes-Barre-Hazleton, Pa.	X	X	
Wilmington, Del.-N.J.-Md.	X	X	
Wilmington, N.C.			(2)
Worcester, Mass.	X	X	
York, Pa.	X	X	
Youngstown-Warren, Ohio	X	X	

[1] In ACG program, declined participation in ACG/DIME program.

[2] Declined participation in nonmail program. Census tract coding guide to be prepared.

[3] In mail census ACG and ACG/DIME program although a nonmail census area.

[4] The three Puerto Rican SMSA's are scheduled eventually to be in the nonmail program. However, at this time only the Santurce and Old San Juan portions of the San Juan SMSA are in the program.

[5] Nassau and Suffolk counties are in the ACG/DIME program, remainder of SMSA is not.

[6] San Bernardino County portion is in the ACG/DIME program, Riverside County portion is not in the program.

Metropolitan Mapping Series

Traditionally the Census Bureau has relied on maps supplied by local governments to conduct censuses. Letters were sent to county and municipal governments a year or two prior to a census, asking for street maps that could be used as enumeration maps. Municipal boundaries, enumeration district boundaries, and other census administrative and statistical area boundaries are added to the maps by the Bureau. They were then reproduced for enumeration and supervisory purposes for the census. This procedure, although cumbersome, did allow the census to be taken more or less efficiently, at least in the less densely populated areas of the Nation. For the 1970 census, this procedure will be used in those portions of the country outside the urbanized areas covered by the Metropolitan Mapping Series (MMS) program.

However, the Census Bureau found while conducting the 1960 census that it was becoming difficult to administer the census because of the great number and differing scales of maps received from the communities in these urban centers. Problems of overlapping maps and control of enumeration districts were endemic. As soon as the census was concluded, proposals were developed to create series of street maps at a standard scale for use in censuses and surveys. These proposals developed into the metropolitan mapping series program.

For this program two basic inputs were needed: U.S. Geological Survey topographic maps for control purposes and maps provided by local planning agencies or municipal governments for defining current street patterns, street names, and other pertinent features.

The early maps produced in the MMS program were created from these two sources by Bureau cartographers at a scale of 1 inch = 880 feet This was later changed to 1 inch = 800 feet. The maps are intended for statistical purposes and therefore are not of engineering quality, although they are generally sufficiently reliable for most planning and administrative purposes. The maps were designed to contain all existing vehicular thoroughfares and their names, railroad tracks, major drainage features, shorelines, lakes, and prominent landmarks that occupy large areas, and have clearly defined boundaries. They also contain State, county, congressional district, minor civil division or census county division, municipal, and census tract boundaries. Block numbers, census tract numbers, State and county names, and place names and codes are also contained on the maps. Maps have presently been drafted for the urbanized portions of all 233 SMSA's. Figure 39 illustrates a portion of an MMS map at half scale (i.e., 1" = 1600').

Once the maps were developed it became imperative to develop an updating procedure so that the maps would be reasonably accurate at census time. The decision was made to involve county and regional planning agencies in this program because of their obvious interest and expertise in the field and their metropolitan-wide coverage. The Department of Housing and Urban Development recognized the usefulness of these maps to planners and made available "701" grant funds for local review. The appropriate agencies were contacted, and all cooperated in the local review process.

The MMS maps will be used for administrative purposes in the 1970 census and basic reference sources for the local coding phase of both the ACG and ACG/DIME geographic base file programs. A reduced scale version (1" = 2000') of the maps will be published for use with the 1970 census results. Preliminary maps at both scales are available at a nominal cost. Final maps will begin to become available in late 1970 or early 1971. A catalog listing available Census Bureau maps, including MMS maps, will be issued during 1970. Inquiries concerning the MMS and other Bureau mapping programs should be directed to Central Users Service, Bureau of the Census, Washington, D.C. 20233.

The MMS will be updated after the 1970 census for future censuses and surveys Coverage of the series may be extended so that they will be of greater use, not only to the Census Bureau but also to map users around the country.

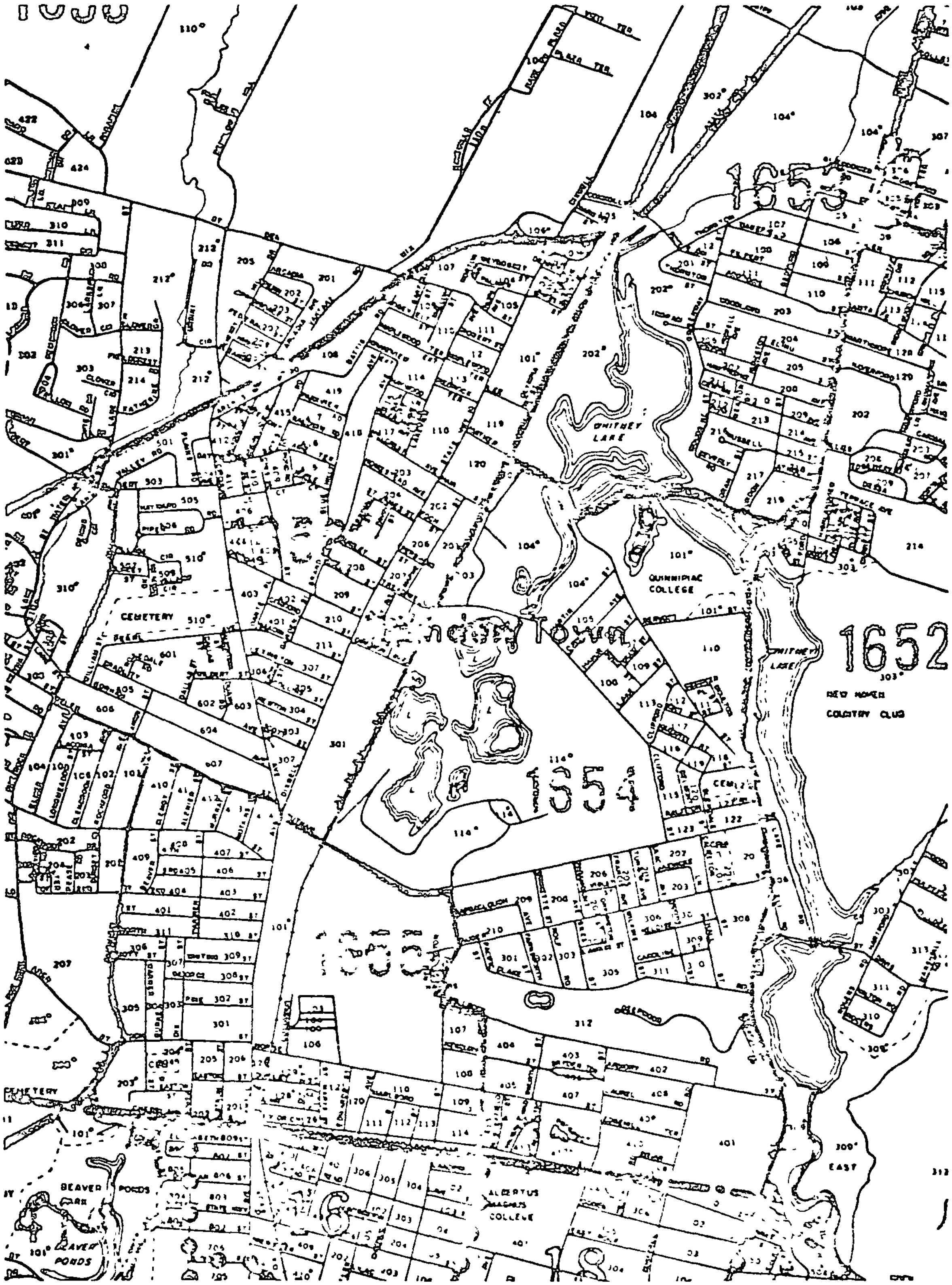


Figure 39. Metropolitan mapping series

Data Available From the 1970 Census

A variety of data products and services will be available from the 1970 Census of Population and Housing data base. A set of standard tabulations will be available in a variety of media including micro-film, punched cards, printed reports, and computer tapes. Public use samples containing data for samples of individuals from the basic record tapes, but without identifying information, will also be avail-[illegible]. In addition, other services may be obtained at user request and expense, including special tabulations of the basic record tapes, and computer-generated analytical reports.

As in earlier censuses, printed reports will be the end product of the 1970 census and will be available for purchase and in libraries. Printed reports offer the two main benefits of low cost and easy accessibility. On the other hand, printed reports will not be available until several months after the tabulations have been completed; and they will contain restricted [illegible]graphic and subject matter detail.

A major element of the 1970 data delivery system will provide magnetic tapes of summary census data. These will have essentially the same subject matter scope as the printed reports, but with greater detail. T[illegible] summary tapes will begin to become available in [illegible] summer of 1970.

The first summary tapes available will be the First Count Summary Tapes, containing final complete-count population and housing data for States, counties, congressional districts, minor civil divisions, places, and enumeration districts or block groups. The subjects tabulated will include age, sex, color, marital status, relationship to head of household, population under 18 and over 65 years of age by household relationship, family type, crowding, tenure of occupied housing units, vacancy status, units in structure, rooms, plumbing facilities, basement, telephone, and value or rent. In addition, many tabulations are cross-classified by race.

Second Count Summary Tapes, containing complete-count population and housing data at the level of census tracts, minor civil divisions, and larger census areas, are expected to begin becoming available in late 1970. These tapes will include the same subjects as the first count tapes, but in much greater detail.

Third Count (Block) Summary Tapes, containing complete-count population and housing data for city blocks, are expected to begin becoming available in the spring of 1971. These tapes will contain fewer data than the tapes for the first or second counts; however, substantially more data will be carried on the block tapes than was available in 1960 or is expected to be printed in 1970.

Fourth Count Summary Tapes, including data down to the census tract or minor civil division level, are expected to be available in mid to late 1971. They will contain 20-, 15-, and 5-percent sample population and housing characteristics such as occupation, income, education, and household equipment and facilities.

Fifth Count Summary Tapes, which contain population and housing sample data summaries for ZIP code areas, are expected to be available in late 1971 or early 1972.

Since summary tapes must be created by the Bureau of the Census in order to process the census, the tapes for all five counts will be available to users at the cost of about $6[illegible] per reel. Sets of summary tapes may be ordered for each count by State. The number of reels varies by State and by count. The Third Count for California, for example, will be about 30 reels in all.

Microfilm will also be a generally available medium for obtaining general census tabulations. Printed reports issued from the 1970 census will be available on microfilm. The contents of the summary tape files will also be available in this manner.

An extended array of sample data on magnetic tape for public use is planned for the 1970 census. These tapes will be made available to users at the cost of reproduction, together with appropriate documentation. Although public use samples are limited in geographic detail, they are most useful to researchers interested in the relationships of demographic variables for individuals rather than aggregates.

Special tabulations of census basic record tapes will also be made available. For this service, computer programs are created to meet the specific needs of the user in obtaining data summaries for geographic areas not recognized in general tabulations and/or including subject breakdowns or cross-classifications not appearing in general tabulations. Users are charged for planning, programing, clerical, machine, or other costs necessary to provide this service.

Another type of special service is matching studies. Matching studies take a series of individual records furnished by a user, link them to their respective census questionnaires, and prepare summaries of the census characteristics of the individuals. Individual data are never released.

Other special services for the 1970 census may include computer-produced analytic reports, computer graphics, and software packages for census data analvsis. Further information on 1970 census data products and services can be obtained from the Central Users Service, Bureau of the Census, Washington, D.C. 20233.

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

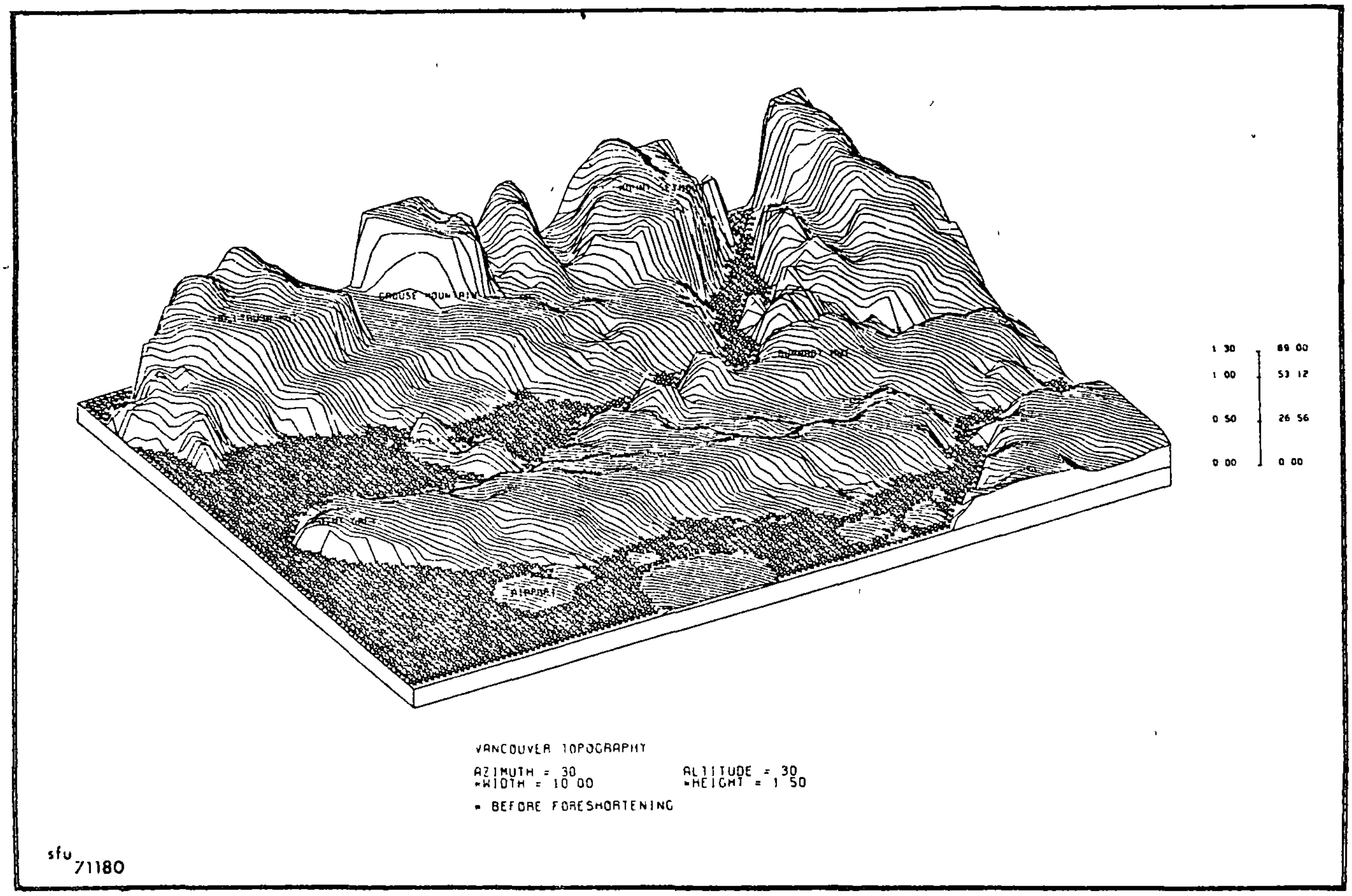


Figure 1.1 Perspective View of Greater Vancouver, View from S.W. A typical example of the advantages of computer cartography: the repetitive computation of the picture coordinates of a point in three-dimensional space (approximately 10,000 times in this case), followed by controlled plotting. Program 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

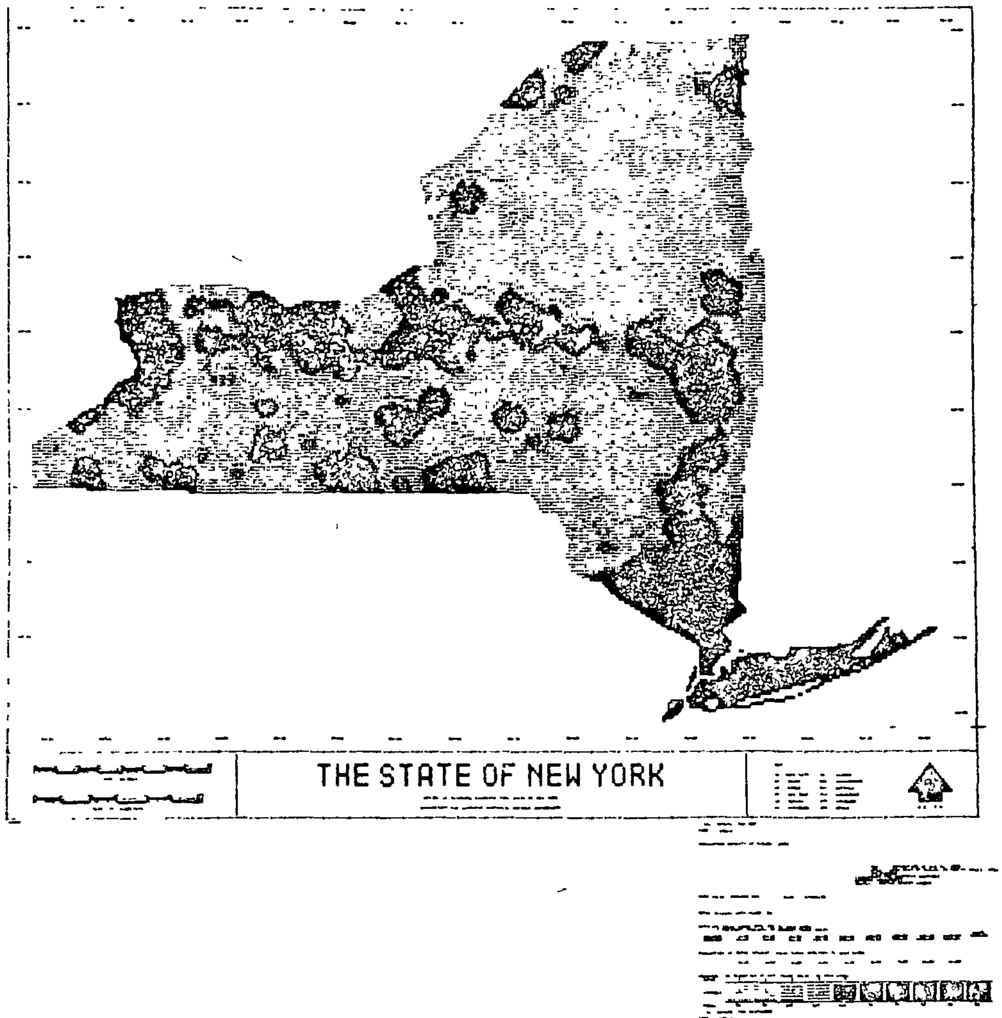


Figure 1 2 New York State, Population Density, 1990. A good example of a line-printer map. Program: SYMAP; data: Office of Planning Coordination, State of New York; production: C. Steinitz, Harvard University

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

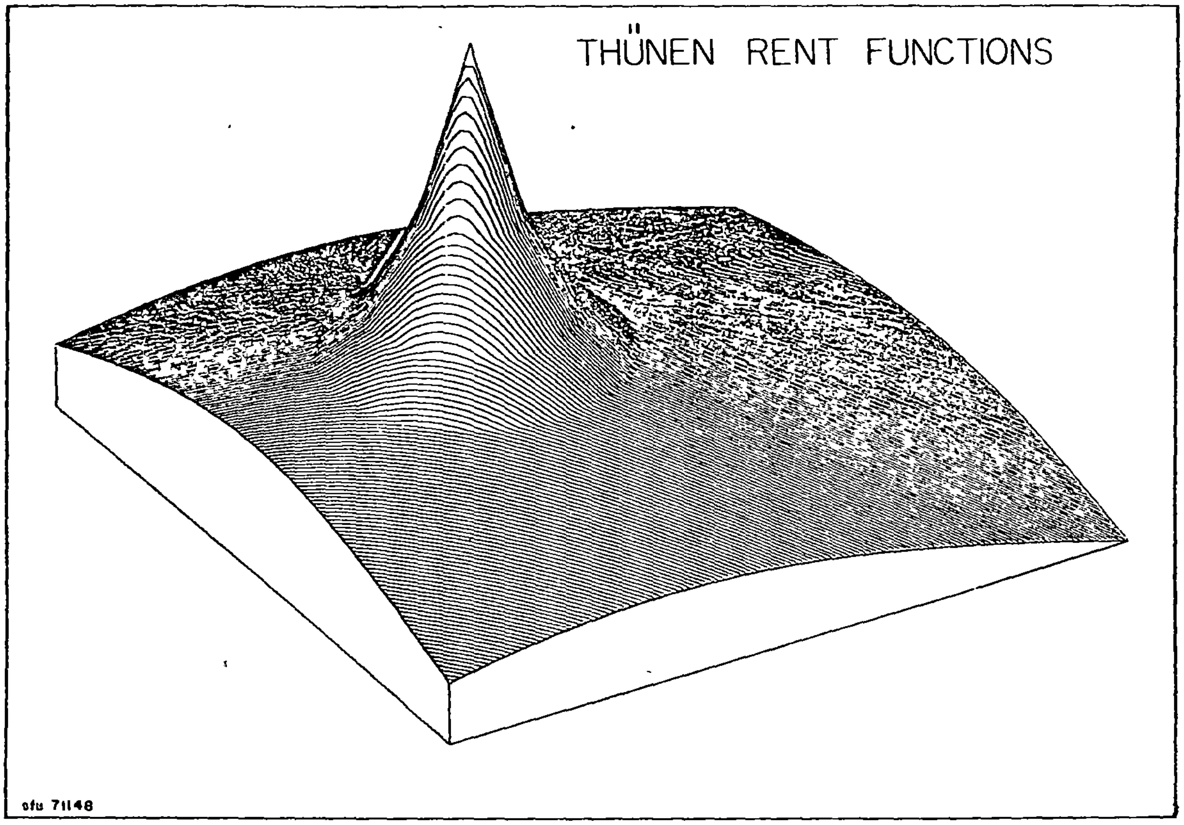


Figure 1.3 Three-Dimensional Rent Functions. The graphic representation of geographical models can aid our understanding of them. Programs· SPATFUN, SYMVU; production· T.K Peucker

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"

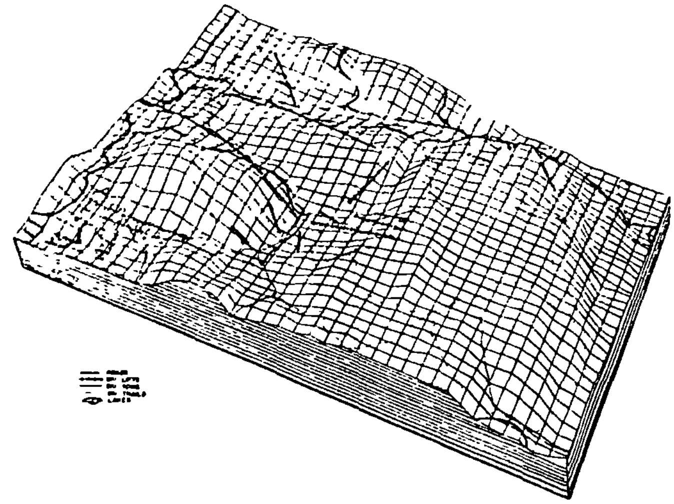


Figure 1.4 Ski Area. Perspective plots can aid planning and advertising. Program: VIEWBLOK; data: D. Douglas, University of Ottawa

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 = II = 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)

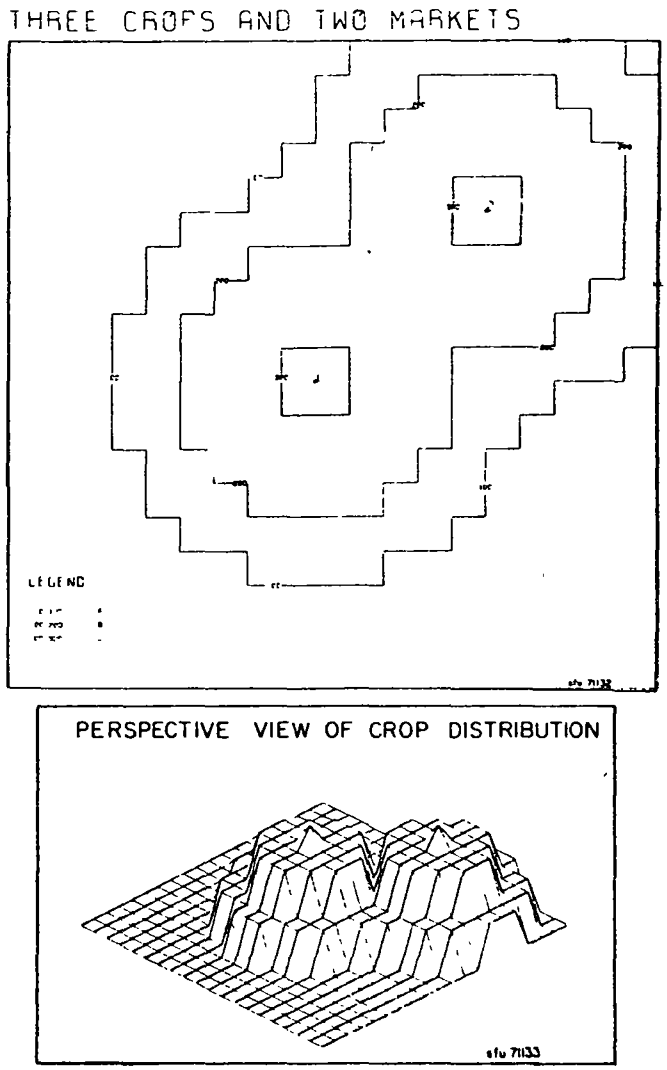


Figure 1.5 Crop Distribution Simulation. The simulation of land use with two markets based on the Thünen theory. Programs: SIM, PERS, CNTOUR: production: S. Witiuk, Simon Fraser University

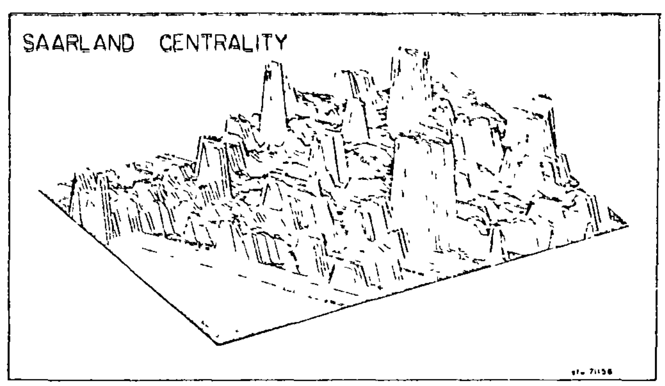


Figure 1 6 Saarland, Centrality The results of a factor analysis of 103 variables and 356 communities Factor 1, centrality Programs FAKAN, SYMVU, data and production W.D. Rase

---

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

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.

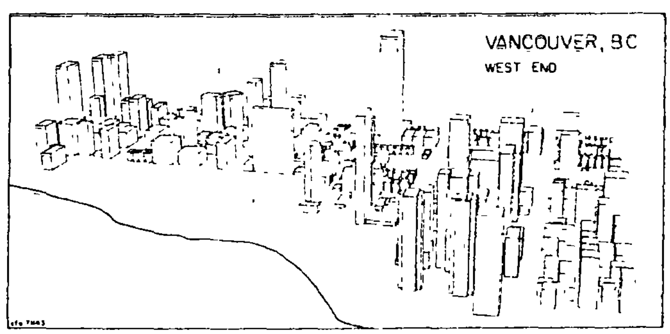


Figure 1.7 Vancouver, West End Vancouver's major high-rise apartment area All houses are represented by blocks The data bank of houses can easily be changed to show residential development The coastline is drawn in manually. Program BLOCKS, data· A Furguson coastline L Nelson, both from Simon Fraser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

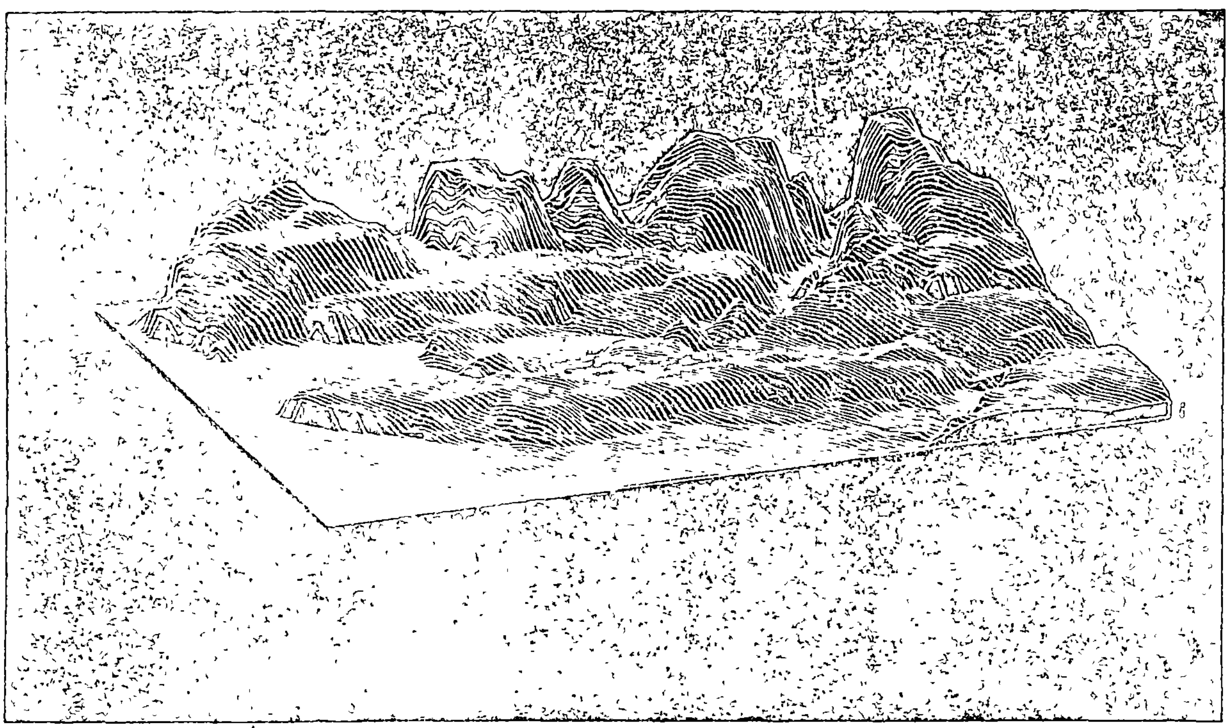


Figure 1 8 Vancouver This negative of a plot is an example of the signal means of cartographic representation Program SYMVU, data D Mark, production W D. Rase

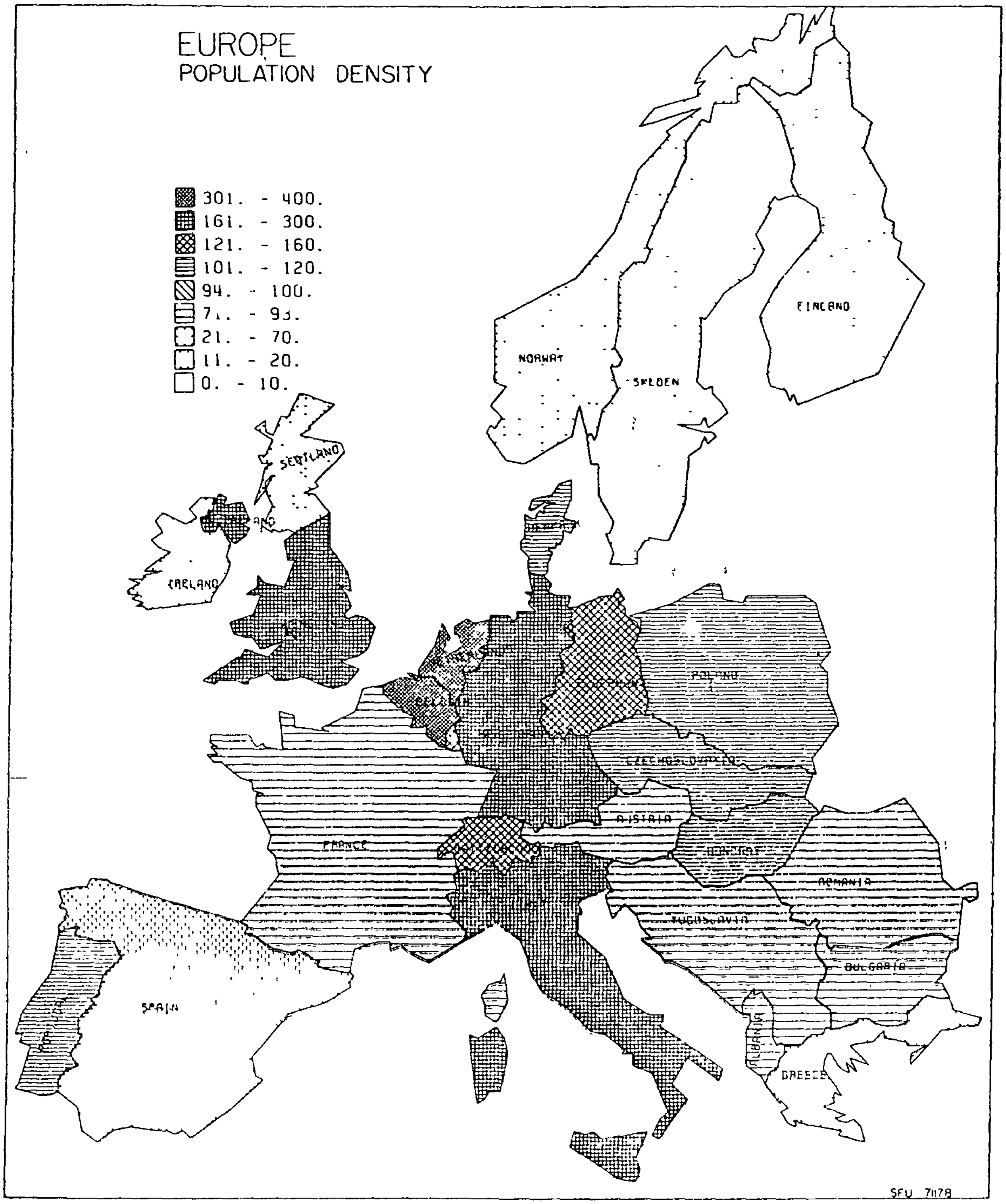


Figure 1 9 Europe, Population Density The plotter produces very fine shading but the computing and plotting time is prohibitive if used on a day-to-day basis Program CALFORM, data G Hayward, production C Johannson

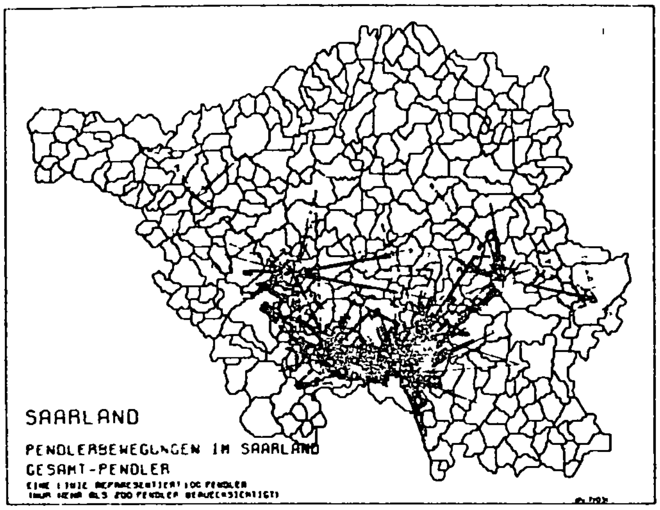


Figure 1 12 Saarland, Commuting. The commuting pattern in this German province. Program. PFEIL; data: W D. Rase

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)

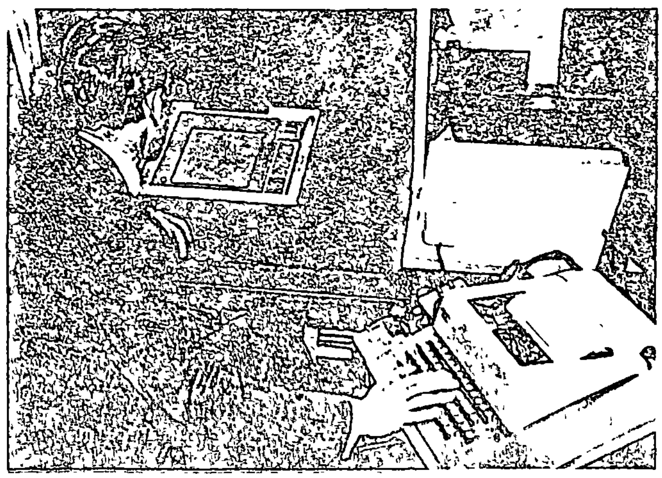


Figure 1.14 Interactive Storage Tube The tube is in an almost horizontal position for working convenience The operator has his right hand on a teletype console for alphanumeric interaction with the computer Between the teletype and the tube is a "mouse" for two-dimensional interactions with the computer 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 disc. This process would normally be the result of some previous computations 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

SKI-AREA STUDY

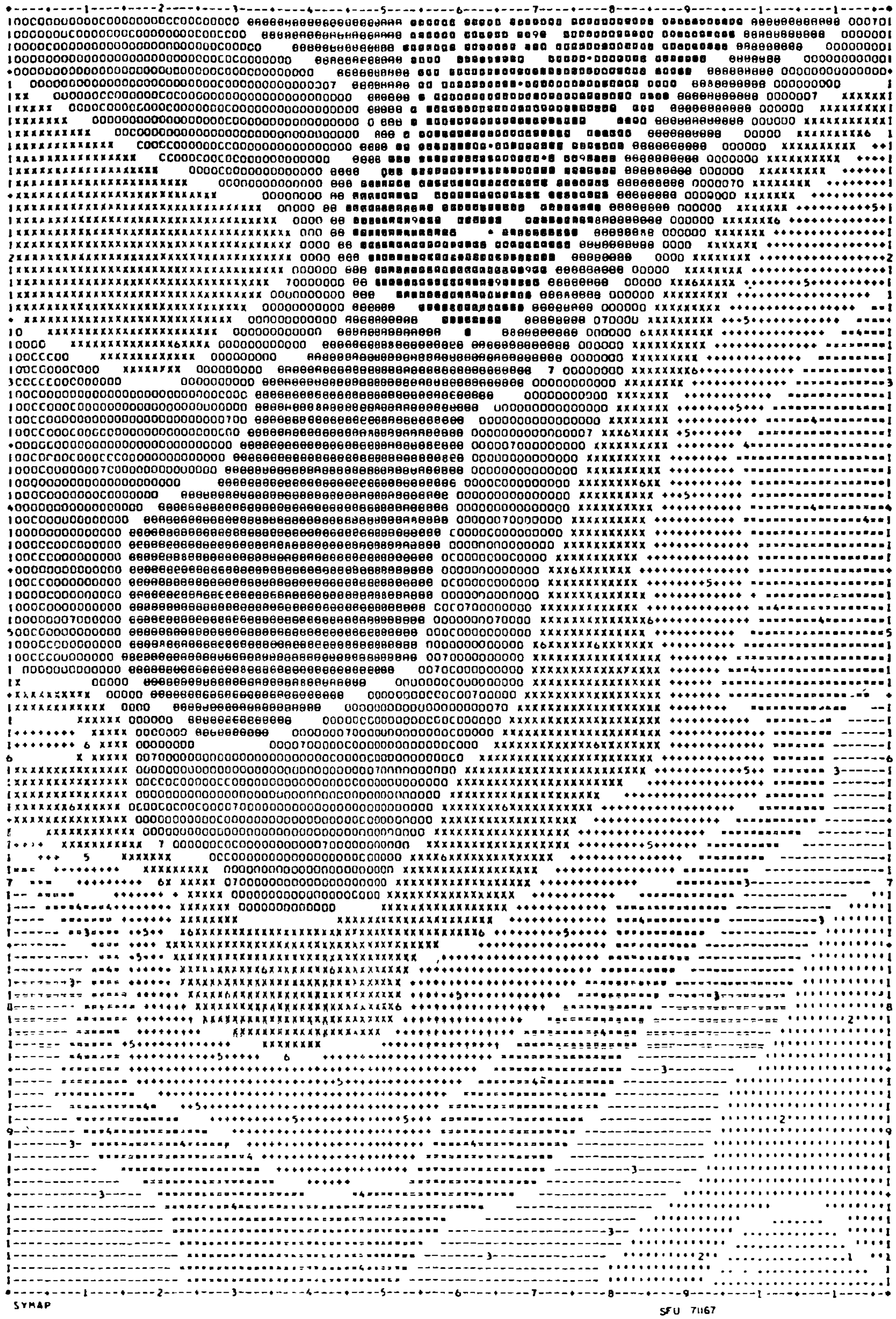


Figure 2 1 Grouse Mountain, North Vancouver Ten class intervals have been used for the contouring The map shows the different combinations of letters used The numbers show the location of data-points and their level Program SYMAP, data D Mark

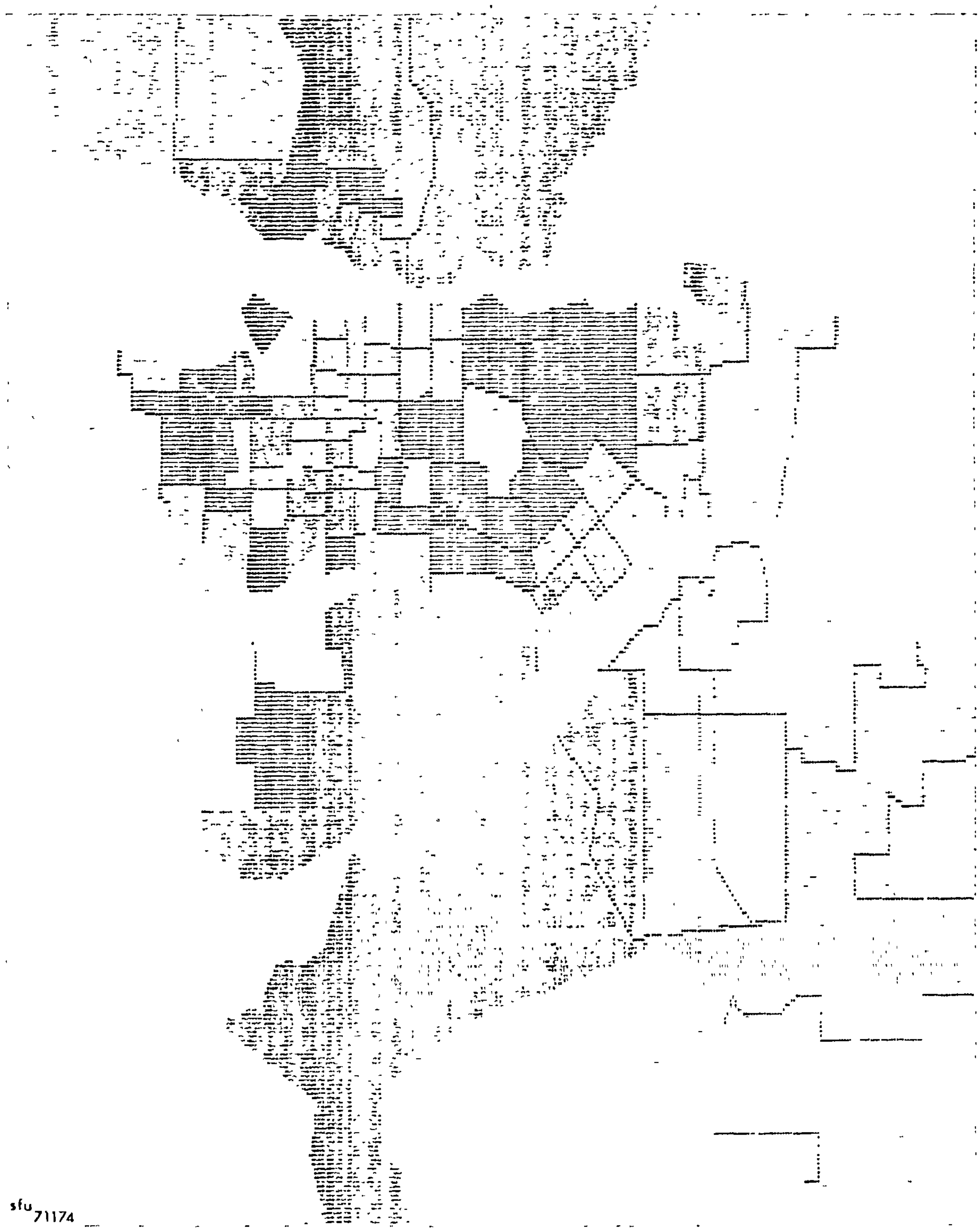


Figure 2.2 Vancouver A choropleth map showing the results of a factor analysis of the census of population factor (79 variables, 118 census tracts). Data T.K Peucker, production W D Rase

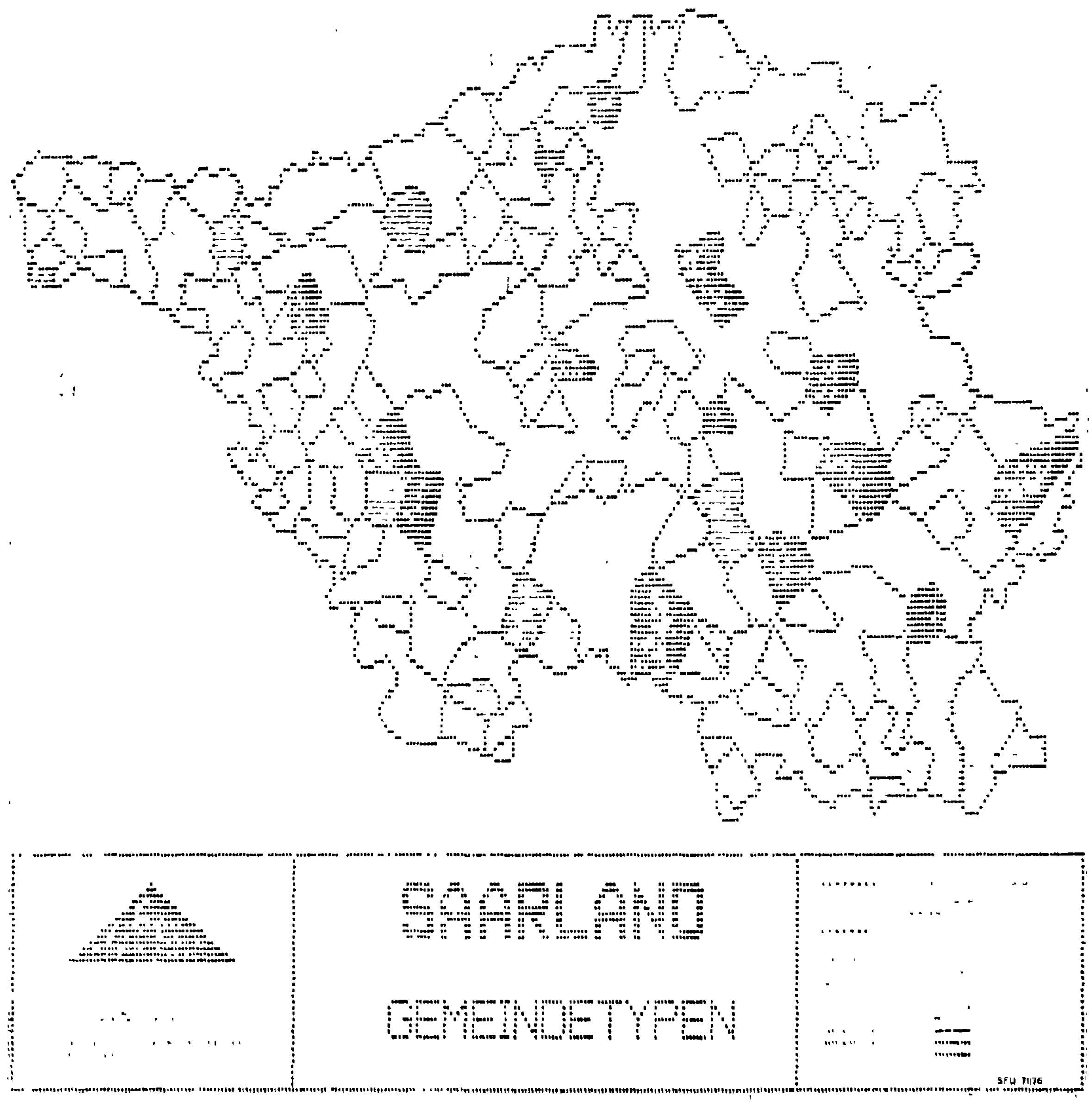


Figure 2 3 Saarland, Centrality. This example shows that the proximal-type map can be very suitable for classification mapping. The areas of little and high variation are well delineated by the density of borders Program SYMAP; data W.D. Rase

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.

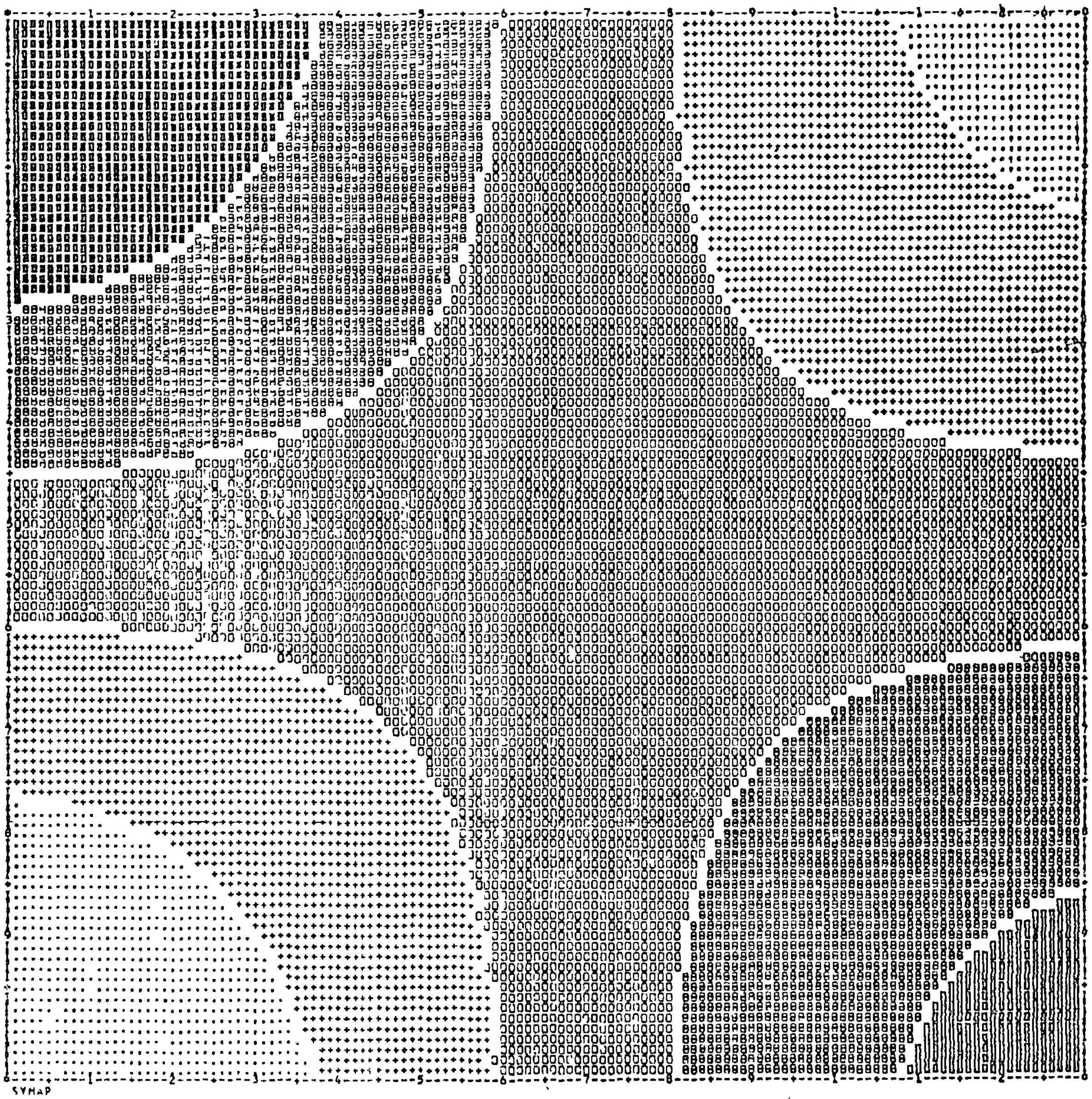


Figure 2 4 Saddle Point. The data-points are at the four corners of the square. No electives of the map package are used. Program: SYMAP; production: G. Hayward, Simon Fraser University

In addition to the outline, the base-map may also have internal boundaries such as those of states The conformoline-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)

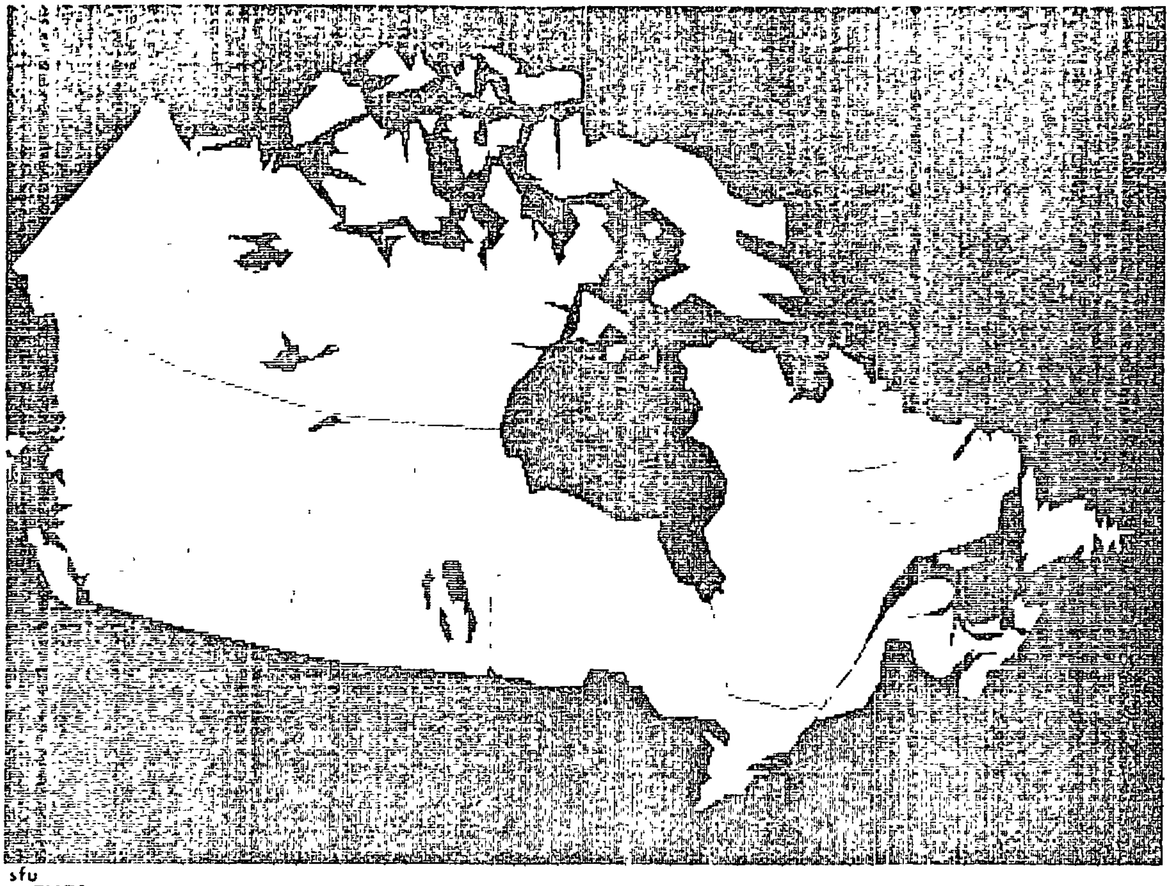


Figure 2 5 Canada Outline. The use of the outline package and the size elective.
Program SYMAP, data D. Hatlelid, Simon Fraser University

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.

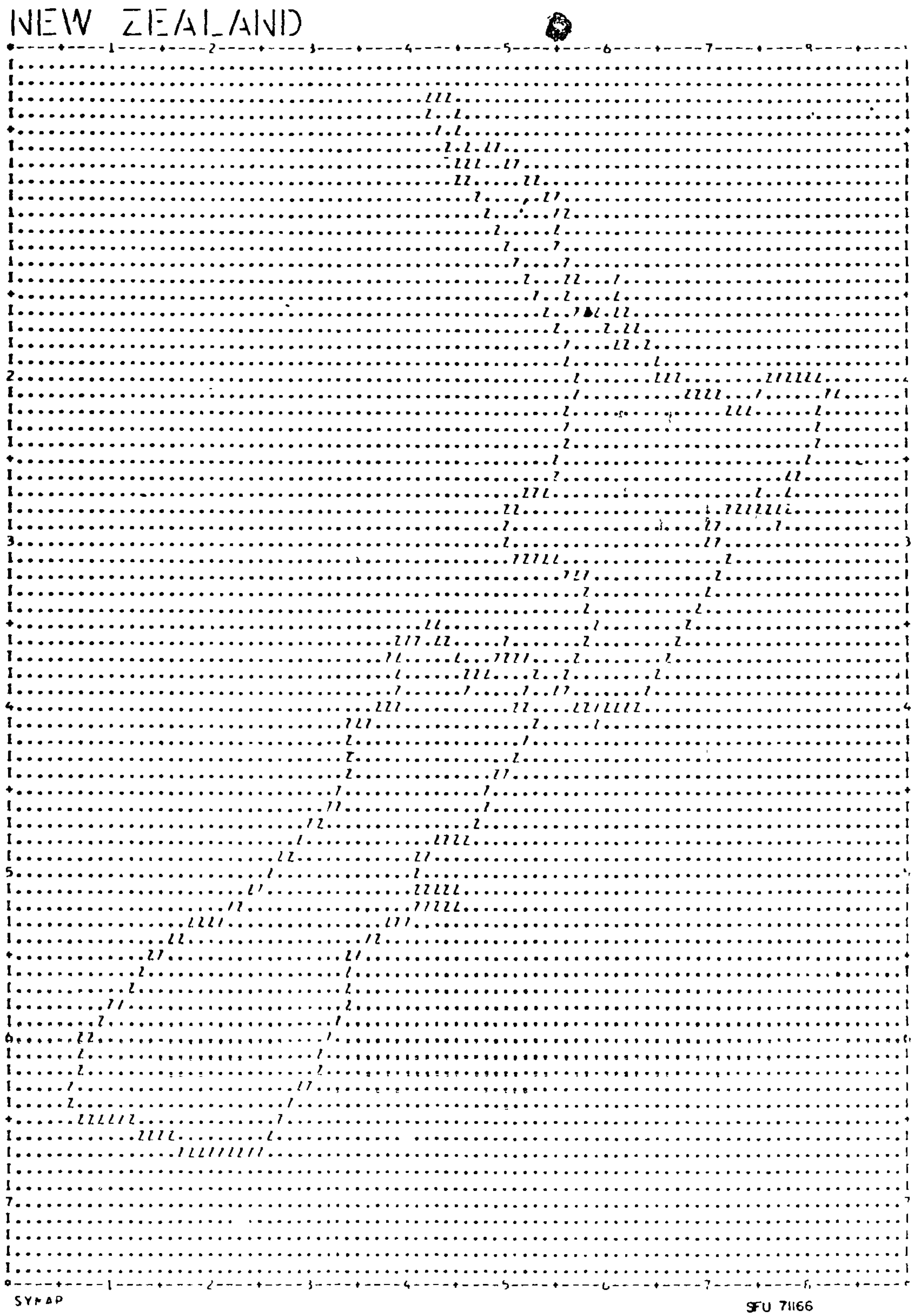


Figure 2 6 New Zealand The use of the otolegend Program SYMAP, data T K Peucker

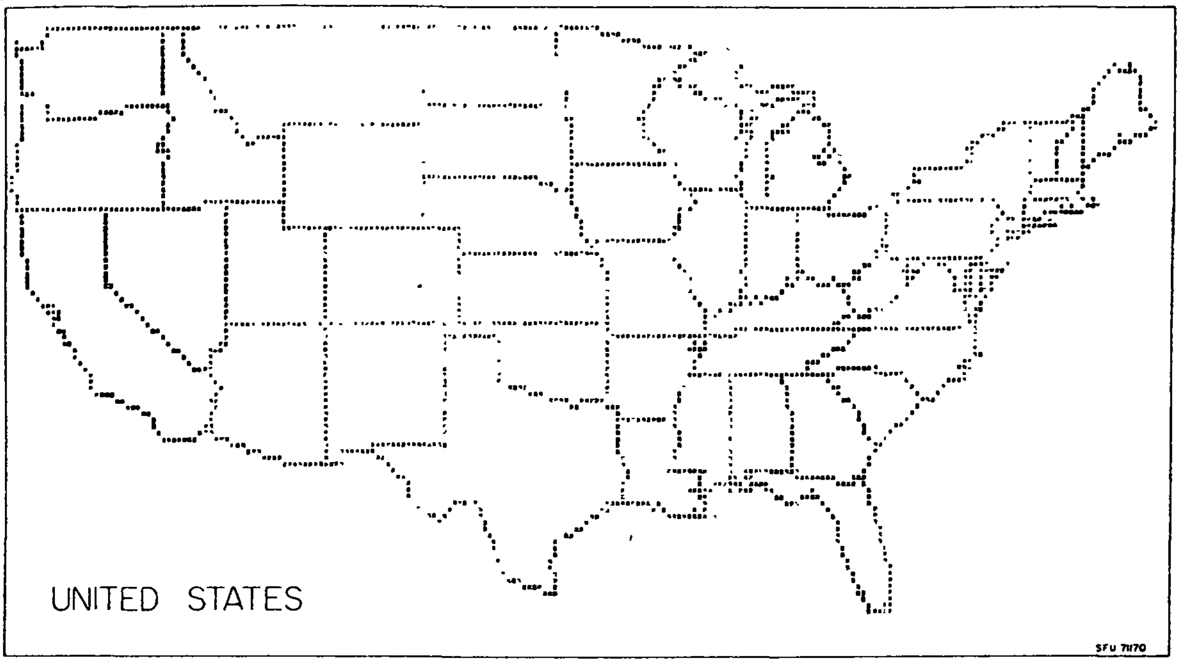


Figure 2.7 U S.A All the boundaries are produced using the otolegend package. Program: SYMAP, data T K. Peucker and R Mercready, Vienna, Virginia; production G. Hayward

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.

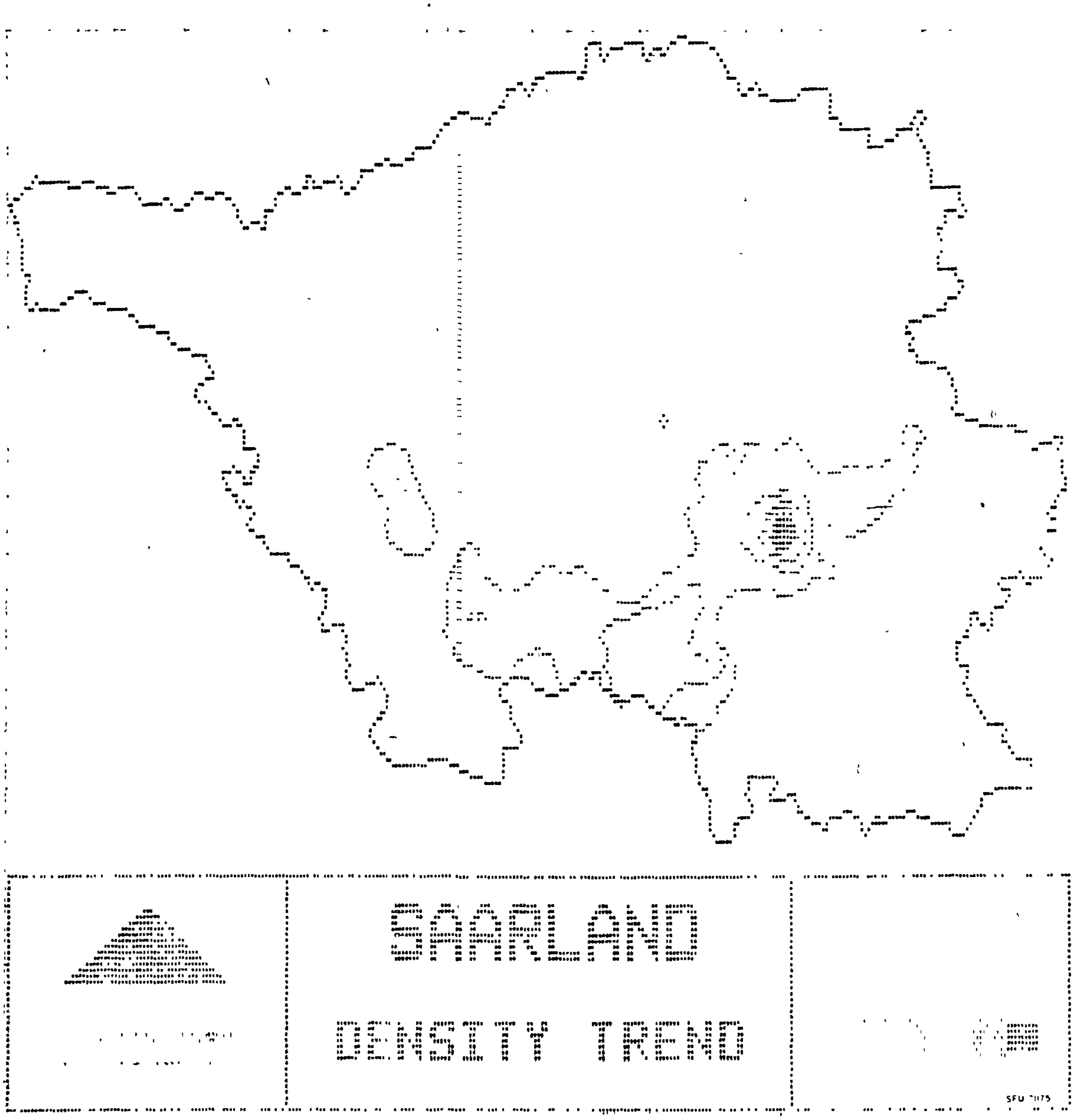


Figure 2.8 Saarland, Population Density In this case, the class intervals should have been altered The default option gives the urban population with four intervals and the rural density with only one It would have been interesting to differentiate since there are distinct variations closely related to the distribution of forest Program SYMAP, data W D Rase

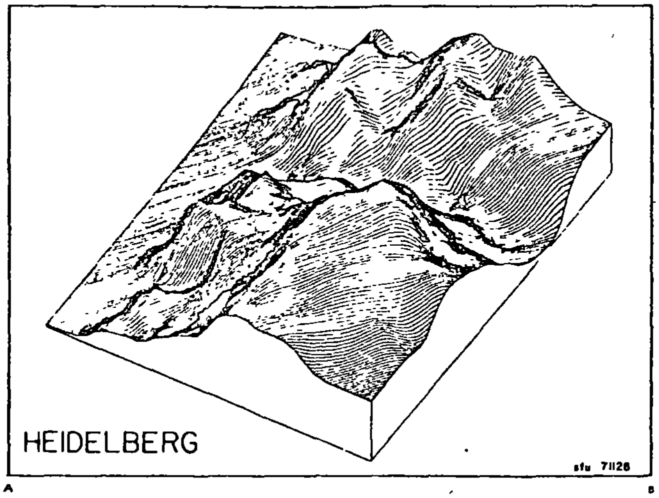


Figure 2.9 Heidelberg, View from S.E. The surface was digitized irregularly, interpolated by SYMAP, and the regular grid used for the perspective plot. Approximately 400 original points, 10,000 grid-points. Programs: SYMAP, SYMVU; data: T.K. Peucker

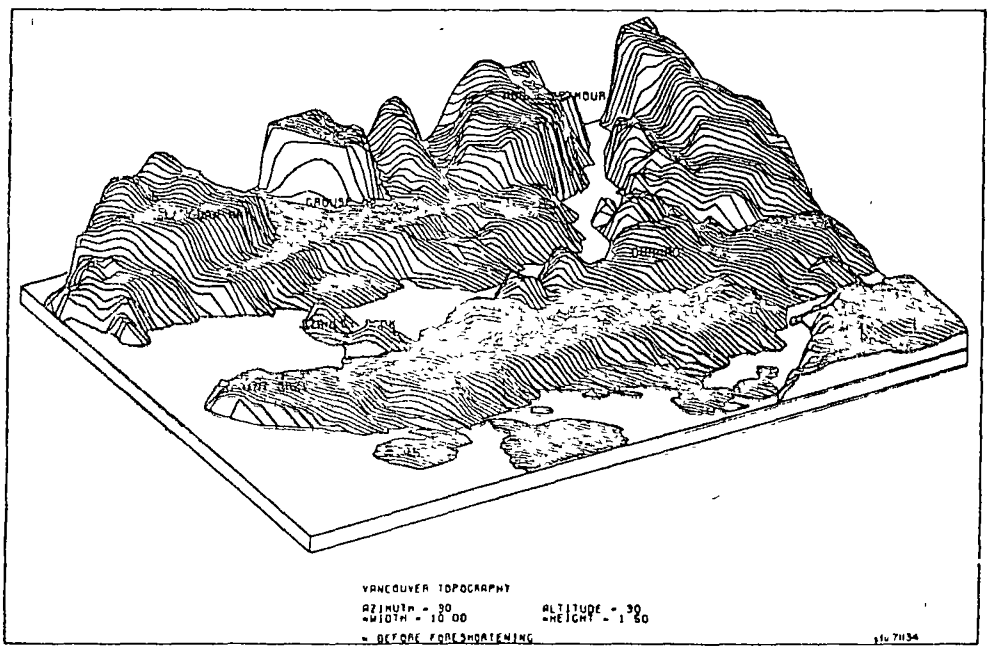


Figure 2 10 Vancouver, View from S.W Approximately 640 points were digitized, interpolated to 10,191 grid-points, and then plotted. Programs SYMAP, SYMVU, data. D Mark, production: W D Rase

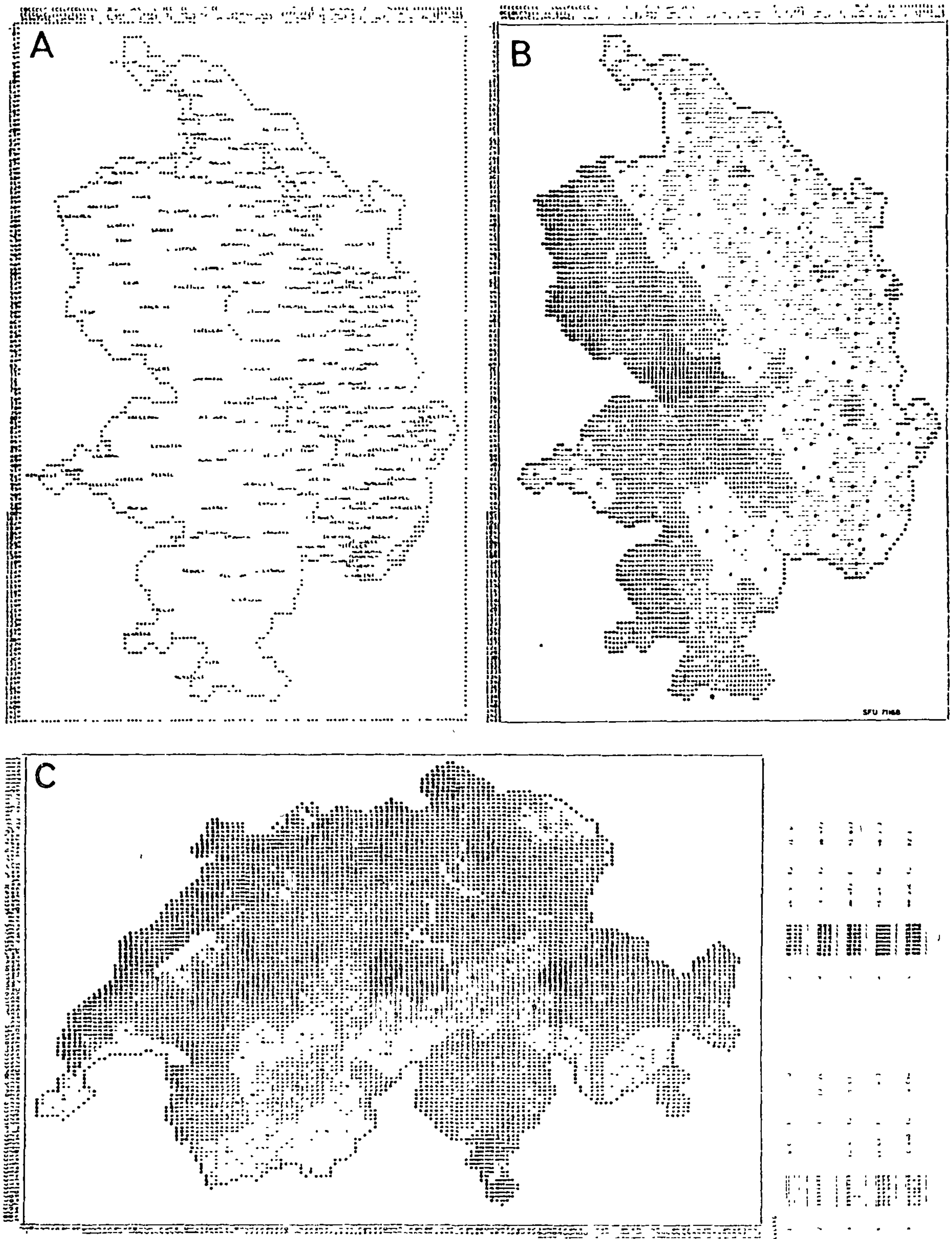


Figure 2 11 Switzerland–Three Maps by GEOMAP A Identification map, B unproductive and uncultivated land as percentage of total district area; C forest as percentage of total district area Since the program can print only a page width, the data have been rotated 90° to allow for a larger map Data D Steiner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

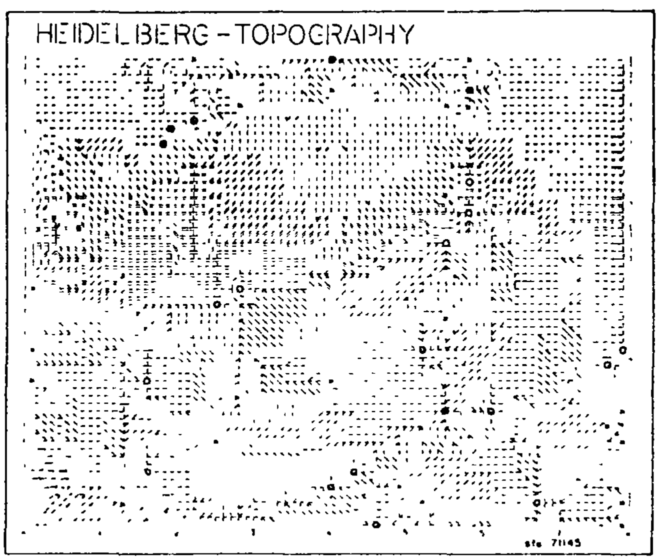


Figure 3.1 Heidelberg, Topography: A study of the surface of Heidelberg by a slope-analysis program. The specially marked points are peaks Program: WATERSHED; data: T.K. Peucker, production W D Rase

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

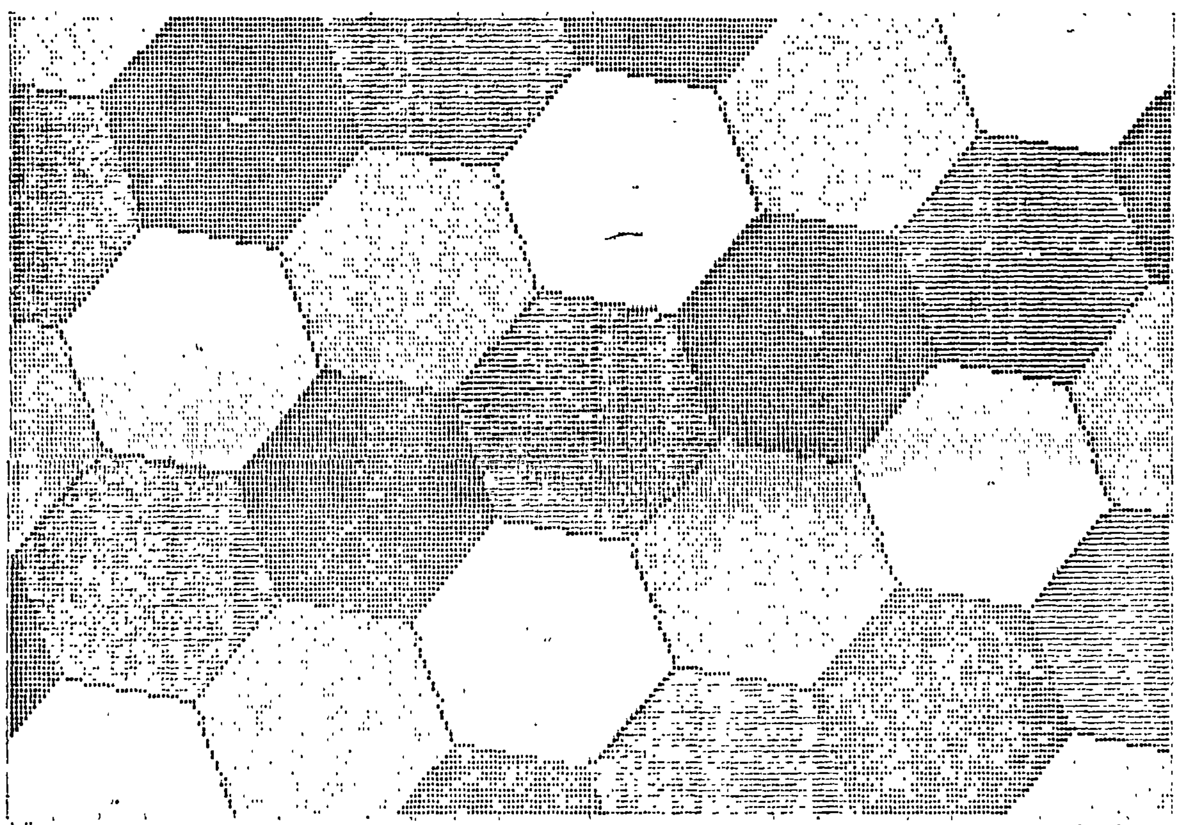


Figure 3 2 Hexagons. Produced by the proximal-map electives. The data-points are arranged in regular triangles Program SYMAP, data W.D Rase

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.

[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

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)

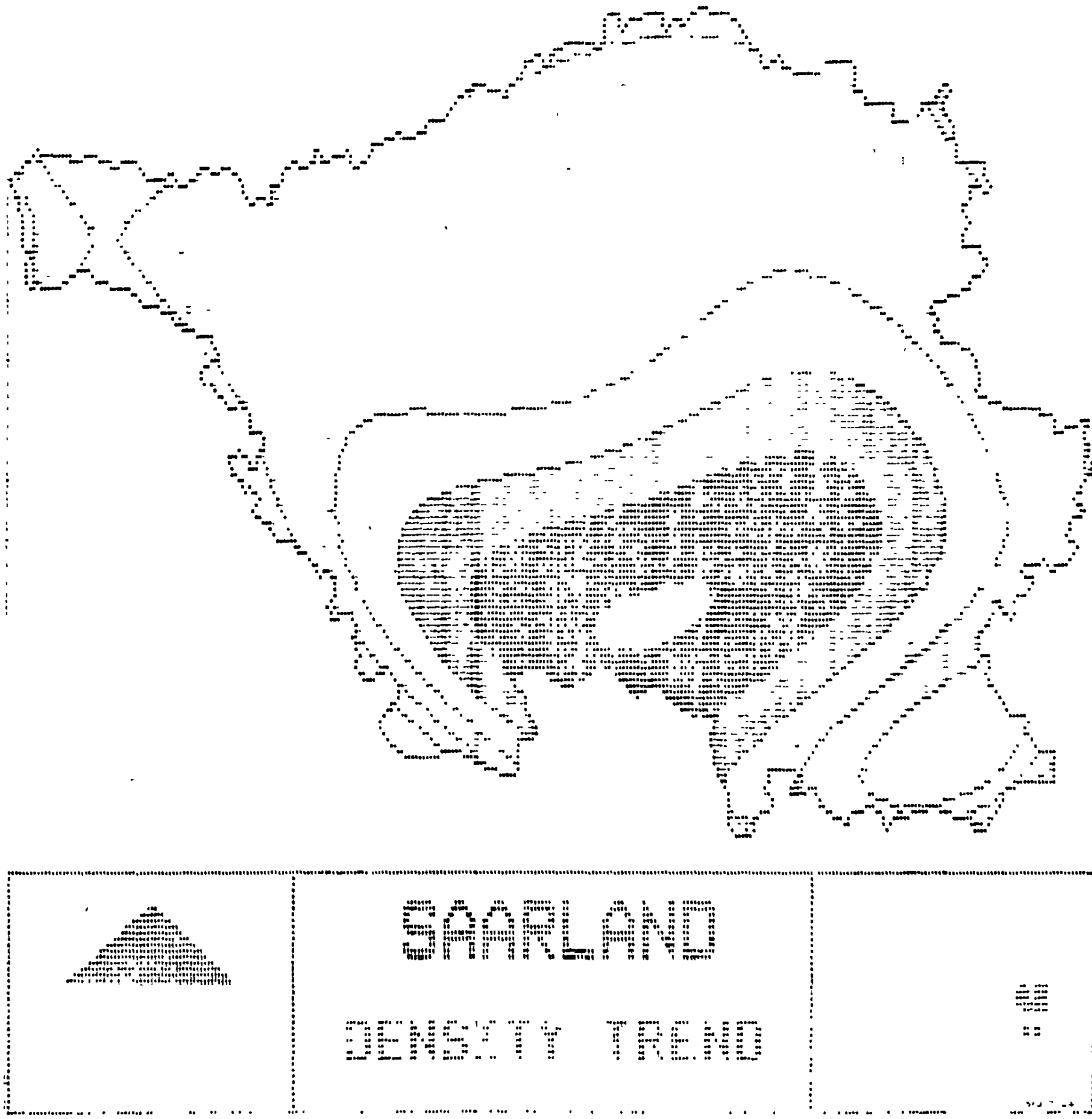


Figure 3.3 Population Density, Saarland A fourth-order trend surface
Program SYMAP with trend-option, data W D Rase

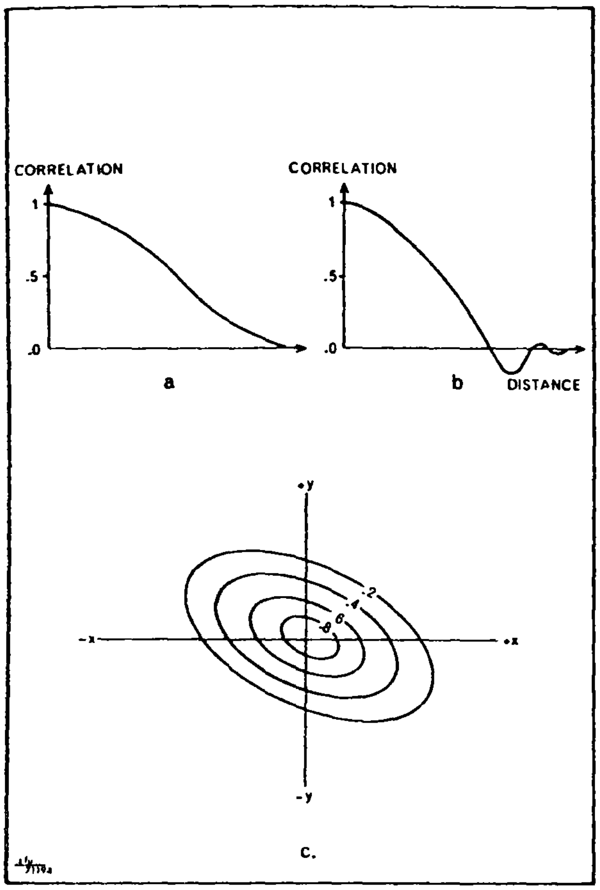


Diagram 3.1 Autocorrelation Function

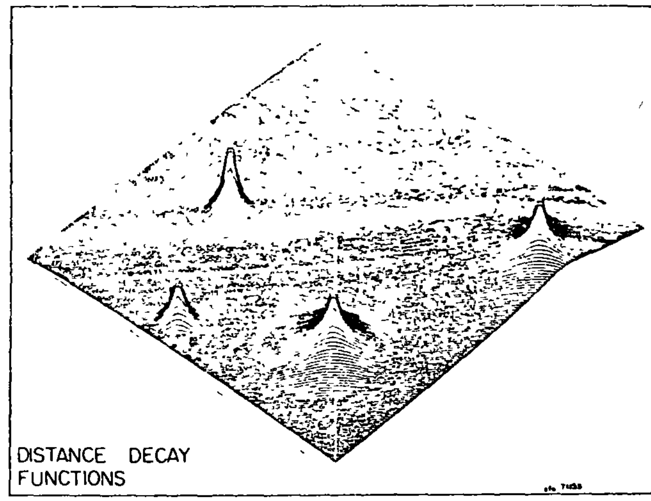


Figure 3.4 Distance Decay Function With Four Centers Programs: SPATFUN, SYMVU; production: T.K. Peucker

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.

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.

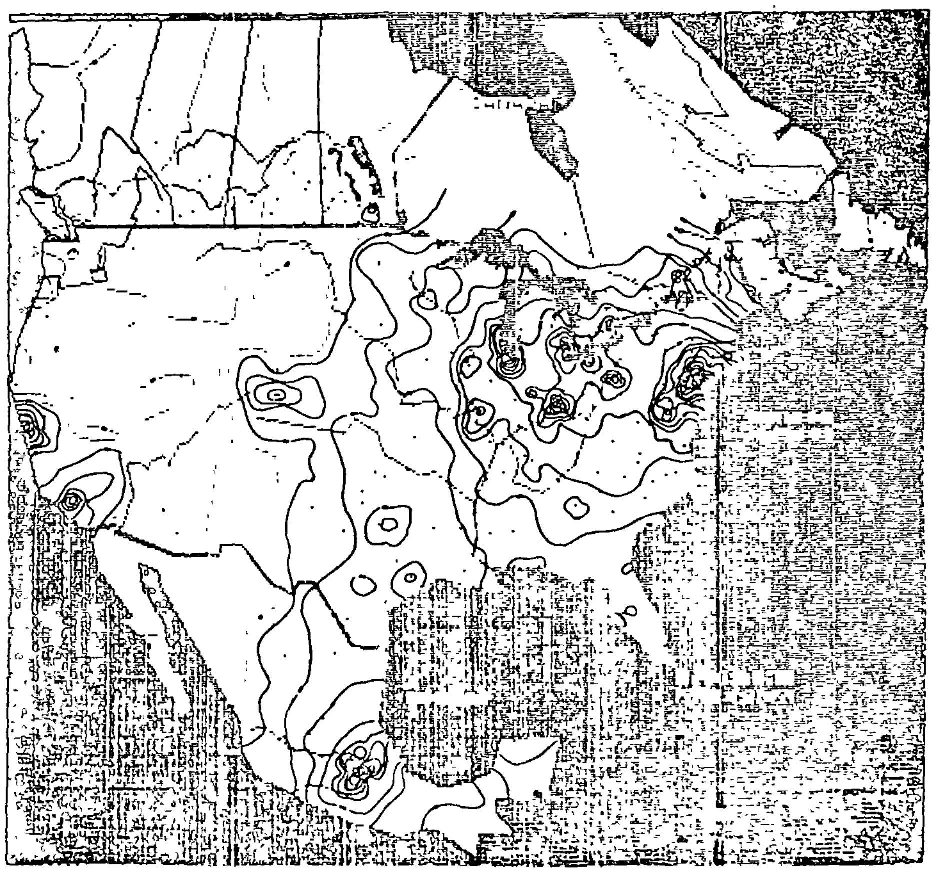


Figure 3.5 North America, Population Potential The concept of surface-specific points and lines can be exemplified by the population-potential map Source Douglas, David H , *Illustration of Regional Disparities in North America by the Use of the Gravity Model,* M.A Thesis, Carleton University, September 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 autocorrelation 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.

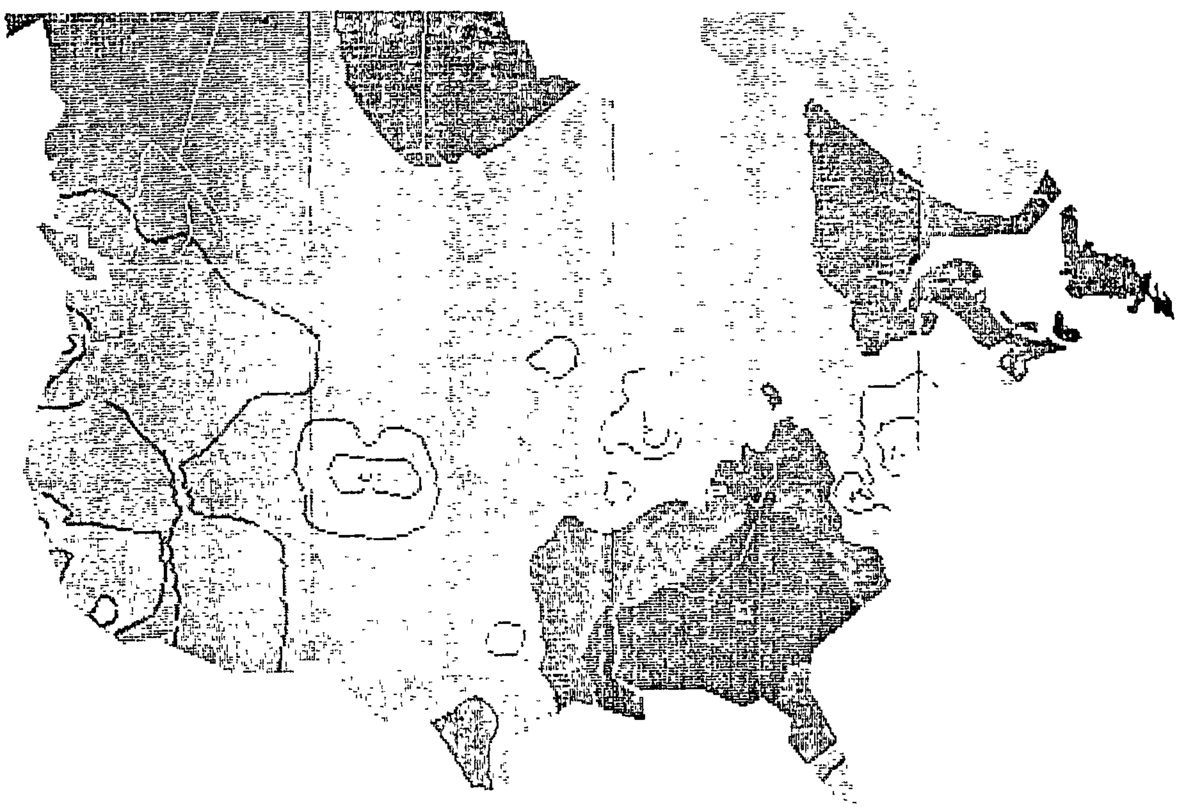


Figure 3.6 North America, Sales and Population Potential. The map shows the weighted difference of the population potential and the retail sales potential (Harris' market potential). The result is a map of purchasing power disparities (darkest area = negative). Source as in Figure 3.5

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.

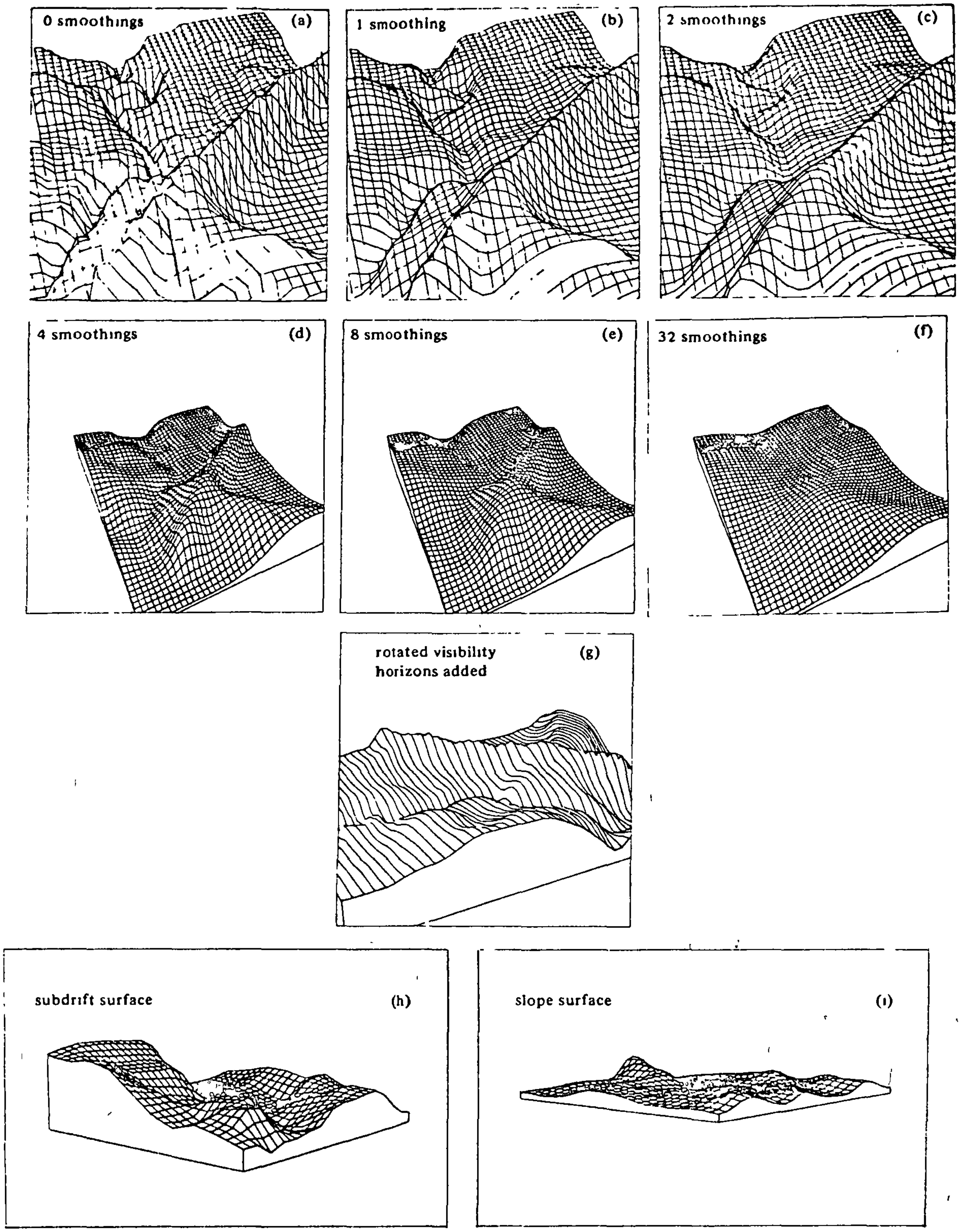


Figure 4 1 Surface Processing. Graphs a, b, c, show Leamy Creek in a close-up with 0, 1, and 2 smoothings, d, e, f show the same surface with 4, 8, and 32 smoothings. Leamy Creek is rotated in graph g and visibility horizons are added. Graphs h and i show a subdrift surface and its slope surface, respectively. 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).

[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.

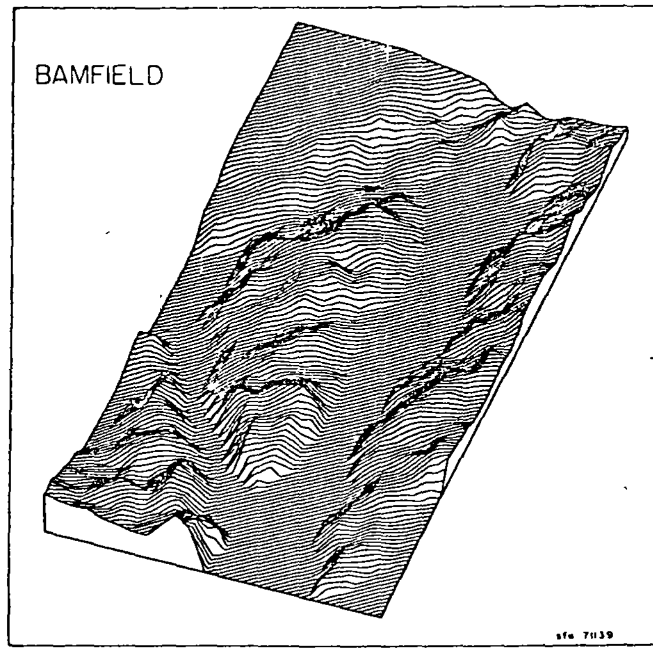


Figure 4.2 Bamfield, Linear Interpolation The elaborate SYMAP-interpolation algorithm is used only for every second point down and third point across as default. This example shows that this can create unpleasant patch-surfaces Program SYMAP, data: G. Neilly, McMaster University, production T K. Peucker

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

[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.

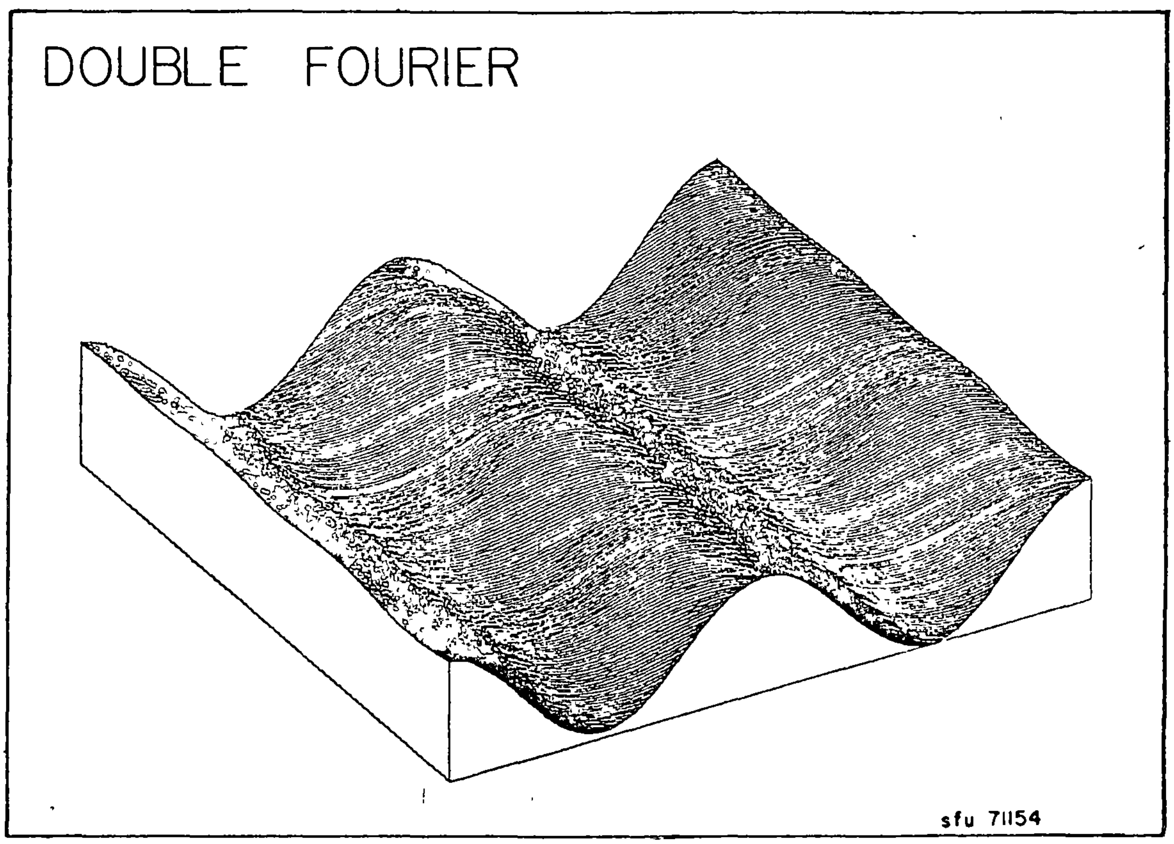


Figure 4.3 Double Fourier Series. Program SYMVU; production. W D. Rase

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 discriptions 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

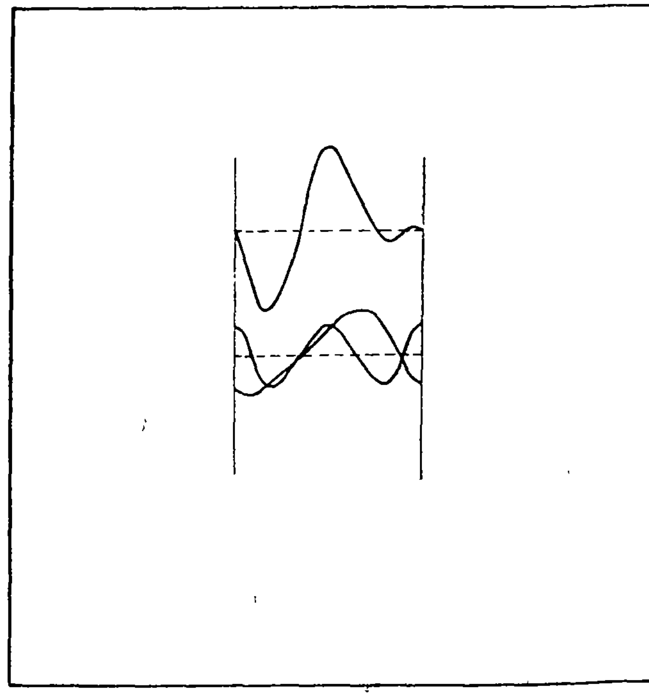


Diagram 4.1 Wave Components

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

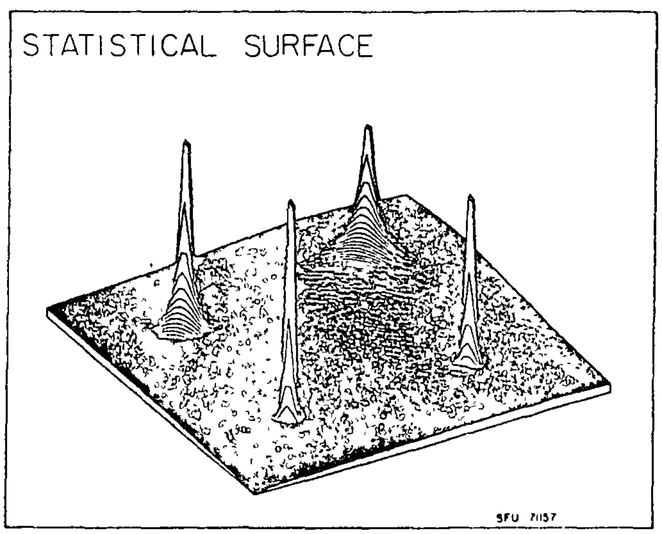


Figure 4.4 Statistical Surface, Distance Decay Function Programs SPATFUN, SYMVU; production T.K. Peucker

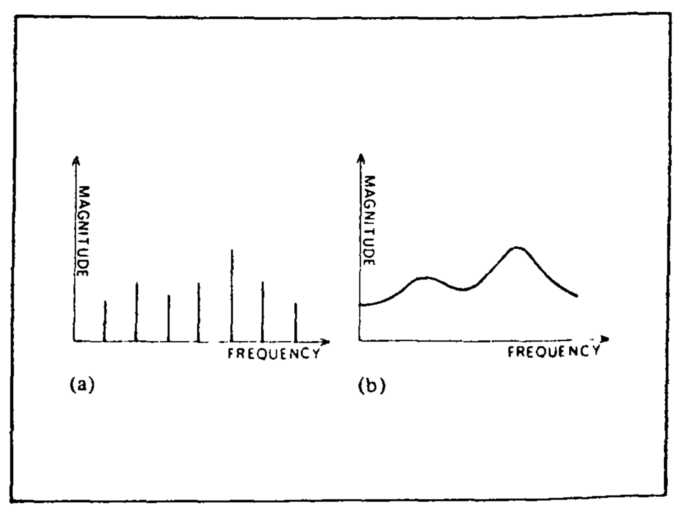


Diagram 4 2 One-Dimensional Spectrum

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

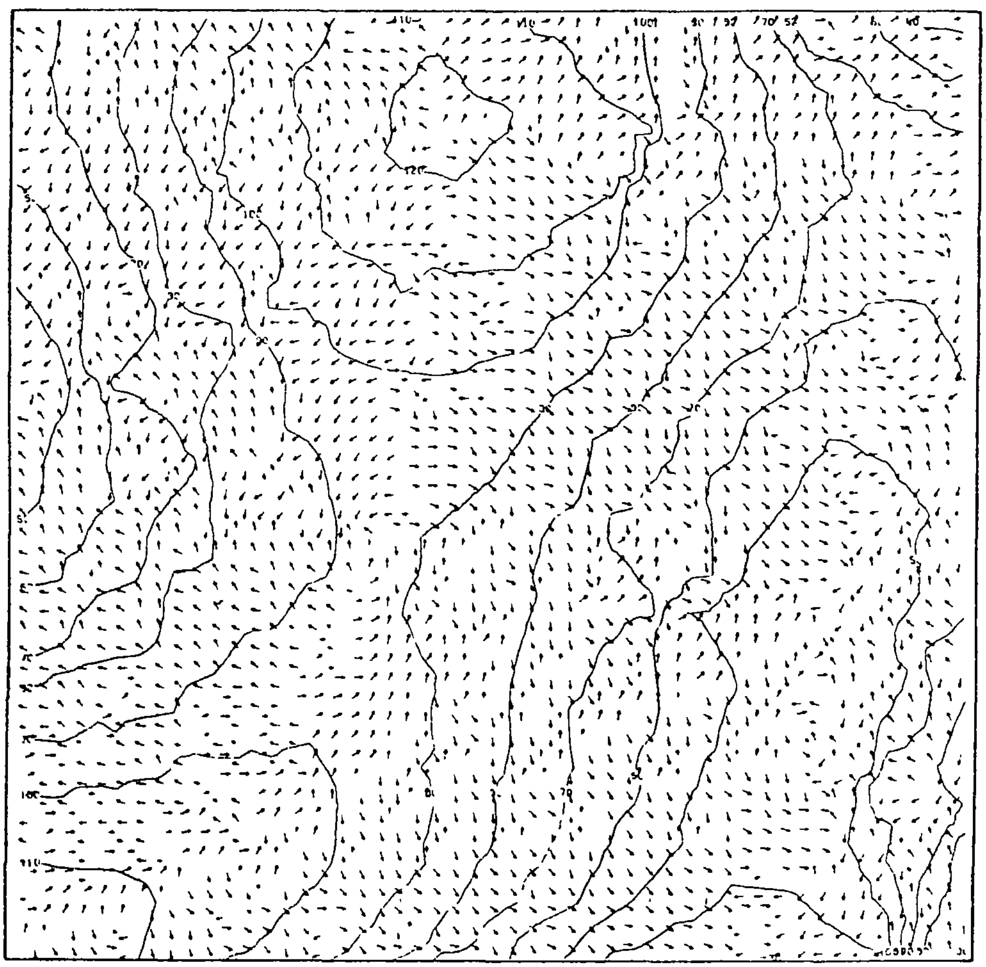


Figure 4.5 Grouse Mountain Slope. Program· SLOPE; data: D. Mark, production: W.D. Rase

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

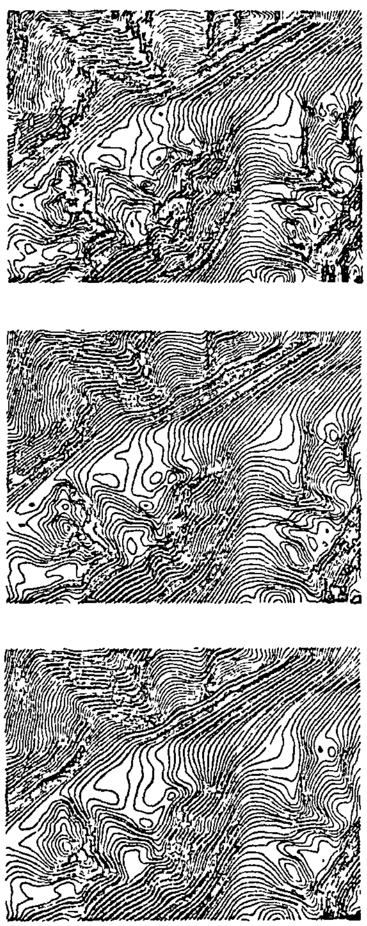


Figure 4.6 Lake Louise West, Smoothings This example shows that filtering has a scale-reduction effect. Program· INCLIN, data· T.K. Peucker

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.

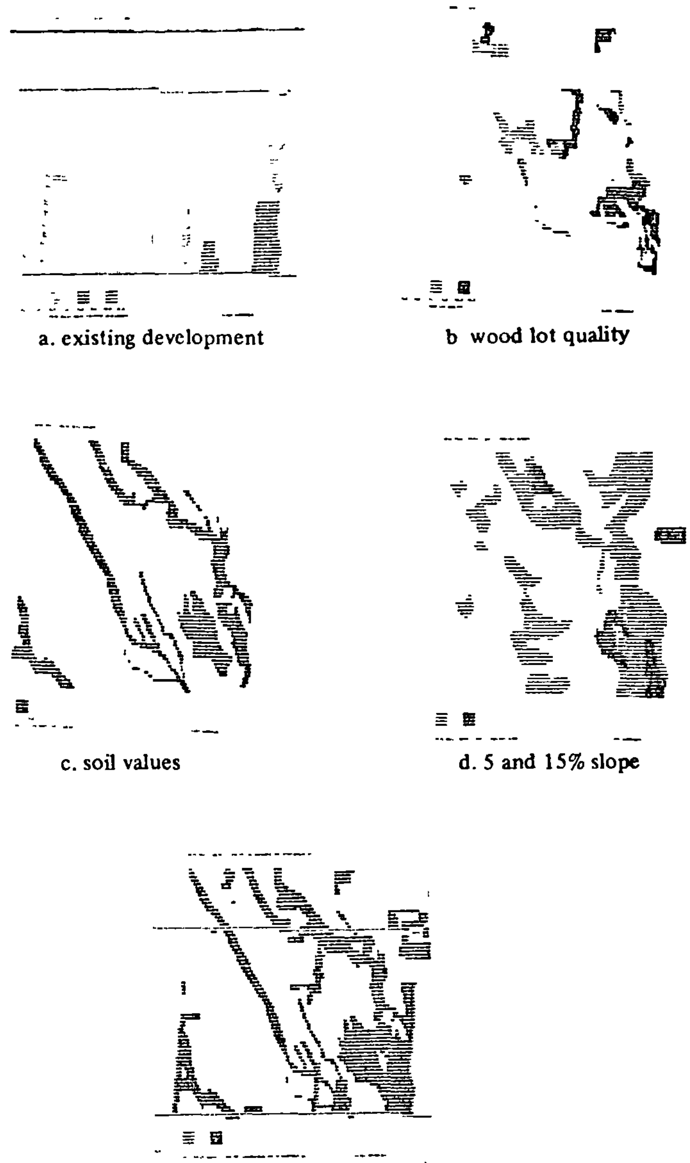
The other approach is to take advantage of the fact that even a conventional computer can perform limited logical and shifting operations on each binary digit of one memory "word" in a parallel fashion. In this case, a surface is represented by one or a stack of "bit-planes" (i.e., grid values may be 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)

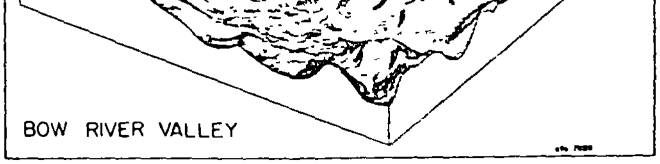


Figure 5.1 Bow River Valley, View from N.W. Shows very clearly the glacial features of the area Visual effect given by 298 lines running parallel over the surface. Program SYMVU; data. G Neilly, McMaster University

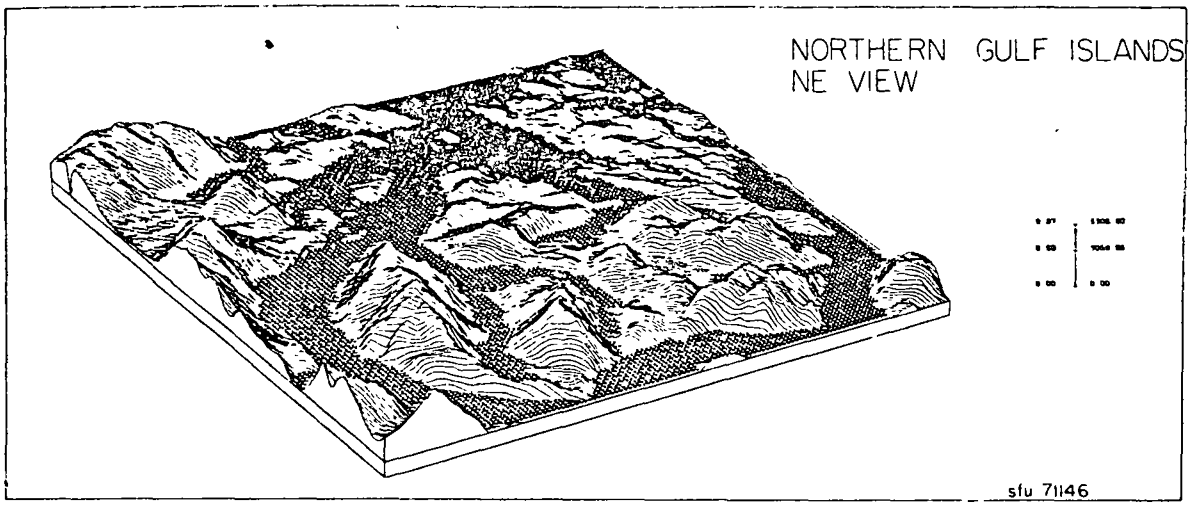


Figure 5.2 Northern Gulf Islands, Water Given Background Symbol. Program: SYMVU; data: W. Wolferstan; production: W.D. Rase

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," *I.B.M 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.

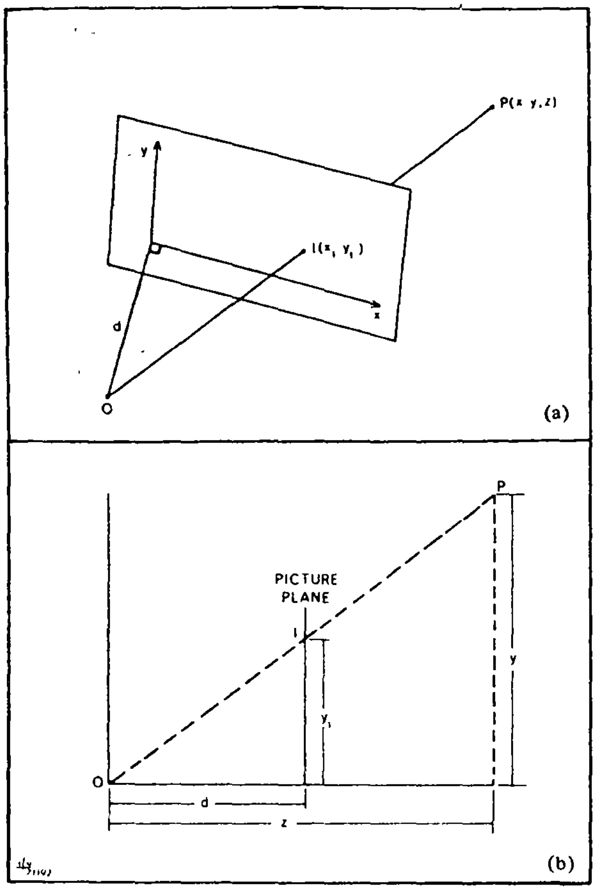


Diagram 5 1 Picture Plane

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

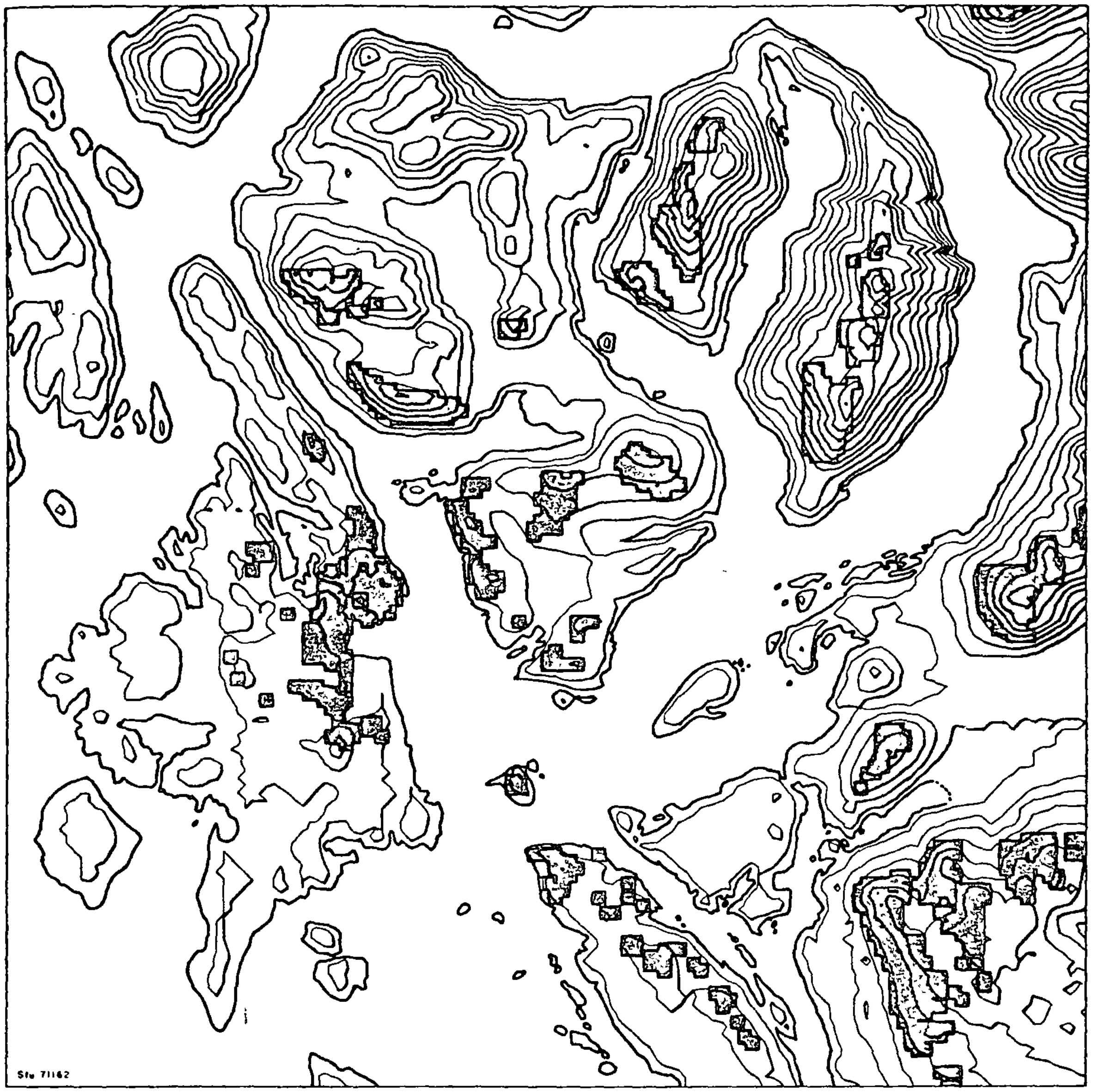


Figure 5.3 Northern Gulf Islands, Visible Area From One Point. Programs: VIEW, SHDCTR, data: W. Wolferstan; production: W.D. Rase and G. Brady

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.

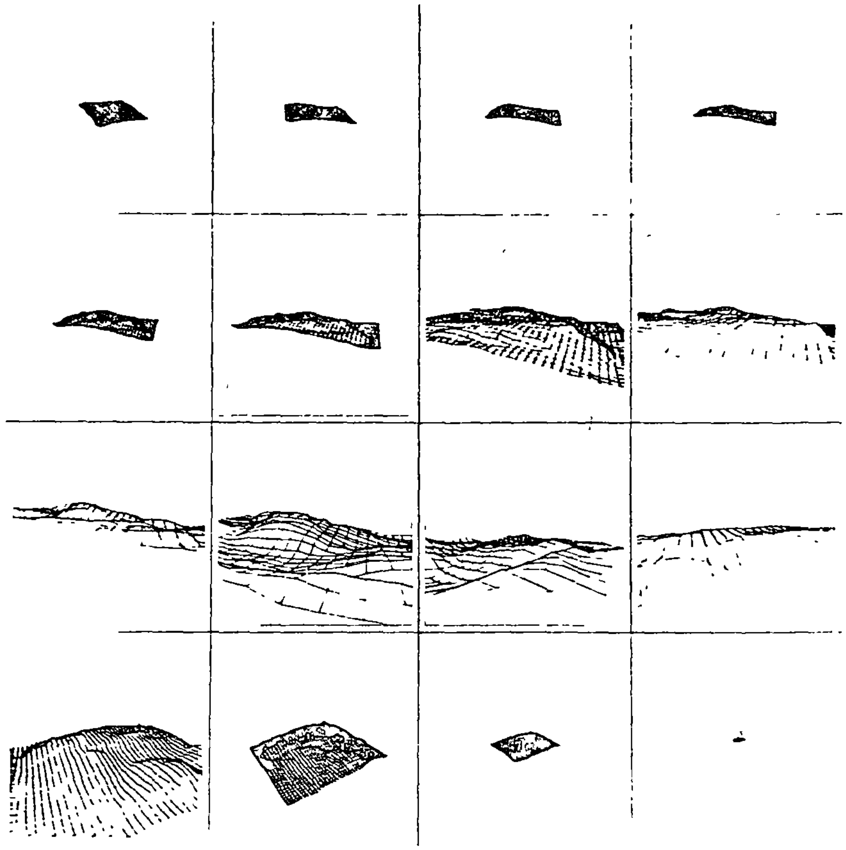


Figure 5.4 Camp Fortune, Views Seen by Approaching and Passing Aircraft The turn to look back is at the 12th graph. VIEWBLOK computes the drawing of one set of parallel lines and flags the points (1-visible, 0-not visible) for the perpendicular lines or other features Program· VIEWBLOK, data D Douglas

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.

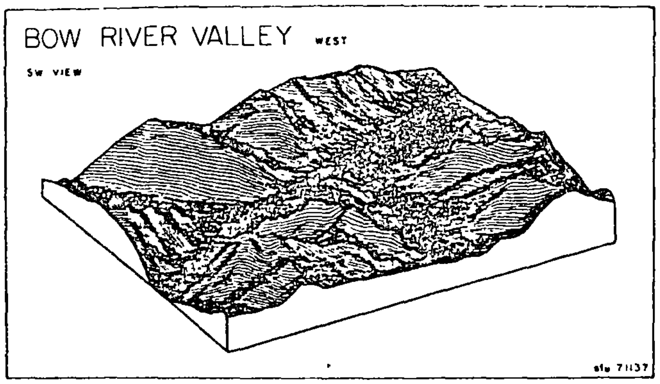


Figure 5.5 Bow River Valley, West. Shows the amount of area which can be hidden in a perspective plot. Program: SYMVU, data: T.K. Peucker, production: W.D. Rase

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.

[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.

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.

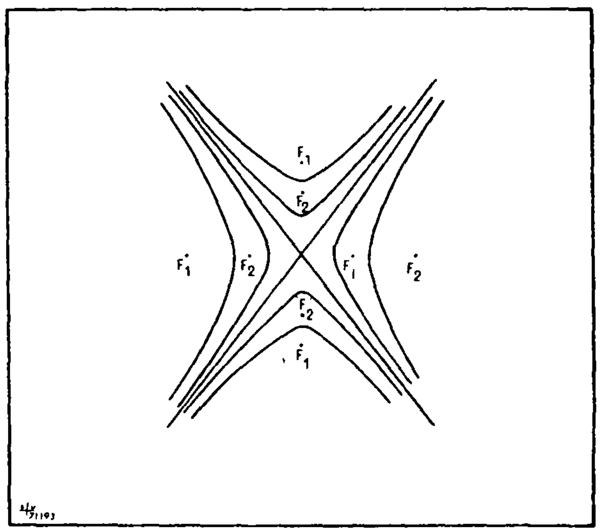


Diagram 5.2 Contour on a Saddle-Point

[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 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.

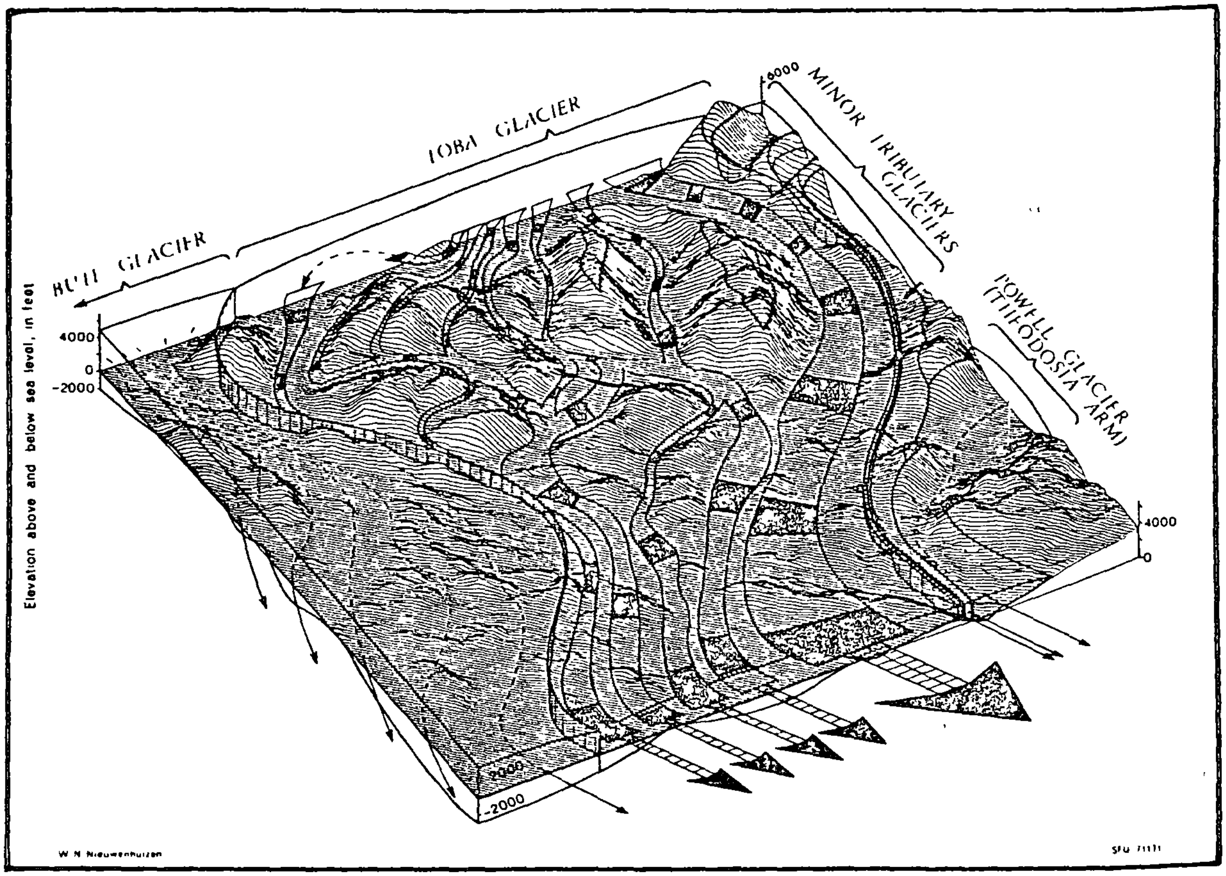


Figure 5.6 Northern Gulf Islands, Ice-Flow Map. A computer-drawn view used as the basis for the analysis and description of ice-flow patterns. Analysis and ice flow W. Nieuwenhuizen, Simon Fraser University; program· SYMVU; data W. Wolferstan

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

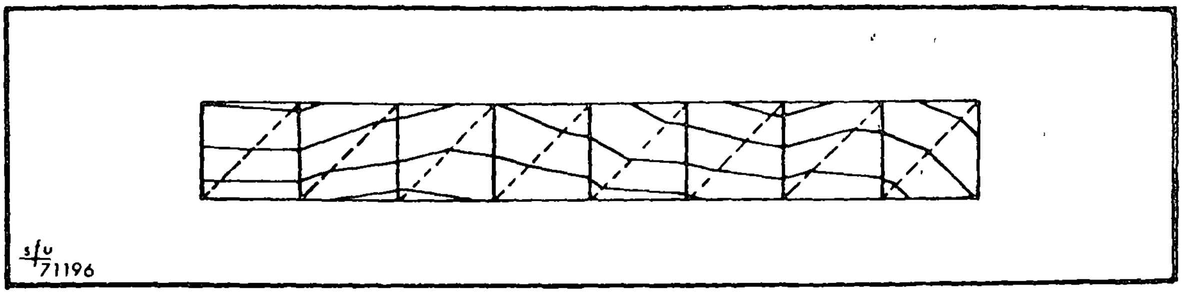


Diagram 5.3 Contour Scanning

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)

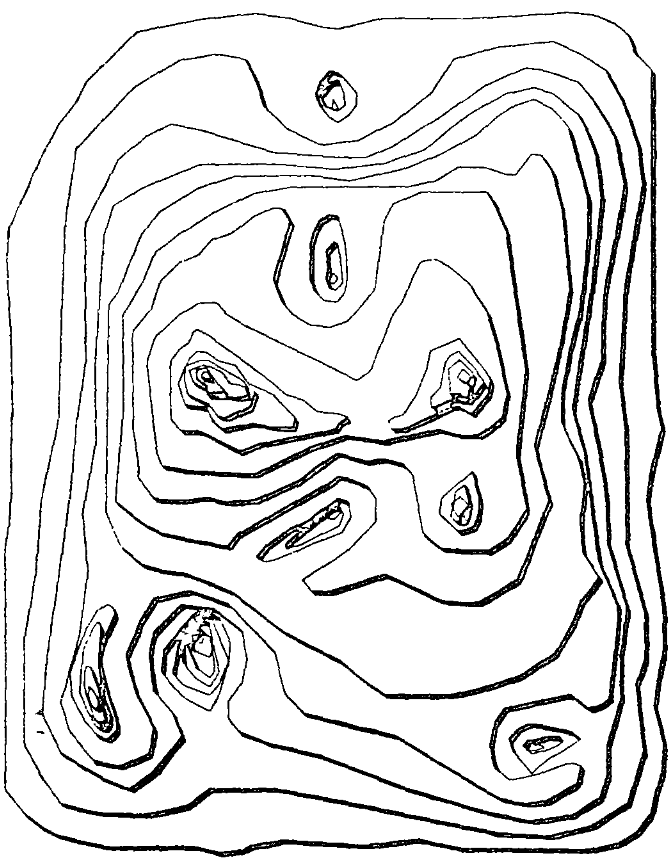


Figure 5.7 Theoretical Surface, Relief Contour Method Program: SHDCTR; data: D. Wolfe, Simon Fraser University; production. T.K. Peucker

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.

---

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

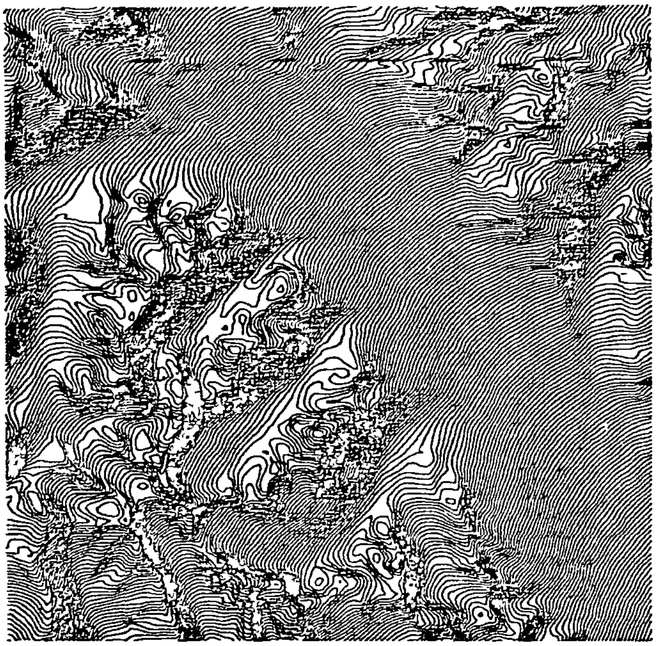
$$Z^r_{ij} = A + B\,(Z^s_{ij} - A) \text{ where}$$

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

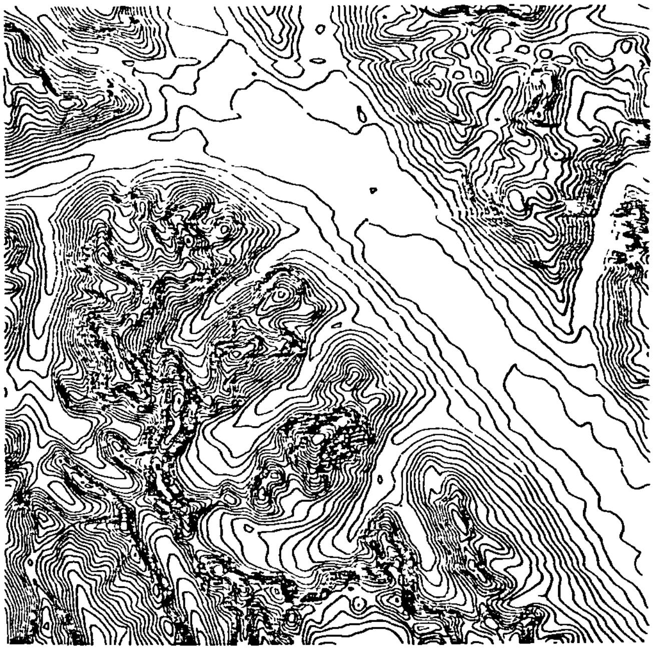
Z^s_{ij} = the smoothed value of the point ij

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, YOELI

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

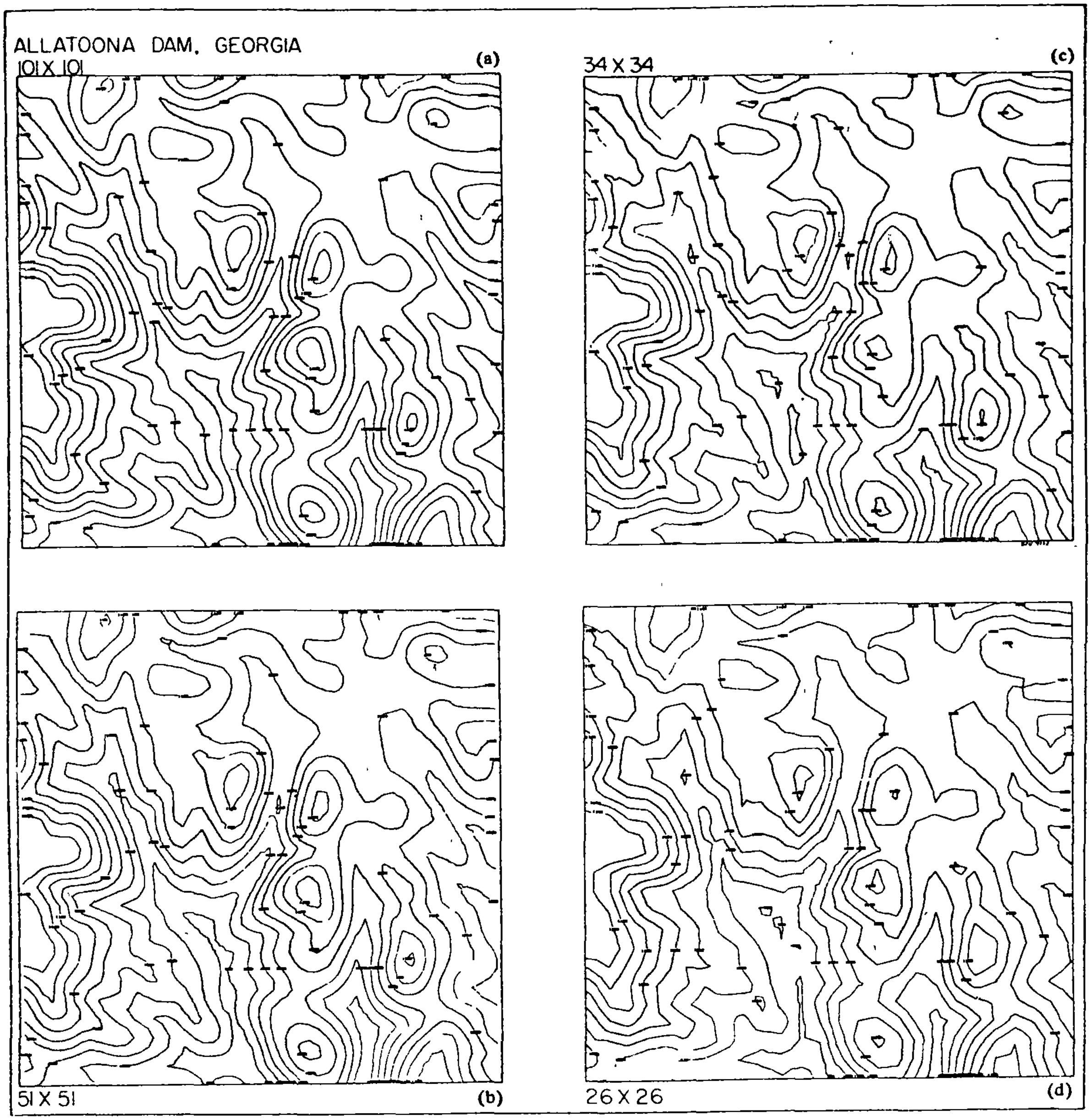


Figure 5.9 Allatoona Dam. Contouring With Varying Grid Sizes. Program: KOPPE, CNTOUR; data courtesy D. Edson, U.S. Geological Survey; production: T.K. Peucker

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

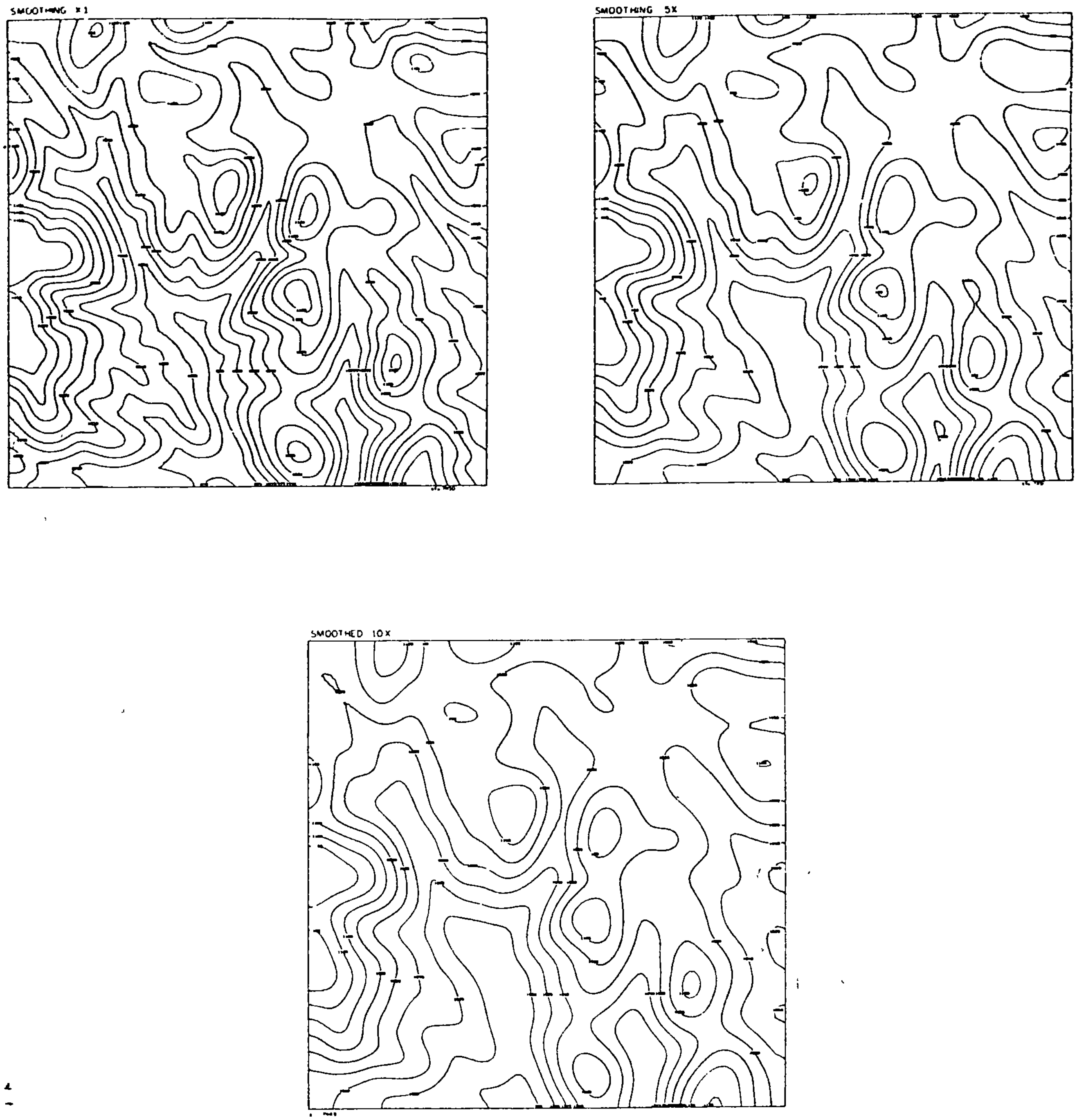


Figure 5.10 Allatoona Dam, Contouring With Smoothings Programs SMOOTH, CNTOUR; data courtesy D. Edson, U.S. Geological Survey; production T.K Peucker

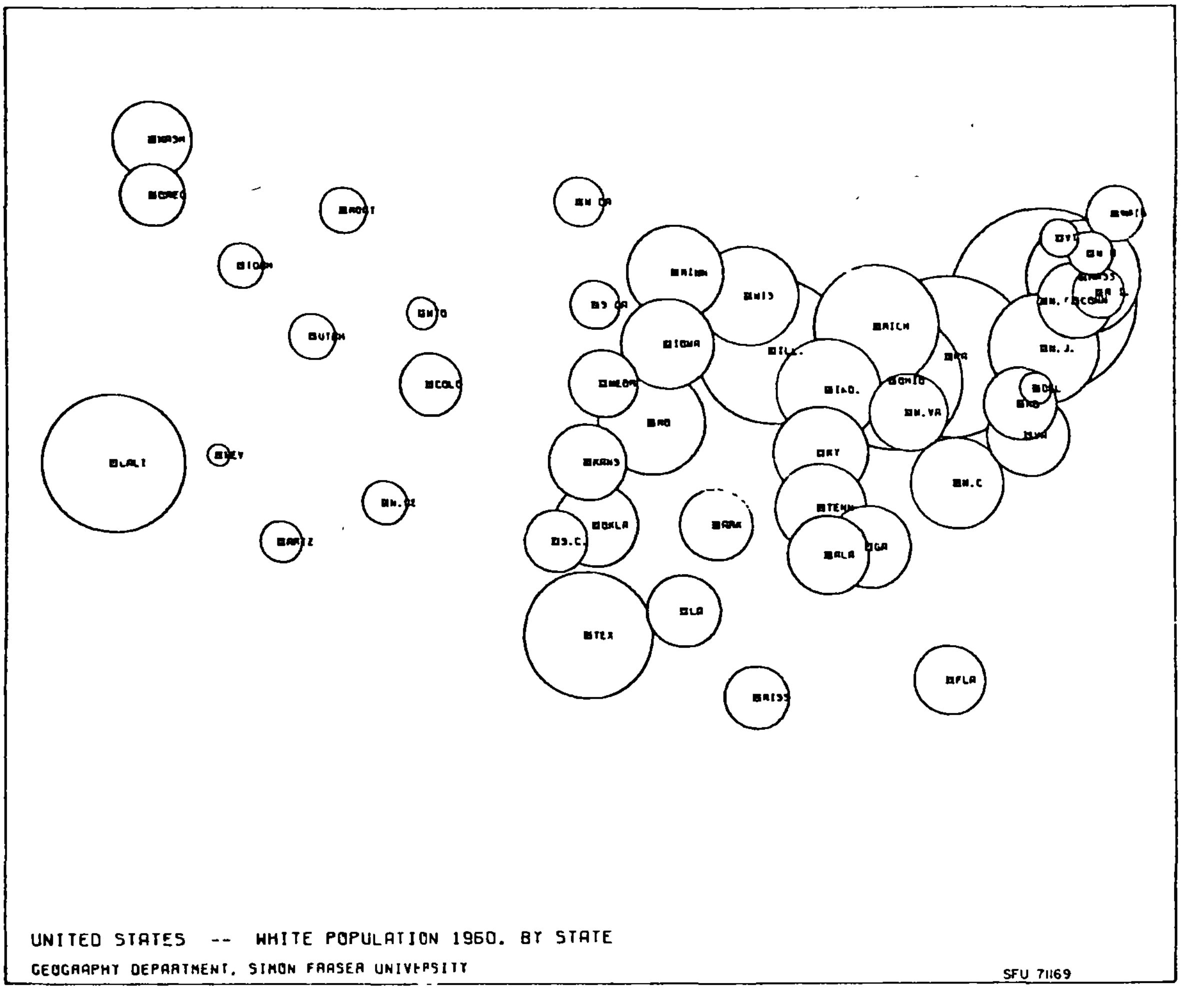


Figure 5.11 Display of Absolute Data Overlapping circles with removal of overlapping parts. Program· SIRKEL; data: D Hatlelid.

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, J H. and R L Gard (1968), "Computer-Assisted Cartography, A Graphic System for Chart Composition and Revision," *Information Systems Symposium*, Sept 4–6, 1968, Washington, D C.

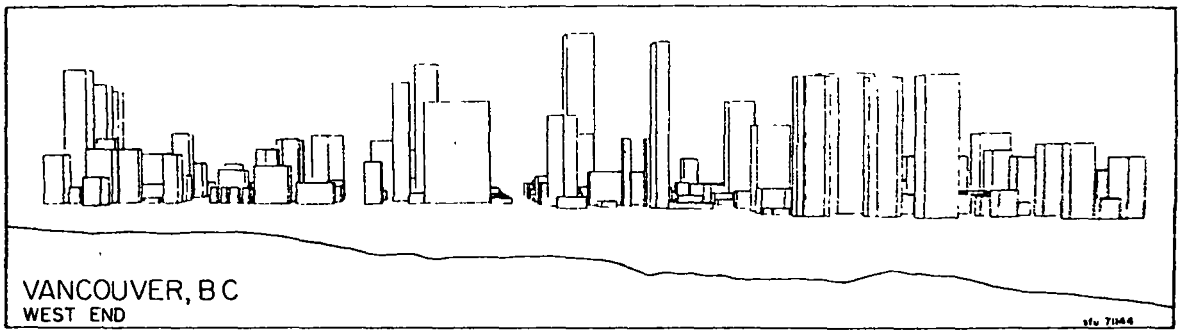


Figure 5.12 West End, Vancouver, Advantages of Hidden Line Removal. Program: BLOCKS; data· A. Furguson

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-databank. 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.[63]

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

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

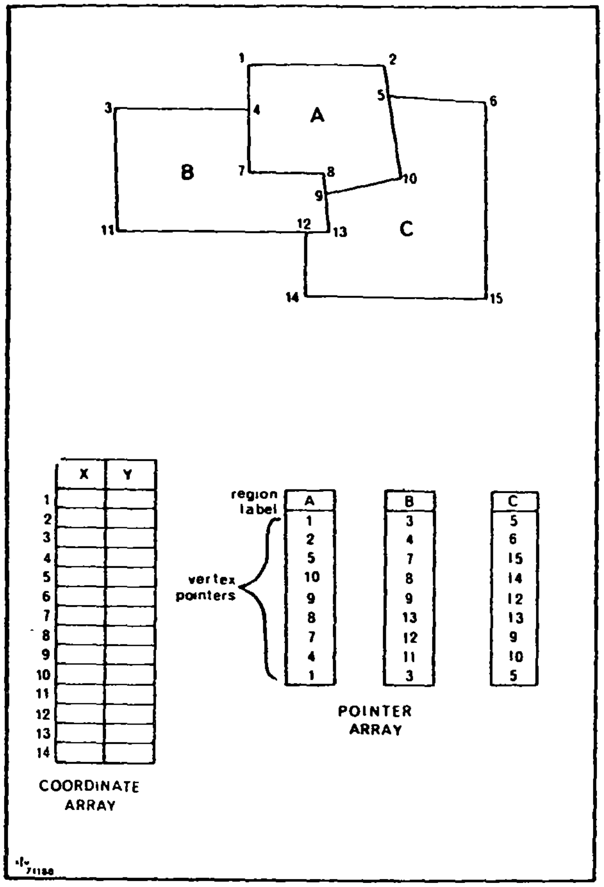


Diagram 6.1 Boundary Encoding

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 Deecker, 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).

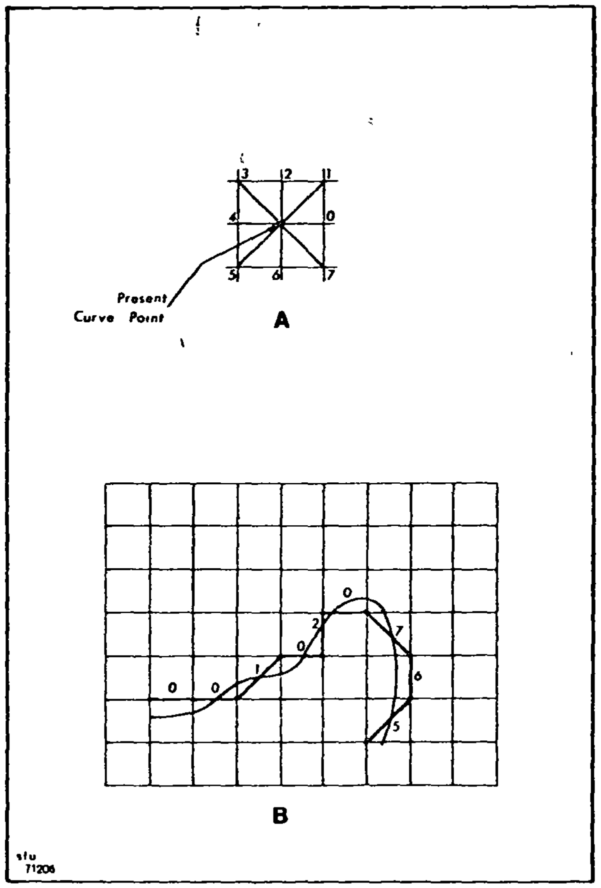


Diagram 6.2 Chain Encoding

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,

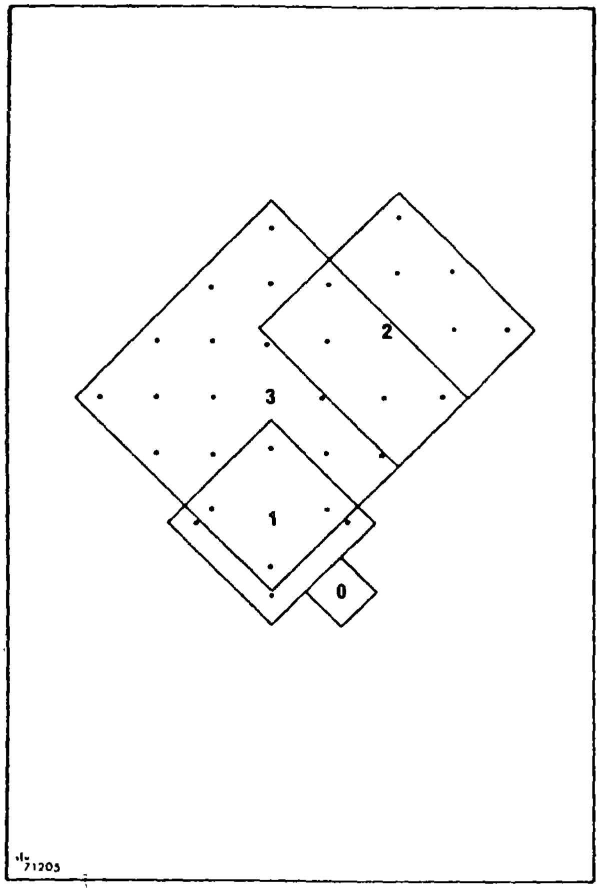


Diagram 6 3 Skeleton Encoding

[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

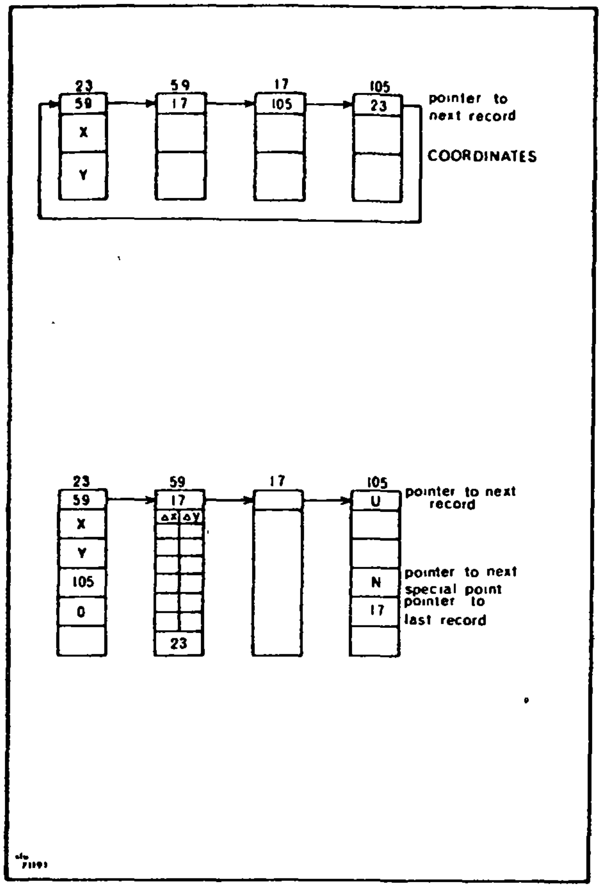


Diagram 6.4 List Structure of Line Encoding

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

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

[73] In Tomlinson, 1970

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

[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

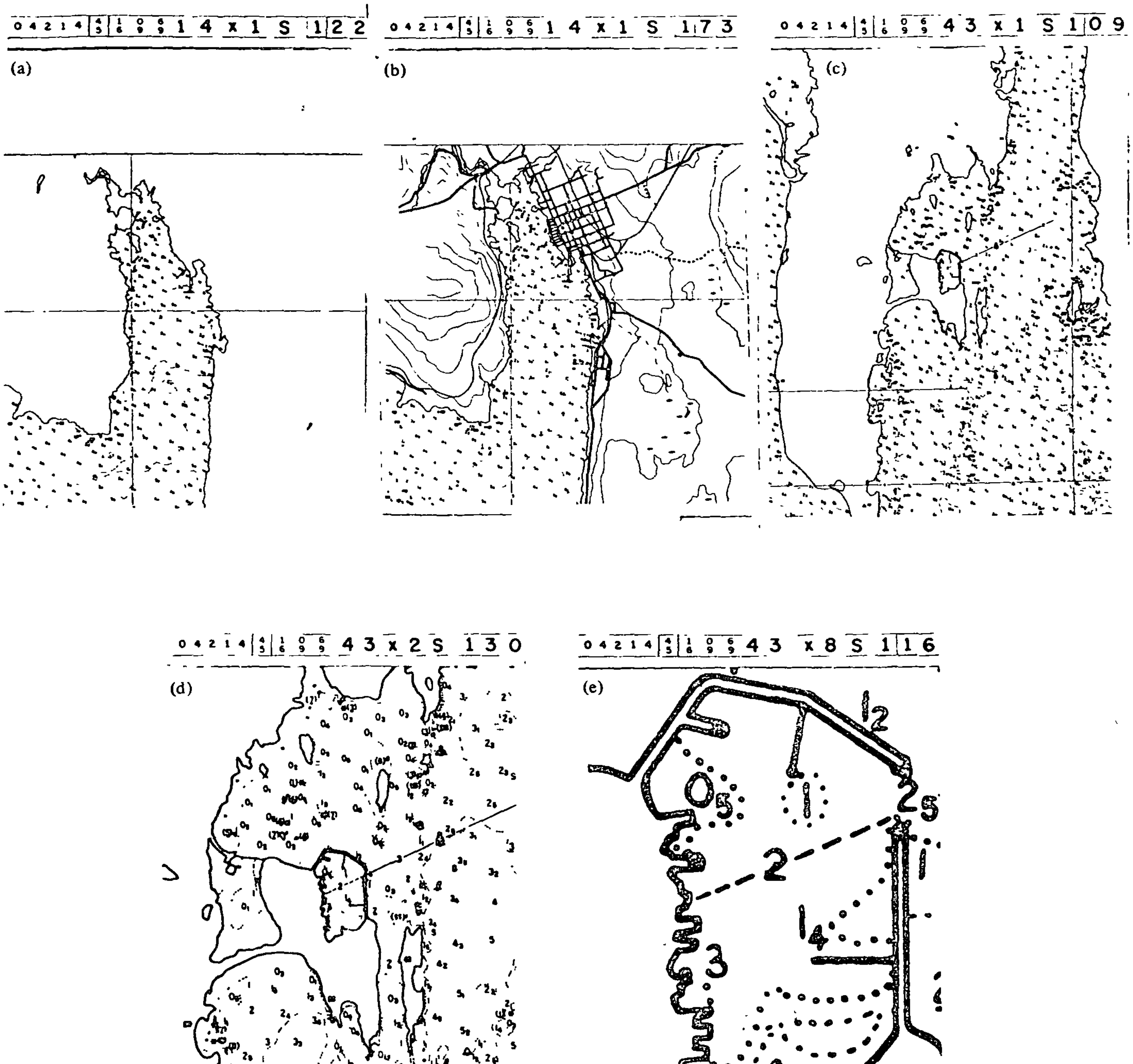


Figure 7.1 Computer Aided Compilation Operator can change lines, add data sets (b), zoom the picture to see a section in more detail, (c, d, e,), etc Courtesy A.R Boyle, University of Saskatchewan

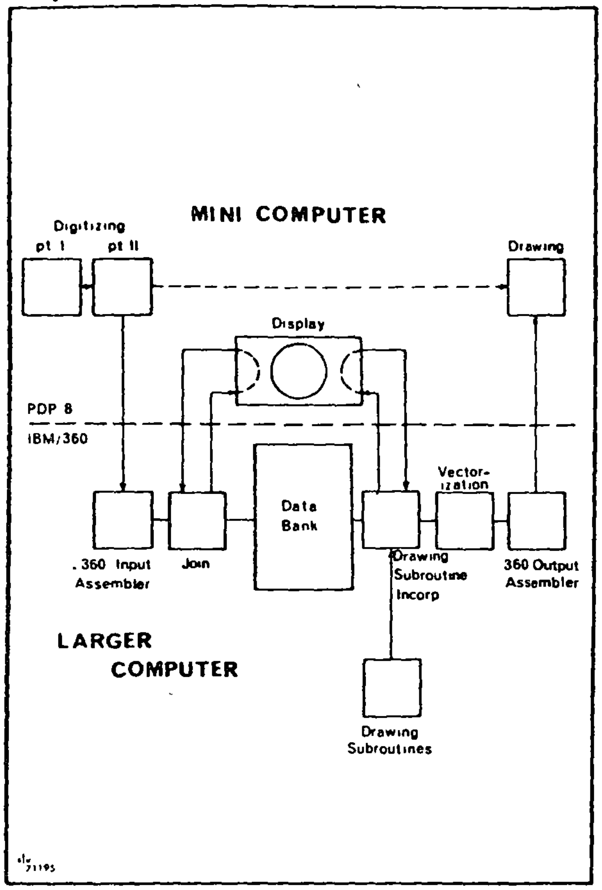


Diagram 7 1 The Automated Cartography System for the Hydrographic Service of Canada

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

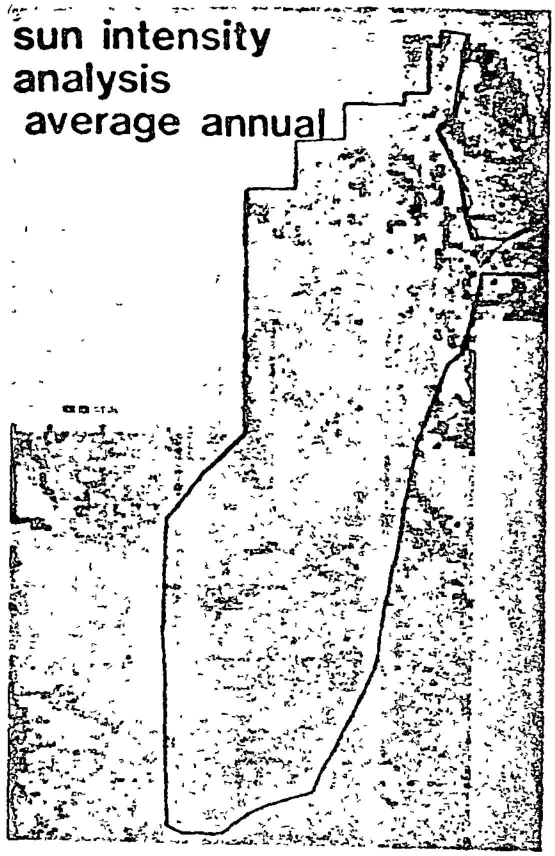


Figure 7.2 Northstar Development, Sun Intensity Analysis
Courtesy J. Dangermond, Redlands, California

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, T 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

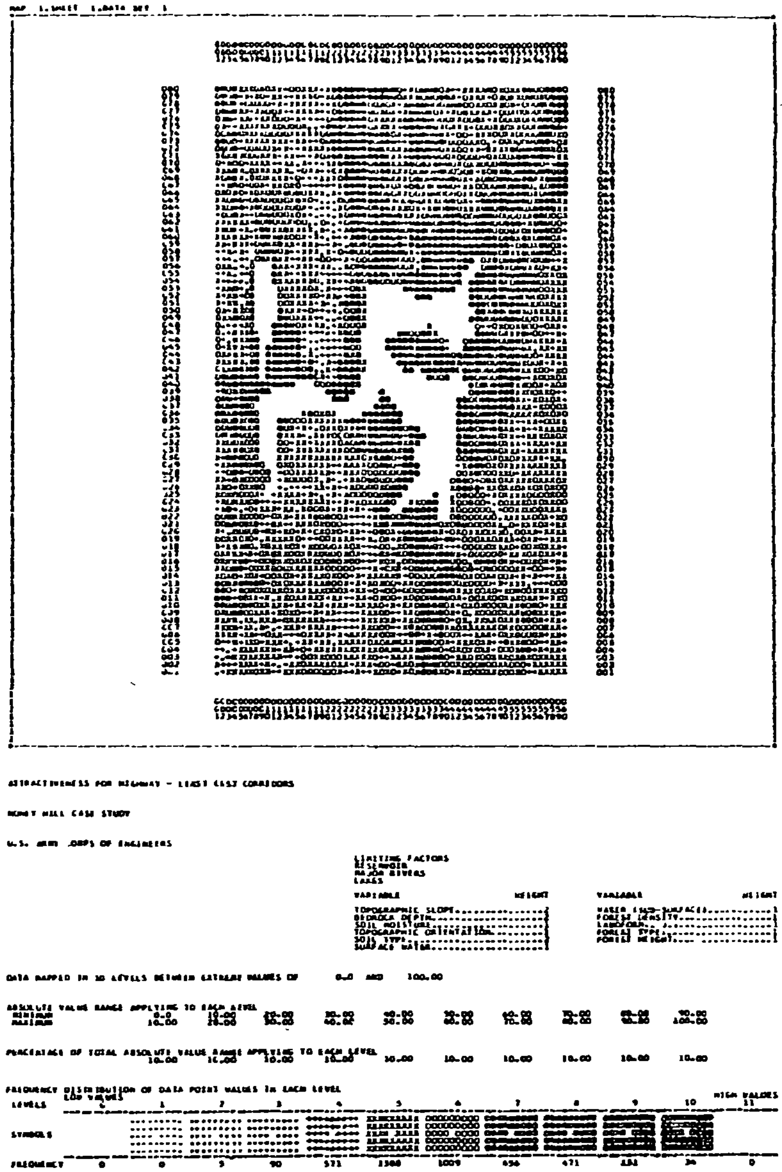


Figure 7.4 Honey Hill, Attractiveness for Highway. Courtesy C Steinitz, Harvard University

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

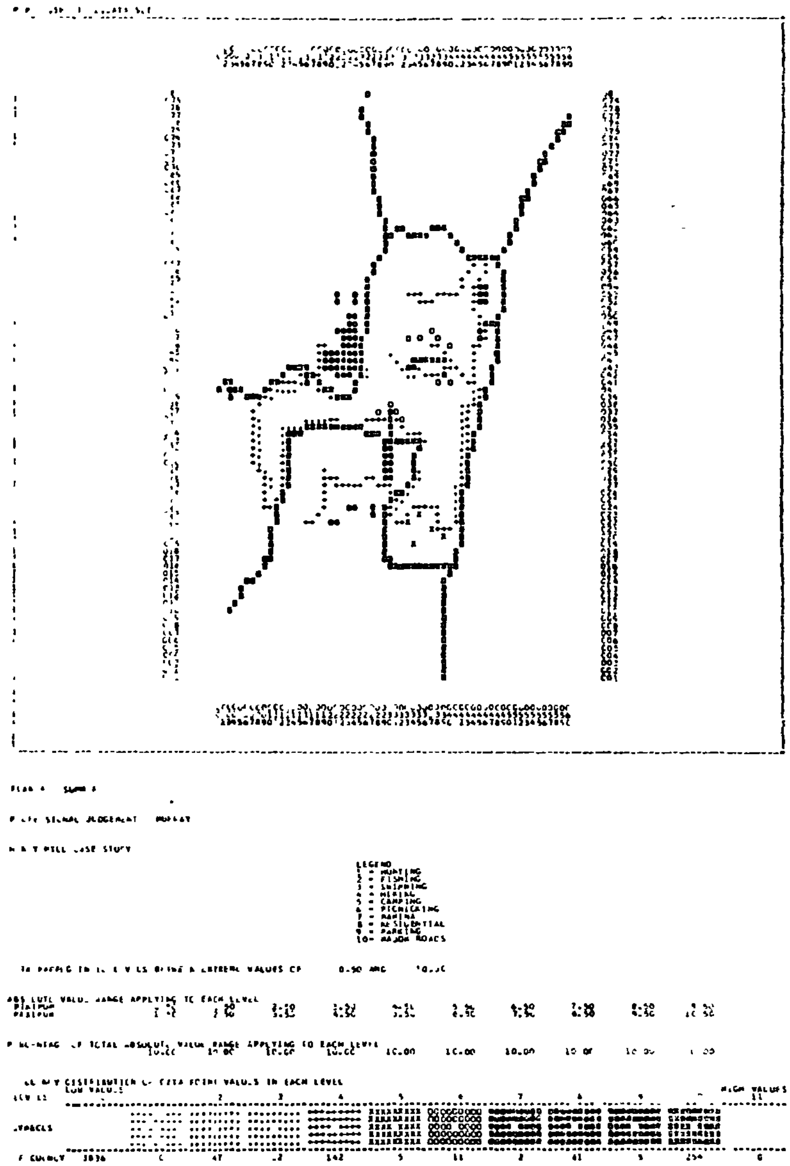


Figure 7 5 Honey Hill, Professional Judgment, Murray Courtesy C. Steinitz, Harvard University

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.

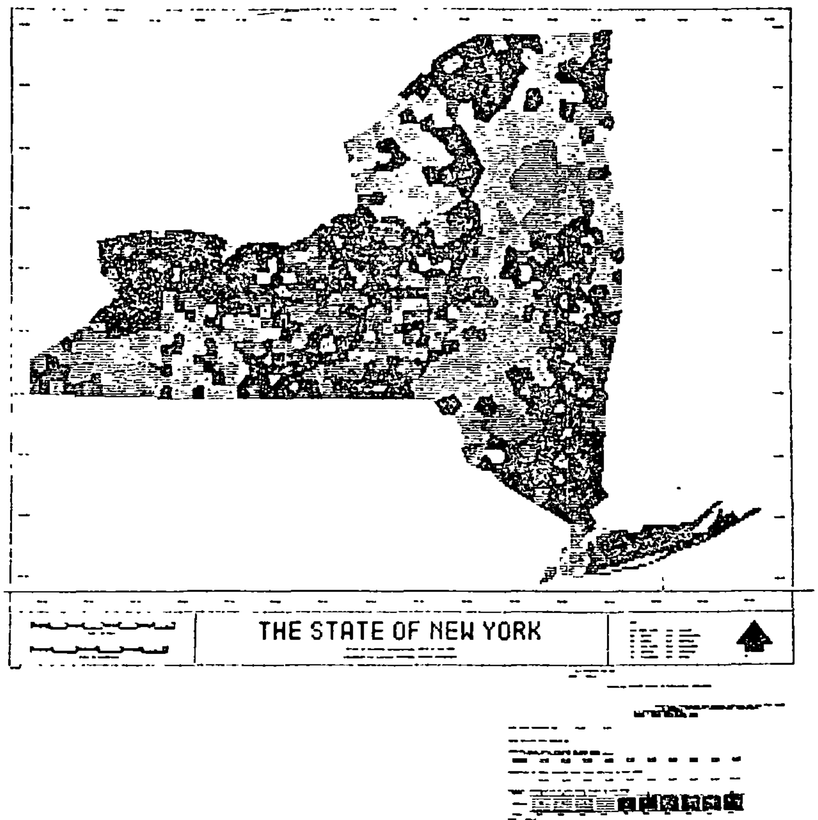


Figure 7 6 New York, Predicted Population Change 1960-1990. Courtesy C Steinitz

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.

Hägerstrand, 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 av 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 $7\frac{3}{8}$ by $3\frac{1}{4}$-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. Hatlelid, 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

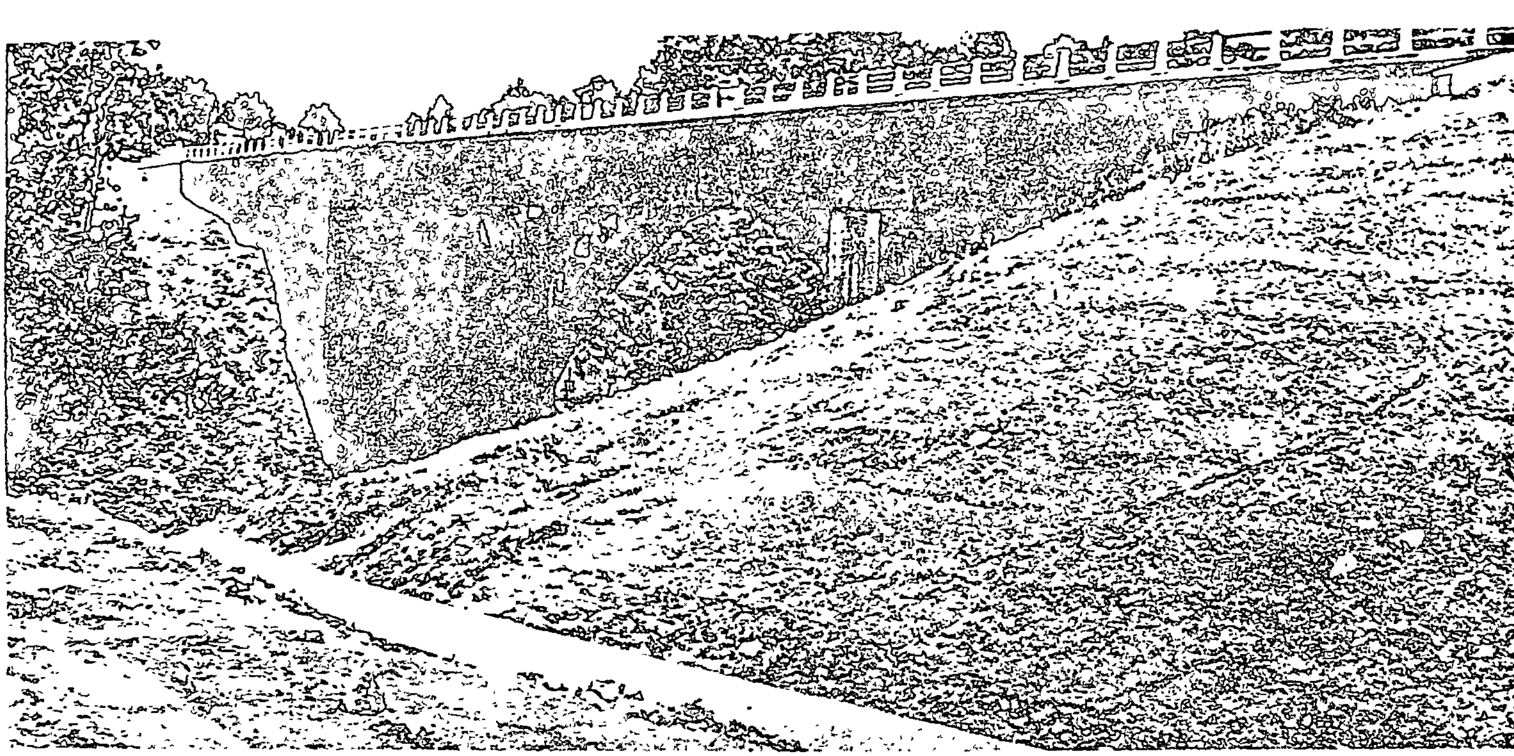
SYMVU – Plots perspective views of regular grids Authors F. Rens, State University of New York, Buffalo and Laboratory for Computer Graphics and Spatial Analysis, Harvard University.

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



Falsework eliminated in bridge erection

The U S. Highway 50 bridge near North Vernon, Ind., is one of the first applications in this country of the precast, prestressed concrete box girder design—erected without falsework This project points up engineering skills and progress so well that it competed for the Outstanding Civil Engineering Achievement Award last spring.

GENE K. HALLOCK, M. ASCE
Chief Engineer
Indiana State Highway Commission
Indianapolis, Ind

THIS STRUCTURE REPRESENTS a pioneering application in this country of an innovative approach to the construction of economical, long span bridges that makes efficient use of materials and eliminates costly falsework The successful application of this new method requires skillful craftsmanship, precise quality control of materials, use of new adhesive materials, and above all, full use of current engineering technology to design the segmental forms, calculate erection and service stresses, and efficiently plan and carry out casting and erection.

This type of construction, pioneered in Europe and now being used sparingly in the United States, is a testimonial to civil engineering genius. It is an innovation in construction and is very pleasing aesthetically Due to the simplification of construction on the site, environmental problems, such as erosion, and visual eyesores, such as forms, were reduced considerably.

The new structure, designed by bridge engineers of the Indiana State Highway Commission, carried one-lane, two-way traffic while an existing adjacent bridge was being reconstructed While work on the segmental bridge was going on, US 50 traffic was maintained on the existing structure for alternating one-way traffic. The existing open-spandrel arch bridge was reconstructed with precast, post-tensioned members to meet budgetary and aesthetic limitations

The combined roadway width of the two adjacent bridges is 44 ft (13.4 m) The total length of the bridge is 390.7 ft (119.1 m) Combined with the roadway reconstruction, the total resurfacing will stretch 1,550 ft (472 m)

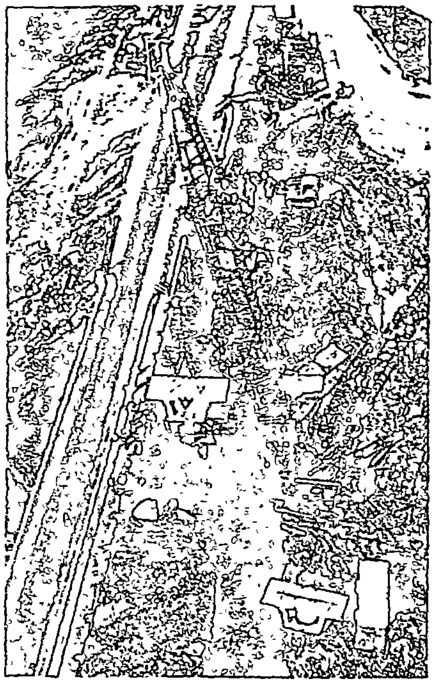
The US 50 bridge, located about 0.7 mi (1.1 km) east of North Vernon, was fabricated under a contract with Construction Products Corporation, Lafayette, Ind The segments, which weighed about 30 tons (27 Mg) each, were cast against adjacent segments in order to insure perfect field fit Approximately 123,000 lb (56 Mg) of reinforcing steel are included in the new structure

The selection of the segmental alternate was based on an economic study by Highway Commission bridge engineers. The cost of constructing a twin to the existing open-spandrel arch structure was prohibitive. The fabrication contract awarded in March, 1974, was for $230,000.

In September, 1974, the Highway Commission awarded an erection contract for around $1,220,000 to Tousley Bixler Construction Company, Indianapolis, Ind. The bridge segments,

This precast segmental bridge near North Vernon, Ind., required about 25% less construction time than its arched companion bridge. Furthermore, the construction cost of duplicating the old arch design was prohibitive

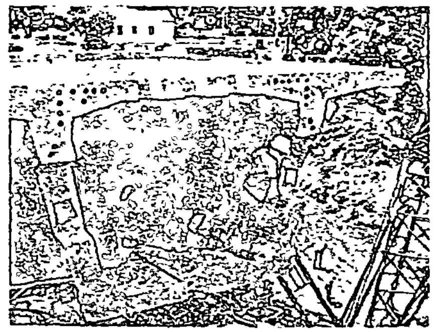
which the contractor started erecting in May, 1975, were added one at a time in each direction from a central pier. They were connected with steel tendons and an epoxy adhesive material. After the segments were in place, a protective membrane and a wearing surface were constructed over them.



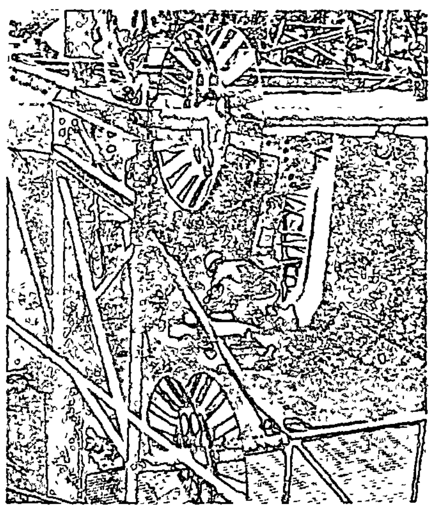
The two 10-ft (3-m) square, hollow piers have 18-in. (460-mm) walls standing on footings founded in sandstone. The hollow stems were filled with dry sand to produce more vertical load to counteract overturning moments. Piers reach about 50 ft (15.2 m) high. This Manitowoc crane is placing the first pier segment. By lifting the segments into place from below the only live load was materials weight, tensioning equipment, and personnel movements. If, instead of using a crane, the segments were lifted by hoist on the superstructure, the live load would have been much greater and a high impact factor would have been needed

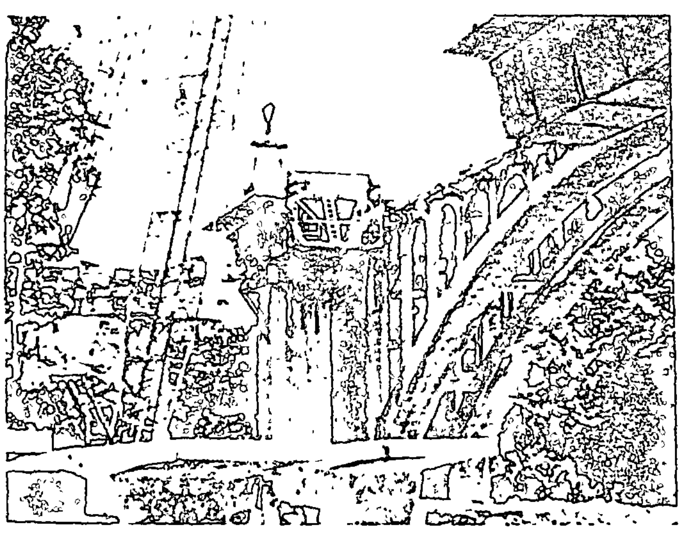
Segmental bridge pluses

Among the many advantages to segmental bridges is the shorter time period required for construction. These bridges can be constructed 20-25% faster than standard bridges because the segments are precast, large, and fit together perfectly. A savings of time means a savings of money in labor costs.



The cross sections used consist of a single cell box connected to the eastbound structure by a longitudinal neoprene expansion joint. Since the longitudinal reinforcing does not cross the joints between segments, no tension was permitted in the concrete. Pier segments have 4-ft (1.2-m) thick diaphragms inside the box sections to transmit loads from the superstructure to the bearings





Epoxy resin jointing was used between the segments to provide a watertight seal and act as a lubricant while joining the segments. Steel tendons were drawn through the channels and tensioned to support the cantilevered segments

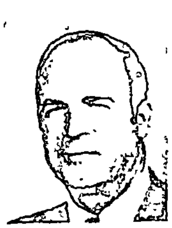
Another significant feature of segmental bridges is that work can be done from the top of the bridge causing less debris to fall into the river or roadway below. Therefore, these structures have less adverse effect on the environment and working conditions.

Segmental bridges also require fewer piers for support, making them particularly advantageous over highways and busy waterways. The hollow segments also allow a more thorough inspection of the structure. Even though the bridge floors of segmentals are expected to require the same amount of maintenance as standard bridges, the support structures are expected to better resist deterioration.

A few civil engineers in Indiana were searching for innovative and better means of constructing bridges in this country. In 1974 a number of them traveled to Europe to view, first hand, segmental bridge construction. These bridges, which had been built during the previous decade, seemed to have good application in a number of situations in Indiana. For this reason, a few were placed under design. The North Vernon structure was the first built in this manner.

The Indiana Highway Commission has awarded contracts for two more segmental structures since the completion of the North Vernon structure. The most recent is for a site on U.S. 136 over the Wabash River at Covington, Ind. It is a six span structure with a maximum span length of 187 ft (57 m). The contract amount is $3,313,388.84. The successful bidder is considering erection by jacking the segments from one end of the structure. This is a method occasionally used in Europe, but which has not yet been used in the United States.

The Commission also has under design a segmental structure over the Indiana Harbor Ship Canal in East Chicago which is expected to cost well over $30,000,000. This design will probably be let as an alternate to a structural steel design. ▽



Gene K. Hallock has been Chief Highway Engineer since January, 1974. Prior to that time, he served as Assistant Chief Engineer, Highway Operations, and in various capacities in the Construction Division dating back to 1953 upon graduation from Purdue Univ.

Computer cartography offers county unlimited combinations and considerable savings

Advanced computer cartographic methods are saving Sacramento County nearly $6,000 per sq mi (2 6 sq km), in their unified base mapping system. How? By computerizing map related information which can be updated, edited and corrected using a CRT unit, stored for later recall, and then computer-plotted to create maps tailored to specific county departmental needs and scales. Two systems were studied and compared to the more conventional method of scribed overlays Results of the study and a brief history of Sacramento County's pioneer mapping system detail the progress of the plan.

VERN W. CARTWRIGHT, P.E. M. ASCE
President
Cartwright Aerial Surveys, Inc
Datamap Systems, Inc
Cartwright Research
Sacramento, Calif

JOSEPH P. ALESSANDRI M. ASCE
Chief, Water Resources Division
Dept of Public Works
County of Sacramento
Sacramento, Calif.

A STUDY TO DETERMINE the feasibility of computer cartography for Sacramento County's unified base mapping system has proved the new technique outstanding in flexibility and convenience and revealed that, in the area tested, savings of nearly $6,000 per sq mi (2 6 sq km) will be realized on the initial mapping phase alone.

Additional savings will accrue through faster and improved information retrieval, availability and accuracy, additional map scales, and the elimination of costly duplication by map making departments within the county.

The enormous benefits and effects of computer cartography are just beginning to become evident—even to the principals who established the pilot program Detailed and accurate data banks (see Fig 1) of information tied to geographic locations will be useful to every phase of government, for maps keep track of people, taxes, property ownership, streets, buildings, utilities and many other entities.

In the past, individual maps were created for very narrow and specific purposes. Local government found itself with a multiplicity of scales and formats suiting a particular user's needs but useless to others. And, as departments and agencies grew, so did the duplication of maps and the cost to taxpayers.

Such duplication is no longer tolerable to county government, so new ways are being sought to increase the cost-effectiveness of mapping systems, and, at the same time, to expand their ability to meet users' needs

Computer cartography offers solutions to many previously insoluble problems. With interactive computer graphics and powerful software, computerized mapping has become a practical reality. Now, data can be economically converted to digital form, stored, changed, corrected, updated and recreated to suit the users' needs

This new technology, linked with aerial photography and photogrammetry, is being used in a pilot program comparing a conventional mapping system using scribed overlays with a Complete Map Information Management System (CMIS) (see Fig. 2) created by computer cartography

The beginning—1970

As their first step toward the control of proliferating systems, and to improve unsatisfactory mapping procedures, the County of Sacramento's Department of Public Works, in 1970, retained a consultant (Gennis, Grey & Justice, Sacramento) to analyze their existing systems and the needs of all map-making and map-using departments and agencies within the county A feasibility study was completed and a unified mapping system recommended.

The study (CIVIL ENGINEERING, ASCE, December 1973) outlined a controlled property line base mapping system tied to the California State Plane Coordinate System which could be combined with selected scribed overlays to form composite maps with any desired combinations of information. With this system, a user agency could select overlays of underground utilities, district boundaries, addresses, coordinated parcel numbers or other types of data.

The consulting firm also recommended the establishment of permanent monuments referenced to property lines and suggested that computerized graphic data storage and retrieval using the county's computer system be added later. Equipment then on the market permitted the computerized recording of X and Y coordinates of all points of interest for later retrieval by computer-driven plotter But the state-of-the-art at that time still required the preparation of manuscripts and map data into separate overlay form before the information could be recorded on punched cards or magnetic tape.

Pilot program

Funding to begin the proposed mapping program was not immediately available However, in 1974 the need for infiltration and inflow studies in connection with a regional wastewater treatment program provided the impetus to start the mapping effort A portion of the urbanized county area was selected to be mapped and funding was provided for the preparation of the controlled property line base maps showing underground sewers, drainage and water systems as well as two-foot contours, orthophotographs and in some cases planimetric detail

By this time computerized mapping techniques had matured to the point where the Department of Public Works felt that computer cartography should be tested against conventional mapping techniques A portion of the initial area was assigned to map production using computer graphics and a portion to conventional methods so the systems could be compared and the results analyzed for cost effectiveness and efficiency.

The area selected for the comparison study consisted of 96 map sheets, each covering 2,000 X 3,000 ft (610 X 914 m) at 1" = 100' (1 1200) scale on a pre-determined grid 39 of the sheets were to be pencil manuscript property base maps and overlays

for data bank digitizing, four sheets were to be created directly by computer cartography from source material such as sewer system maps and Assessor's maps; and the remaining 53 sheets were to be prepared using conventional methods with scribed overlays Orthophotographs of the entire 96 sheet area were to be prepared to aid map interpretation by the general public, to verify accuracy of property line plots and to locate manhole inlets, etc.

Cartwright Aerial Surveys Incorporated of Sacramento was retained to photograph a portion of the entire 96 sheet area and to prepare the topographic, planimetric, and computer graphic maps for the 39 and four sheet areas.

Before Cartwright Aerial Surveys flew the project, ground control was targeted along with other selected points such as manholes, valve boxes, etc. The mapping area was then flown at a photo scale of 1"=500' (1:6000). Glass diapositives were precisely drilled and read on Cartwright's monocomparator to an accuracy of one micron

Analytical aerotriangulation was employed throughout the mapping area and coordinates obtained for all desired points These computed points then supplemented the control for the topographic, planimetric, and computer mapping Control for the project was then computer plotted for each sheet of the 96 sheet area

Conventional mapping

Conventional map compilation was accomplished in the 53 sheet area by Gennis, Grey & Justice and orthophotos with contour overlays prepared for each map sheet. Overlays were prepared from Assessor's plates, recorded subdivision maps, records of surveys, and deed research and were scribed and photograhically combined to produce the desired composite maps The cost of these conventional maps totaled $4,046 for each 2,000 X 3,000 ft (610 x 914 m) area

Pencil manuscript digitizing

In the 39-sheet section, Datamap Systems, Incorporated, a Cartwright

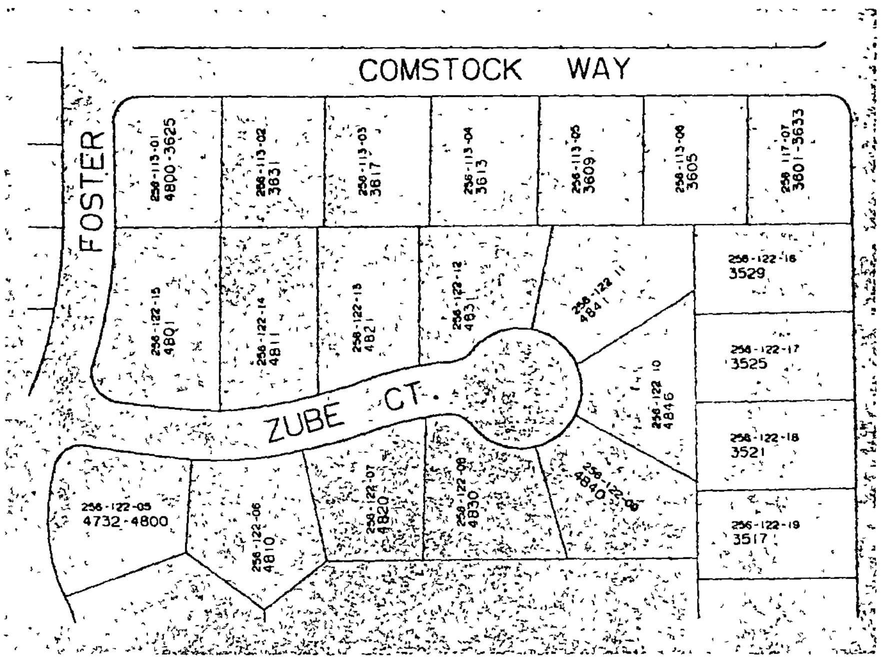


Fig 3 Sacramento County computer plotted planning information map Reproduction of computer plotted map is created by extracting specific information from bank of digitized map data

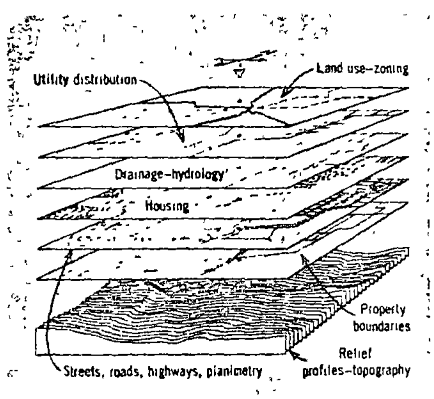


Fig 1 Graphic representation of a Digital Data Bank Data banks are homogeneous sets of data which have been compiled and stored for later recall to produce maps for specific purposes They are the building blocks of modern state-of-the-art mapping They may be recalled randomly for CRT (cathode ray tube) viewing, singly or in combination, edited, changed and re-stored without physically drawing a map Data bank versatility is unlimited Any number can be selected to satisfy specific map requirements The combinations which can be used are near infinity Maps tailored to planimetric detail, topography, land use data, water resources, parcel boundaries and forestry, just to name a few, may be produced from one aerial survey

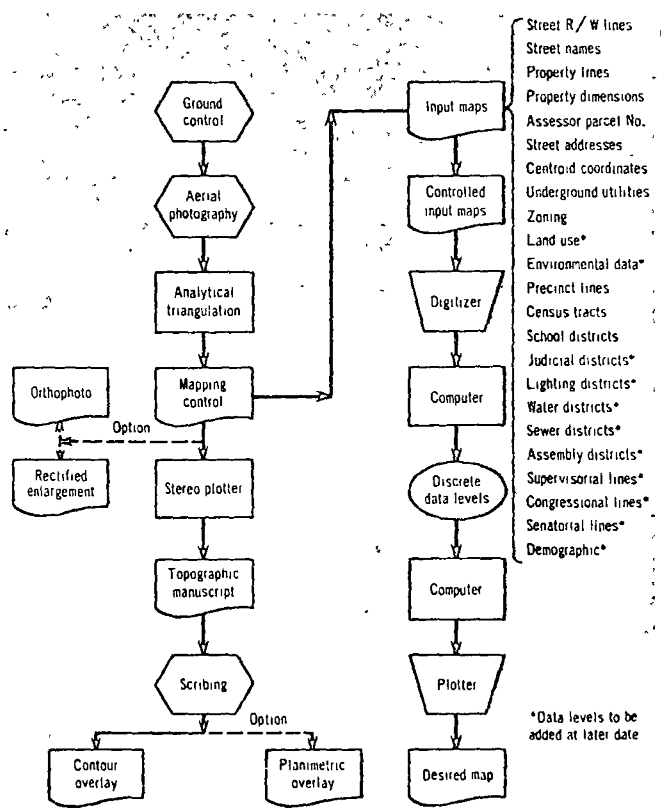


Fig 2 Sacramento County Base Mapping System Flow Chart

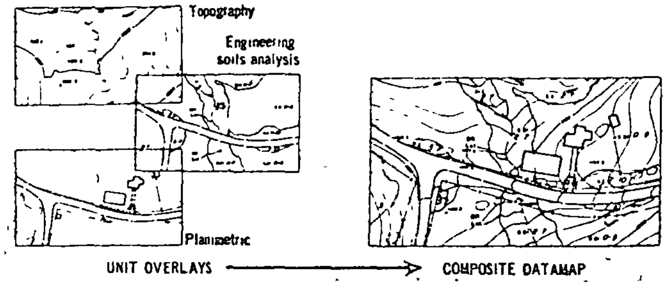
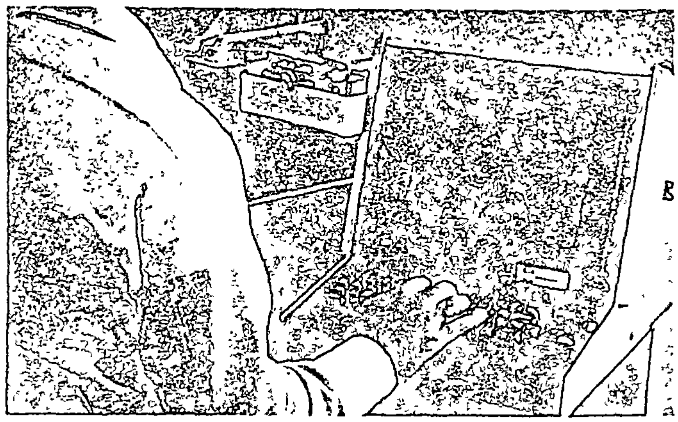
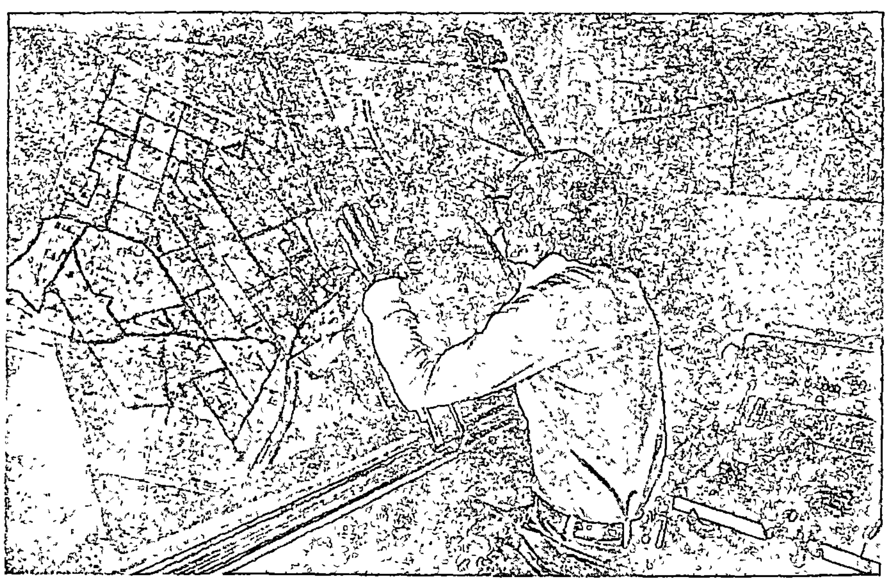


Fig 4 Unit overlays on the left (topography, engineering soils analysis, and planimetric), can be combined into the composite datamap on the right, or numerous other combinations



CRT Editing station All information entering the computer memory is edited and corrected as necessary, at the CRT editing station Data banks can be recalled and changed here without disturbing the other map elements



Digitizing from source material The operator moves the cursor which is mechanically linked to encoders that count increments of movement in the X and Y axes Cross-hairs position the point and a switch located on the cursor record the X-Y coordinate In straight-line work, the cursor is moved from point to point, curved lines are digitized by recording the beginning, the middle, and the end of the curve The computer then fits a curve through the three points

Corporation, digitized road and property line hand-prepared, pencil manuscripts furnished by Gennis Grey & Justice These manuscripts were used as a base and overlays of the sewer system, underground drainage and water distribution systems were digitized to develop discreet data levels of information As these overlays in their initial preparation had not been registered with each other there were conflicts in the positioning of street names, numbers, dimensions, and other information and, as a result, it was necessary to reposition and redraft the annotation before digitizing Special software was written to accomplish this automatically Also, it was necessary to computer-adjust the various overlays for agreement with the property line base maps using the CRT (Cathode Ray Tube) (see photo) Since the CRT displays the data graphically, omissions and errors were easily spotted and corrected In this area of the County data levels were developed for property lines, street names, sanitary sewers, underground drinage and underground water distribution systems

Computerized mapping

Computer cartography was used to the maximum in the four-sheet area, digitizing directly from source material (see photo) without the intermediate step of creating hard copy manuscripts Current Assessor's maps were used as the source documents for the property base. Where possible, coordinates of horizontal control points were assigned to identifiable locations on these maps It was determined that four control points were the minimum required to control each Assessor's sheet being digitized.

Where fewer than the four required control points could be located, a mathematical traverse was employed to supplement the control and to verify its accuracy Each control point was digitized and its real-world coordinate position entered in the computer The computer then checked the validity of the control by determining the distance in inches between the digitized control points and by comparing these values to the distance in feet between the real-world coordinate positions Where the control was valid this comparison yielded a computed map scale that did not vary more than one or two percent from the scale indicated on the source document

When the scale variation was excessive, the following conditions were checked for error.

1) the plotting of the ground control on the source document,

2) the coordinates assigned to the ground control itself,

3) real-world coordinates entered in the computer, and

4) the drafting of the source document

If the control placement, assigned coordinates, and the computer entries were verified as correct, then the original Assessor's map or manuscript was in error In such cases, a special computer program written by Datamap Systems reconciled these drafting errors (up to 10 percent of the scale used on the source document) and costly, time consuming redrafting was avoided

To demonstrate the ability and flexibility of the mapping system and to further test the capabilities of computercartography, the number of overlays for the four sheet area was expanded to include property lines and dimensions, street names, addresses and rights-of-way, tract boundaries and labels, zoning, census tracts,

school attendance districts, voter precincts, Assessor's map pages, sanitary sewers, underground drainage systems, underground water distribution systems and paracentroid coordinates In addition, several of the data levels were divided into sub-levels Sanitary sewers, for instance, included pipe diameters, manhole depth, flow line elevation of exit pipe at manhole, manhole standard symbol, manhole drop connector symbol, flush branch symbol and privately owned sanitary sewers

After digitizing and editing on the CRT, the maps were computer plotted in ink-on-mylar with as many data levels as needed (see Fig 3) Each map was gridded with the California State Plane Coordinate System and indexed on the lower left-hand corner with an alphabetic prefix to indicate the scale of the map series The maps were plotted on a 42 X 76 in. (1 07 X 2 03 m) flat bed plotter to a resolution of 0 001 in (0254 mm) The cost from aerial photography to plotted hard copy map was $2,787 (see Table 1) for each 2,000 X 3,000 ft. (610 X 914 m) area, or $1,259 less than for scribed overlays for the same area

TABLE 1. Base mapping cost summary[1]

Item	Scribing	Computer
Aerial Photography	$ 35	$ 35
Analyties	112	112
Surveying (Ground Control)	900	900
Contour Sheets (2' interval)	385	385
Orthophoto 1" = 100' (incl half-tone positive)	285	285
Sub Totals	$1,717	$1,717
Property Line Resolution	$1,068	$ 675
Data levels		
Street Names	65	35
Sewer	630	225
Drainage	566	135
Sub Totals	$2,329[2]	$1,070[3]
TOTALS	$4 046	$2,787

[1] Cost per map sheet, 2000 ft x 3000 ft at 1" = 100'
[2] Includes scribing
[3] Includes digitizing and plotting

TABLE 2. Comparison of unified base mapping methods

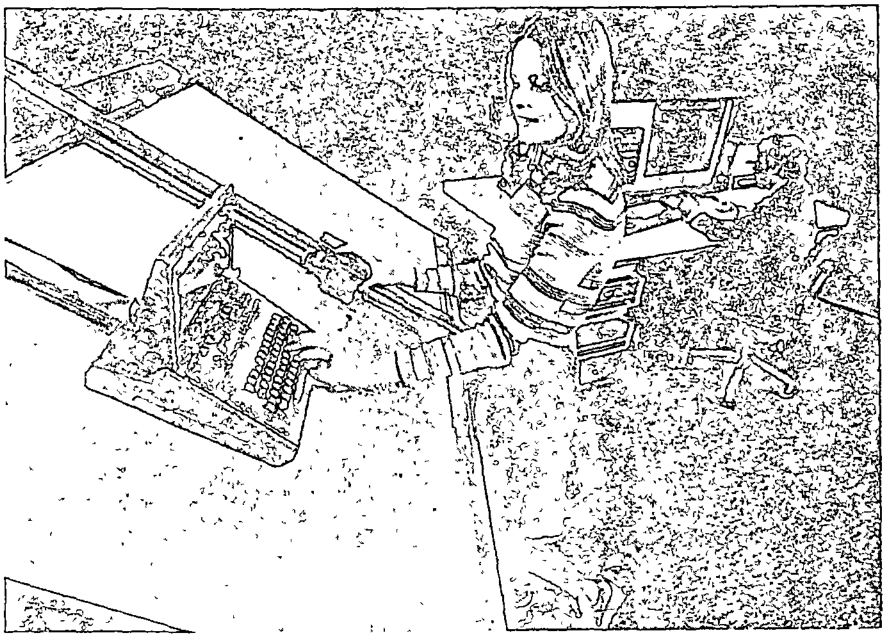
System features	Conventional scribed overlay system	Complete map information management system
Change of scale	Limited	Unlimited
Sheet layout	Fixed	Fixed or Random
Conversion to metric	Difficult	Easy
Ease of updating	Difficult	Easy
Clutter factor	Some	None
Common scale of source documents	Required	Not Necessary
Custom maps	Limited	Unlimited
Production	Slow	Very Fast
Data conflict avoidance	Very Difficult	Easy
Number of overlays or data levels per map	3 to 6	Unlimited
Accuracy	Human Factor	0 001 inch

Advantages

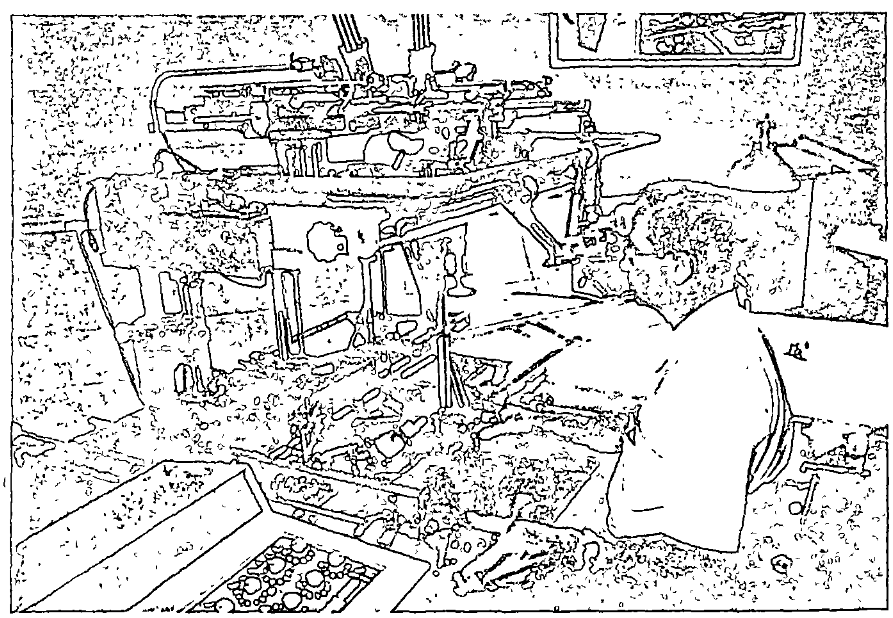
Some major advantages of computercartography over conventional map producing methods were soon apparent For instance, conventional mapping materials such as scribed negatives for overlays require considerable storage space, while the same information stored on computer tape is compact and accessible

Also, while scale changes from scribed overlays could be accomplished photographically by enlarging or reducing, format was locked in and lettering and line widths increased or decreased accordingly Computerized map data, on the other hand, could produce maps with letter and line sizes selected to match the scale of the maps without compromising aesthetic qualities and legibility

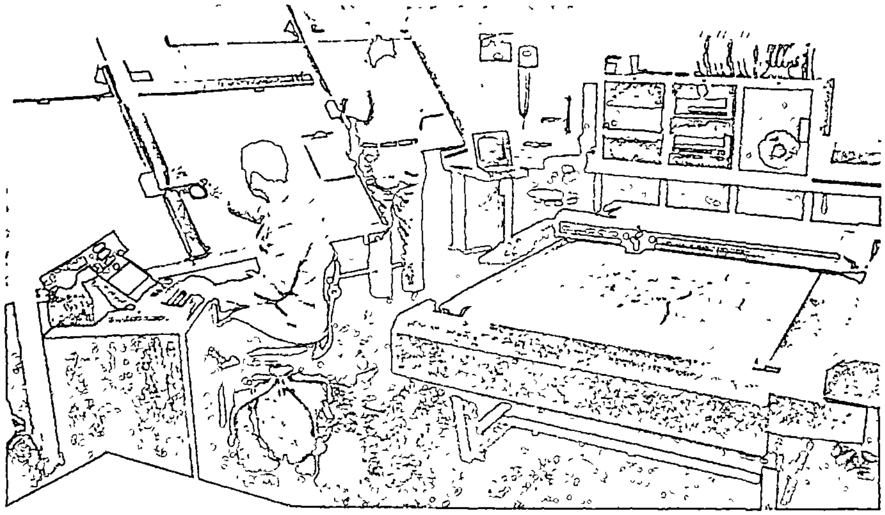
One of the greatest advantages of computer cartography over conventional map-making methods built into the Complete Map Information Management System (CMIMS), is the ease of metric conversion. As the maps are stored in digital form, they can be



Macro keyboard Symbols (macros) after being coded and placed in the computer are accessed by placing the macro indicator on the console keyboard at the desired symbol positioning the cursor where the symbol is needed and pushing the macro switch on the cursor to record Annotation (street names, numbers, dimensions, etc) is accomplished by pre-setting a thumb-wheel switch on the digitizer keyboard console to the desired lettering heighth, and recording a starting position and a second point to indicate vector or direction of the lettering string The annotation is then entered using the alpha/numeric keyboard of the digitizer console



The Santoni IIC stereo plotter used for map compilation



Interactive graphics set-up Moving left to right Macro keyboard used to record symbols, dimensions and other annotations, digitizer with cross-hair positioner to locate X-Y coordinates and cursor arm mechanically linked with encoders that count movements in the X and Y axes, CRT Editing station, computer, disc drives, tape transport, and line printer Foreground Flat bed plotter Selected data banks are combined to form tailored maps on the flat bed plotter Computer driven, the plotter produces the desired scale with extreme accuracy

readily converted by computer and automatically replotted with metric dimensions when the system is adopted.

A similar and even more significant advantage is realized in map updating Handling of bulky overlays required by the scribed system is eliminated in computerized mapping New data and changes in old data are entered into the computer, verified with the CRT, and hard copies plotted as required

Cost comparison

The savings realized through computer cartography compared to the conventional mapping amounted to \$5,843 82 per sq mi (2 6 sq km), or a total of nearly \$6 million to be realized in the 1,000 sq mi (2,600 sq km) of the County of Sacramento

Actually even these amounts do not represent true dollar differences between the systems Only five scribed overlays were prepared for both the conventional mapping and the 39-sheet digitized areas while in the four sheet area 16 discreet data levels and 35 sub-levels were developed

Even if the initial cost savings for computer cartography were not so dramatic, there would still be powerful advantages (see Table 1) According to Brian Richter, Sacramento County Director of Public Works, "Without question the computer can produce a myriad of variable map displays It has all the information and it can adjust scales and formats I think through the pilot program we can begin to sell the system"

Does the County plan to fully implement computercartography immediately since the cost savings are so evident? The County staff is considering several alternatives.

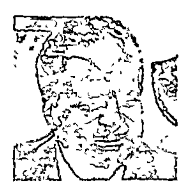
If sufficient funding becomes available to prepare the property line base maps county-wide, the County can then contract the digitizing, compilation, storage and computer plotting to computer graphics firms or purchase the computer graphics equipment and prepare the maps in-house If funding is not available to complete the county-wide system, a gradual shift to digital data base mapping will be accomplished as new mapping is required and funded.

Eventually a new centralized mapping service group might be created within the County and procedures established to facilitate the production of all maps

Even with many questions still to be answered about implementation, Sacramento County's pilot program has answered the big one—the bottom line—"How much does it cost?" Without a doubt, computer cartography saves money and solves problems in addition to providing a versatile and flexible map information management system for the future.

Credits

DATA GENERAL, Nova 1200, 32K core computer; *H DELL FOSTER CO.* Digitizers with macro keyboards, *H DELL FOSTER CO.* 42 X 76 inch plotter, *TEKTRONIX* CRT (cathode ray tube), DIABLO disk drives, 2 5 million words, *WANG,* 9-track, 800 B P.I tape transports, *CENTRONIC* line printer, *H DELL FOSTER CO* monocomparator, *KELSH* plotters, *SANTONI* IIc stereo plotter, *AUTOTROL* digitizers, *ZEISS* RMK A 15/23 (6 inch) camera, *ZEISS* RMK A 21/23 (8½ inch) camera, and a *WILD PUG* III point transfer device with zoom



Vern Cartwright organized Cartwright Aerial Surveys, Inc in 1946, and has since provided aerial photography and photogrammetric services to government, private industry, and professional clients.

Joseph P Alessandri worked for the California Division of Highways and the California Department of Resources before joining the Sacramento County Department of Public Works in 1963 He has been Chief of Water Resources since 1970

Do You Know That

. . . . the 55-mph national speed limit has resulted in safety benefits of major proportions? A by-product of the 1974 energy crisis, the reduced speed limit throughout both 1974 and 1975 yielded a reduction of 17% in the number of people killed on the highways as compared with 1973 The decrease in the fatality rate occurred despite numerous obstacles which could have produced a trend in the opposite direction The National Highway Traffic Safety Administration of DOT cited such problems as a growing number of vehicles, drivers, and miles traveled, an increase in youthful drivers, alcoholic abuse, and a widening disparity in vehicle mix (more heavy vehicles and more small cars on the road)

. . . the United Nations says the world water supply is all wet? The U N reports that at least one-fifth of the world's city dwellers and three-quarters of its rural people lack reasonably reliable supplies of drinking water The report, entitled "Resources and Needs Assessment of the World Water Situation," warns that it might become necessary to relocate populations closer to sources of readily available water In light of increasing population and industrialization, the report notes that water cannot be viewed as an "inexhaustible gift of nature" The document predicts "a critical shortage of water of suitable quality to sustain future growth unless water management is radically improved" (Excerpted from *New York Times,* July 22, 1976, page 2)

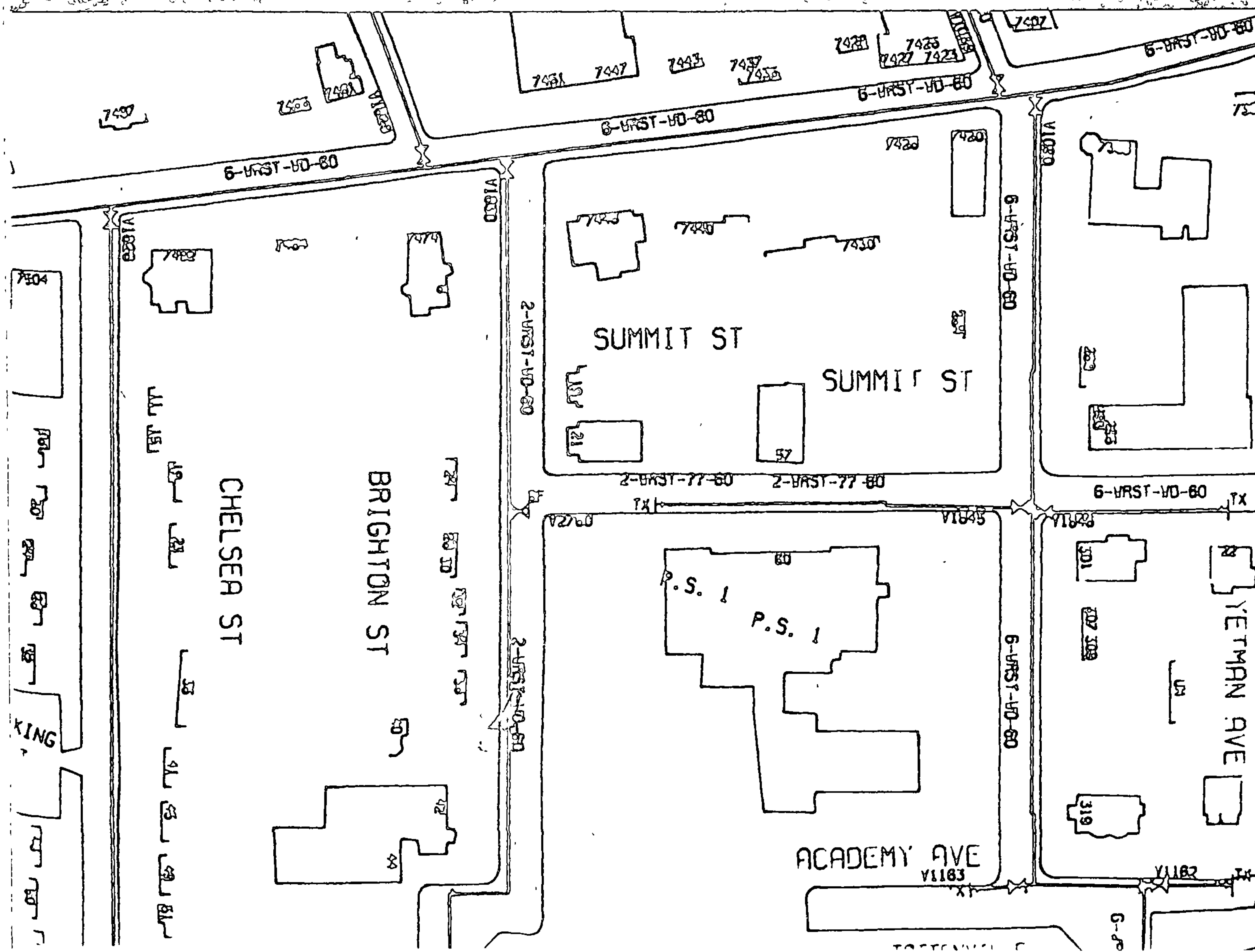
Civil Engineering

Making maps by computer

For making maps, digitizer-computer-plotter systems are increasingly popular. Here are case histories of three applications—by a gas utility for keeping track of 650 miles of distribution pipe and thousands of building-service lines; by a small consultant for making subdivision maps; and by a state highway department for mapping 500 miles of highway a year.

In all three cases, significant reductions in cost of making maps have been achieved or are in prospect. However, not all civil engineering organizations using such systems have saved money. So we offer some guidelines to those thinking of purchasing such hardware or services.

KNEELAND A. GODFREY, JR., M. ASCE
Editor
CIVIL ENGINEERING—ASCE



Gas utility pioneers computer mapmaking

New York City's Brooklyn Union Gas Co has one of the nation's most sophisticated mapmaking-by-computer systems It should quickly pay for itself in savings resulting from the more precise location it gives BU of its 650 miles of gas mains on Staten Island All the map information is stored not only on maps but on magnetic tape, permitting easy analysis and reporting in non-map ways

Q What was wrong with the existing maps?

A About 20 years ago, BU acquired a smaller gas company that served Staten Island Showing location of the Staten Island gas distribution pipeline system were no conventional maps, but 18,000 not-to-scale sketches in hundreds of field books The sketches were kept by order of date of installation of the pipe (starting nearly 75 years ago) There was a geographic index to the field-book sketches

These sketches showed pipe locations referenced to a landmark such as a house front, pole, curb or hydrant Unfortunately, the location of many curbs, for instance, has changed since then, and the sketches had not been updated

This map inadequacy cost money In its service area, also covering Brooklyn and Queens, BU has 3,600 miles (5,700 km) of pipe, and 550,000 service connections to buildings (Staten Island has about 20% of the total system)

Each year, BU must locate and repair leaks, attach and install new underground facilities, locate and replace certain existing facilities To make these alterations they must dig about 40,000 holes

Doing this costs an average of $300 per 3x3 ft (0 9x0 9 m) hole, or about $12 million a year.

Q What can be done to cut this cost?

A Better maps—more accurate and more complete—might trim it by reducing the number of holes dug per repair.

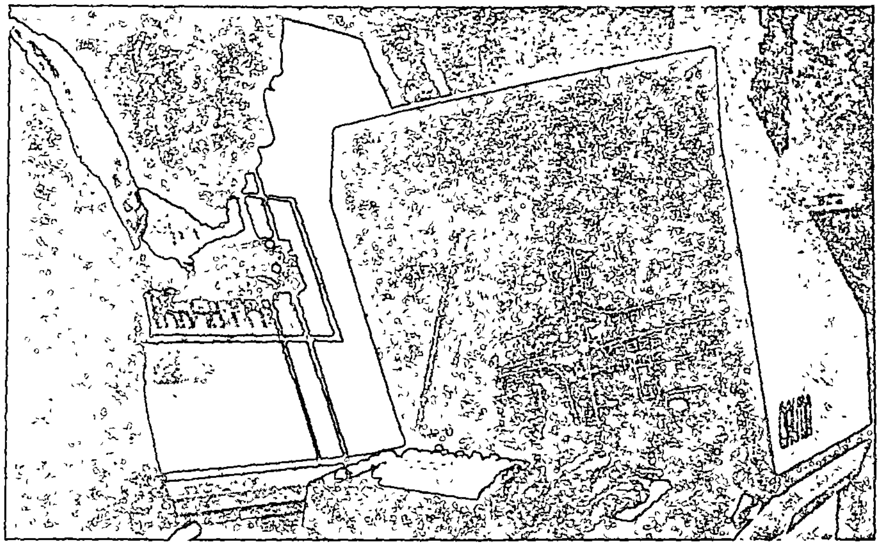
Originally, BU planned to create maps conventionally That would mean making a series of acetate overlays, one for each class of map data—streets, pipes, etc. But analysis of costs and benefits led instead to selection of a new computerized approach (See pictures.)

Q What did it cost?

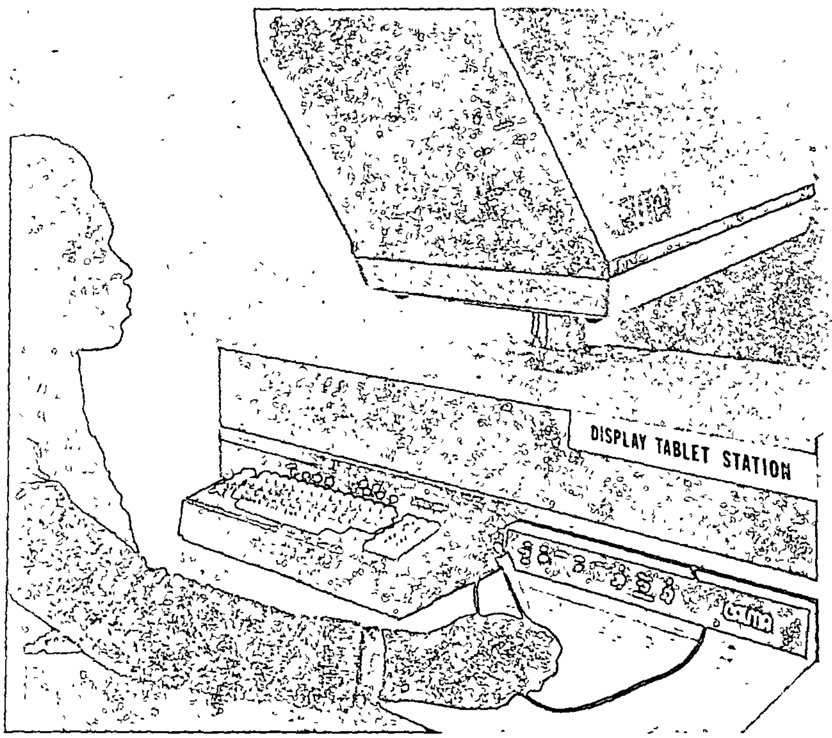
A Raytheon prepared the digital data base, containing the information that directs the computer-driven plotter in making the maps This cost $1 25 million. For another $250,000, Calma provided computer hardware and programs with which BU itself will update the data base and map.

Q Will the system pay for itself?

A BU hopes to recover this front-end cost in a very few years, in savings



TV screen (CRT) shows map of street intersection and gas pipes on Staten Island Man adds information using "digitizer" in his left hand When digitizer's cross-hairs are centered over object that he wishes to add or remove, he keyboards appropriate instruction Instantly, the changes he's made show up on CRT screen



Previous photo shows station used in inputting to computer memory the information from conventional maps Pictured here is station used for changing maps already recorded in the computer Key input devices are "pen" in operator's hand and tablet under it As operator moves pen image of it moves on CRT In this way, he instructs computer to insert remove or change information in computer memory

resulting from digging fewer holes. Other savings should come from non-map uses of the data base (See box.)

BU set the goal of finding the leak or facility, in the first 3x3 ft hole dug, 90% of the time. Based on this requirement, it was determined that the maps must permit the field crews to locate pipe in the ground to within 1.0 ft (0.30 m) horizontally, 0.9 ft (0.27 m) vertically.

A skilled BU office man, using information phoned in by field crews equipped with leak detectors, plus the new maps and his judgment, should be able, usually, to locate a leak rather precisely.

Making the new maps

Q. How was the new mapping system created?

A. First step was to get an accurate geographic data base—street and building locations. This might have been created from existing city maps. But it was found 25% less costly to shoot new aerial photographs and make new maps from these.

Using a stereoplotter and pairs of aerial photos of the same ground, distortions of ground locations on the photos due to elevation differences and film-plane tilt were corrected.

On these corrected photos, pipe loca-

Leaks, corrosion, explosions—what do they have to do with gas pipeline maps?

Federal agencies are promoting communication among public and private utilities having underground facilities. Goal is to minimize damage to (or damage caused by) these systems. An example is gas pipelines.

Q. Are leaks and explosions of gas pipelines a concern?

A. Historically, the gas industry's record has been good but not perfect. There are two major causes of leaks—corrosion of in-place pipes, and accidents during subsurface construction by an outside party.

Q. How serious is the corrosion problem?

A. At a few times and places in years past, very serious indeed. A landmark case, reports Frank Fulton of DOT's Office of Pipeline Safety Operations, came in Natchitoches, La., in 1965. A large, high-pressure gas transmission line ruptured due to stress corrosion. Escaping gas drifted over to nearby homes, and ignited (didn't explode). Several houses burned, and 17 people died.

Q. What has been done about gas pipeline corrosion?

A. Federal regulations now require that, in the case of steel pipelines, gas utilities:

(1) install cathodic protection systems, which neutralize the corrosion potential of the soil and groundwater;

(2) periodically test that cathodic system for adequacy, and

(3) periodically run a "leak survey". Where a truck can run atop the ground over the pipeline, the survey is run fairly speedily. Sensitive equipment in the truck measures natural gas concentrations, and this helps the utility to pinpoint location of the leak.

Excavations and gas pipeline leaks

Q. What about construction-caused gas pipeline leaks?

A. To minimize chances of a construction accident, utilities must use one or more of the following approaches.

- Post markers along the line, indicating presence of a gas line (the usual practice in open, undeveloped areas).
- In cities, the feds and others are encouraging adoption of a "one-call system". Two days before a contractor's backhoe hits the earth, he (or the agency for whom he's working) must call the "one call" operator. This person in turn calls all utilities and public agencies whose pipes or lines are buried beneath or near the area to be dug. Then, before excavation starts, each affected agency sends crews to the site to paint lines on the street indicating location of their pipes.

A significant share of gas pipeline explosions have been caused by excavations done by another utility. These accidents are being sharply cut, thanks to "one call" programs recently established in 80 cities.

Number of cities having one-call systems (in which all subsurface pipe owners agree voluntarily to cooperate) has skyrocketed in the past year, from about 20 to perhaps four times that many.

A major reason for this spurt is that in September 1974, at the American Public Works Association annual meeting in Toronto, at the urging of the National Transportation Safety Board, utilities, public works departments and others formed a Utility Location and Coordination Council. ULCC was formed partly to promote the one-call idea (and rationalized mapping systems).

NTSB's Charles Batten tells CE that one-call systems, where effectively operated have "dramatically" reduced the number of construction-caused accidents to gas pipelines.

This is not an idle concern. For example, in January 1968, in Reading, Pa., a backhoe was excavating earth to permit repair of a water main. It struck a ¾ in. (18 mm) gas line, bending it. No immediate problem was apparent. Then, two hours later, in a nearby building, there was an explosion killing all nine occupants. (This case, Batten says, is an example of a number of such accidents which lead to fires or explosions, some time after the construction incident.)

Q. What do maps have to do with this?

A. In some communities, reports Batten, adequate records and maps of gas pipelines are "nonexistent." In some cases, the map may indicate presence of a pipeline, but so imprecisely that one can't be sure which side of the street it's on. If the utility is to mark or paint-stripe its pipe locations, it must know precisely where they are. Some gas utilities do not have precise-enough maps to do this.

As a national demonstration program in cooperative mapping by public and private utilities, ULCC is helping finance a major demonstration at Memphis, Tenn. The city, county, Memphis Gas Light & Water, and South Central Bell are working together, currently they and others have perhaps 10 sets of maps covering the same area—money can be saved by cooperating. Goal of this program is to establish recommended standards and procedures.

Q. Where can I get more information?

A. The APWA Utility Location and Coordination Council's address is 1313 E. 60th St., Chicago 60637.

Nonmap uses of the data base

One of the major reasons for turning to mapmaking by computer is because the data base will make possible nonmapping applications.

Q. Can you give examples?

A. The San Antonio (Tex.) City Public Service Board, a public gas and electric utility, has maps tied to the Texas state plane coordinate system, on which are plotted all electric-line poles and gas mains.

In addition, the city's Police Department uses the data base as an aid in dispatching squad cars. Seems that San Antonio is an old city, of mixed Mexican and American heritage. In some cases the same street changes its name four times within a distance of only 10 blocks Often the police receive an emergency call—and the caller gets the street name wrong, or gives a non-existent street address.

Thanks to computerized street information, the police can correct false information (which is usually not knowingly given incorrectly).

Procedure A telephone clerk takes the call, enters caller's address and nature of call, into computer memory. The computer program checks to see if there is such an address, if it's in the city or a suburb, and if it's a multi-family address (in which case the clerk asks the apartment number)

Computer also determines from which of 212 patrol-car territories the call comes; forwards call to proper dispatcher, and prints out which patrol cars are nearest the caller. (System also keeps track of status of each patrol, so as not to bother already-busy cars)

System will, each month, plot a crime-contour map of each car's territory—by shift This should help policemen be at the right place and right time As crime patterns shift, system will help Police Department to equalize the work load of the 212 patrols.

Thought to be a national leader, this system was developed with help from the H Dell Foster Co., San Antonio. System will go operational next June. For more information Capt. Jacques Hardy, San Antonio Police Dept, P O Box 9346, San Antonio, Tex.

Electric utility analyzes distribution transformer load

Q. Who conducted the study?

A. This has been done cooperatively by four electric utilities—Arkansas Power and Light, Mississippi Power and Light, Louisiana Power and Light, and New Orleans Public Service Inc.—that together constitute the Middle South Utilities System.

Q. How was it conducted?

A. Don Preston of the New Orleans company explains: First, aerial photos were made of the 95 sq mi (246 km²) service area of the utility.

But map information was *not* digitized, except for locations of 20,000 transformers. Why digitize this information? "Its geographic location, as indicated by X and Y coordinates," Preston explains, "was the only unique, identifier number for each transformer that wouldn't change."

Q Why make the network analysis?

A So as to more optimally size transformers—making them big enough to minimize failures, yet small enough to minimize capital cost. In the past, crewmen equipped with meters climbed transformer poles; but NOPSI has too many transformers to gather information for analysis this way.

A better way was found—a statistical correlation between a customer's kilowatt-hour usage of electricity and the transformer peak load. These correlations were fed to the computer, as was information on customer/transformer associations

The utility believes they will save money, for if they can predict which transformers may fail, these units can be replaced on the field crew's regular time, not overtime at night (summer load peak is in the evening, due to air conditioners) (For more information. Don Preston, Mail Unit 35, NOPSI, 317 Barrone St., New Orleans, La. 70112.)

The police in San Antonio will use a geographic data base and computer to more efficiently deploy their 212 patrol cars, perhaps even to realign patrol districts based on geographic distribution of crimes.

Q. Any other non-map uses of geographic-indexed digital data bases?

A Indeed there are But, the potential is just beginning to be utilized. Know of any notable examples?

tions were pinpointed precisely in this way: Before the aerial photos were shot, large rings had been painted around each BU manhole (for valves) Gas pipes generally run straight between these manholes.

The geographic coordinates—north-south and east-west—of each curb, building front, and gas manhole were recorded on magnetic tape (This was reportedly the first big job on which this "digitizing" was done directly from the aerial photos, without first hand-drawing a map of the lines to be digitized. This cut Raytheon's costs.)

Next, locations of pipe fittings and building-service connections on the mains were digitized. This was done by

Brooklyn Union Gas spends $12 million a year to dig 40,000 holes (to repair pipe leaks, add or replace fittings). If pipe locations are pinpointed better, number of holes dug will be cut, saving money. BU's new computer map system will hopefully do that, soon repay its $1.5 million cost.

inputting to mag tape the detailed information in the 18,000 field sketches As a check on accuracy, all the information was plotted.

Q. Did everything fit?

A No Many errors were detected, some in digitizing, but many because errors or omissions had been made years before in the field books

For example, 20,000 ft (6 1 km) of French cast iron pipe was found—even though none was shown in the sketch books

And there were places where an impossible or illogical connection was shown in a sketch—such as the connection of a 4 in (102 mm) pipe to a 2 in. (51 mm) pipe without use of a reducer. To determine what was actually in the ground, in some cases BU has been digging test holes.

In all, some 500 validation checks were programmed into Raythoen's computer, to check completeness and accuracy of the data base

More than once, BU crews have been sent into the field to replace unsafe or inappropriate pipe or fittings or corrosion protection, uncovered in this way.

Next, after Raytheon delivers a map to BU (the job will be completed about mid-1977), BU itself runs field checks. Some errors are being found, reports Eric Osterberg, engineering director of BU's Distribution Department, but generally the maps are "extraordinarily good."

Q In what form are the completed maps?

A Blowups of the stereo compiled photos and maps are being delivered in three forms.

- of land areas 1,000 ft (305 m) square. These maps, to scale 1" = 40 ft, show every piece of pipe and every fitting.
- maps and photos of land areas 3,000 ft (915 m) square, to scale 1 in. = 120 ft; and
- an index map on which are referenced all the second class of maps

Q. How will the digital data base be kept up to date?

A. Unless it's kept up to date, in time the data base's usefulness will drop sharply. Updating will be done using the Calma system in several ways:

- On the "Digitizer Table" (see photo), the operator places a map of the information to be digitized. The map can be to any scale, oriented on the table in any way To input new information, the operator spots his "cursor" precisely over each key point on the map, pushes appropriate buttons (indicating the type of information being recorded), and then pushes a button recording the new information, and erasing the old from the mag tape. This approach is used if entirely new streets are being recorded from builders' plot plans.
- But on Staten Island, most streets are in place, and already recorded on the BU data base So changes are made in another way: The street or intersection in question is brought up on the cathode ray tube (CRT). The operator revises this electronic "map" using a "pen" and a typewriter-like keyboard.

A check of what he's input to the mag tape—its nature and accuracy of positioning—shows immediately on the CRT, and errors readily corrected.

Q. What hardware is in the BU-Calma map-updating system?

A. Key items of hardware in the system include a mini-computer by Data General; a disk drive by Century; a tape drive by Wang; a tablet-editing station by Calma; a plotter and controller by Calcomp, and CRT's by Tektronix.

But perhaps more important to the success of the system than hardware is the computer programs, because they make it easy to update a map. Until two years ago, updating a digital data base was a "nightmare," says Osterberg, and no programs had what BU considers adequate validity checks. (See box.)

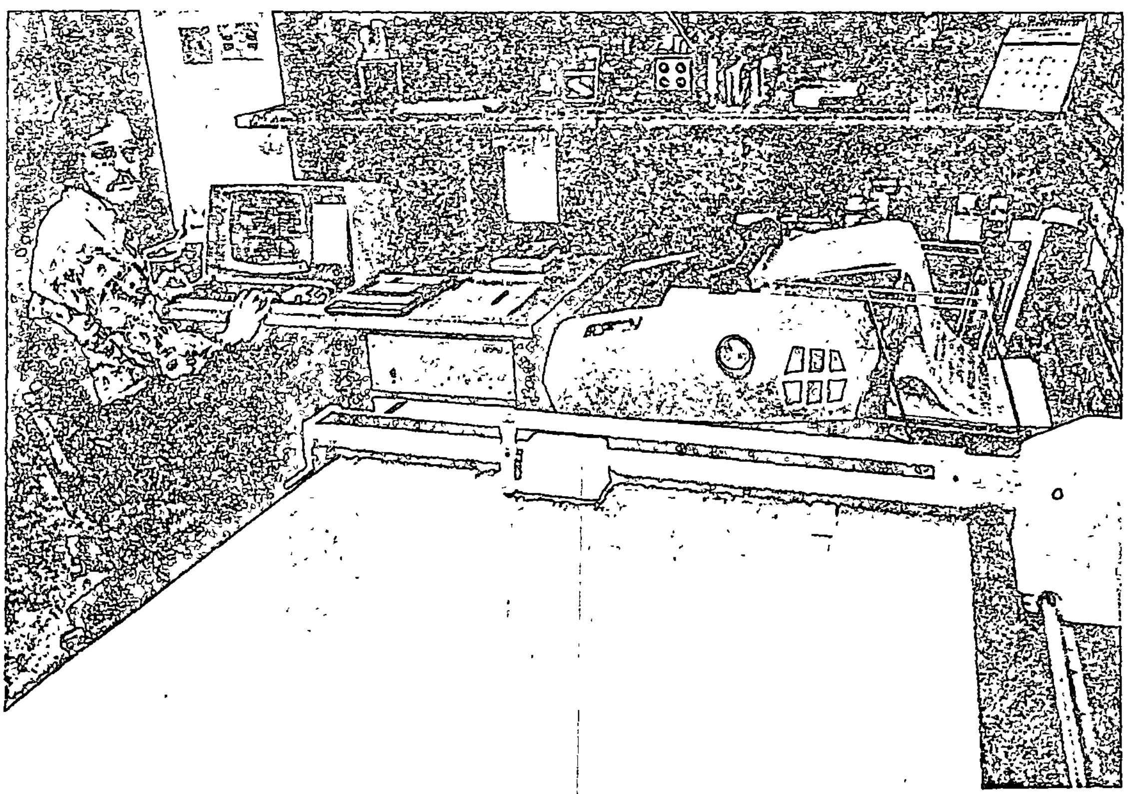
Subdivision maps made by computers

Three years ago, Development Consultants, a small civil firm in Houston, began using a computerized mapping system. Perhaps even owner Nolan Purser, PE, didn't realize how it would pay off:

1 *Speed of map making* has increased almost unbelievably. A dramatic example was a 715-lot subdivision.

Data Processing Manager Don Quinn did 90% of the office work himself in one week In pre-computer days the job would have taken five draftsmen six months.

Lots in the subdivision average 65 x 120 ft (20 x 37 m). "The total job involved computations, platting of two sections with all annotations including



In computer mapping system for subdivision work, key hardware elements shown are, from rear keyboard and CRT alphanumeric line printer (gives non-map output from computer), and plotter (automatically draws maps) Not shown is fourth key item digitizer table (used to transfer information from conventional map to computer)

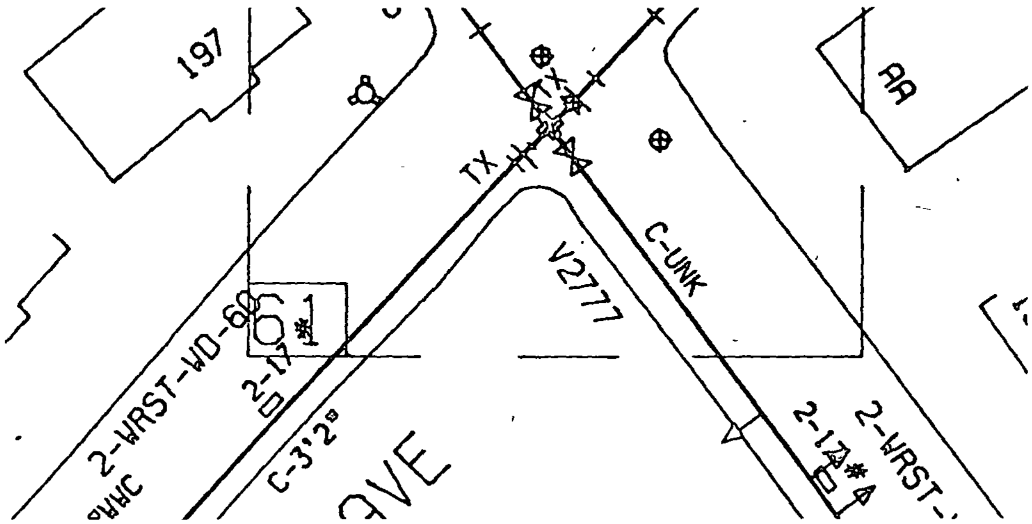
Plotters—many types

A computer-driven plotter draws faster than a draftsman—from 10 to 50 and more times as fast.

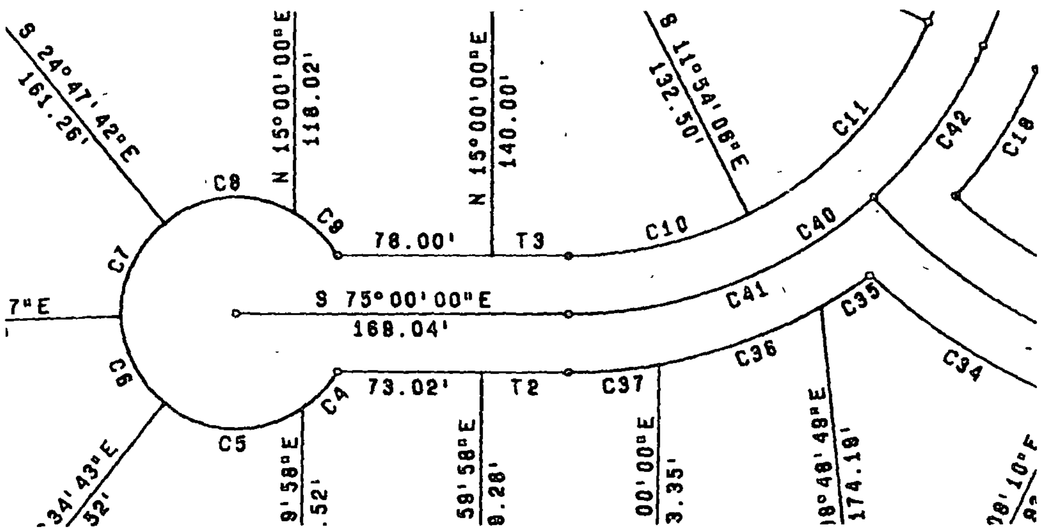
Many plotters draw a line less precisely, or make curves less smoothly than a draftsman can. However, the lower quality work is acceptable for many or most purposes

Other plotters do as good a job or better than a draftsman. But they are very costly.

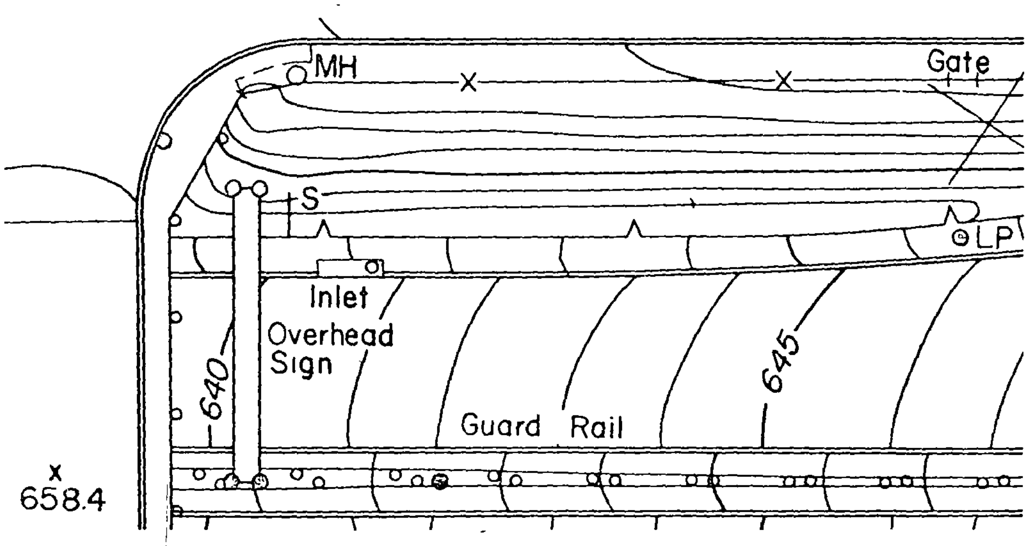
Here are samples of line work by plotters used in the computer systems described in this article. The line-work is reproduced to scale.



Section of map drawn by Calcomp's new vertical plotter, which is part of mapping system at Brooklyn Union Gas Co Speed 30 in /sec with ballpoint, 10 in /sec with liquid ink Resolution 0 0005 in Maximum size 33 x 60 in Cost $30 000 for plotter alone, plus $27,500 for 8K programmable controller (which frees the central mini-computer in the mapping system to do other things)



Sample of map drawn by Wang plotter used by Development Consultants, Houston Cost is modest—$7,000—as plotters go It draws maps up to 31 x 48 in, and is used for all mapping jobs at this firm Speed is described in accompanying story, "Subdivision Maps "



Highest in cost—$150,000 to $200,000—is photohead plotter But, as this sample shows, line quality is also the highest (as good or better than lines drawn manually) Map was plotted commercially using magnetic tape supplied by Texas Highway Dept

dedications, title blocks and all general notes plus all plans and profiles drawn by the plotter All this involved 35 plan sheets—yet my total time for drafting the plans was less than 24 hours," reports Quinn

That's one man-week to do a job that formerly required 125. Even with computers, how can a job possibly be speeded this much? (For example, others have told CE that computer mapping systems speed work by from three to 10 times.) Among the reasons why Development Consultants has speeded development mapping much more than this are the following:

- Where such map elements as street cross-sections are standardized, the computer program can be written so that, to draw all lines on the street, all you need input is the centerline of the street. This speeds drafting
- Development mapping is repetitive Maps using the same lines have to be made several times (1) boundary map of the total subdivision or development, (2) maps of portions of the total, as they are laid out into individual homesites, (3) maps used as the basis for earthwork construction of streets (and water, sewer and drainage systems), and by utility companies in laying electric, telephone and gas, and (4) a title map of each individual parcel (Purser's firm may or may not do this last)

Would you believe that a computer map system could cut map-drafting manhours for a 715-lot subdivision from 2½ man-years to one man-week? That's the report from Houston's Development Consultants.

In computer mapping, once the X and Y coordinates of a point are recorded on magnetic tape, 90% of the work in mapping that point is already done This is not true with conventional mapping—you must re-plot the same point each time you make a new map Thus, when you have more than one use for the geographic data, savings in time and money multiply.

2 *Cost of map-making*, in the case of Development Consultants' subdivision maps, is sharply lower than for maps made conventionally

To make the same number of maps that now cost Purser about *$1,000 a month* for the $30,000 Wang mapping system, would, he calculates, cost about *$10,000 a month in draftsman salaries* (To the computer-system cost, add the $3,000 a month salary cost of computer operating personnel)

Computer mapping—growing popularity

Computerized design and mapping is taking off in the civil engineering field. Among reasons for this, reports software supplier Hector Holguin of El Paso, Tex., are these:

- Hardware costs have plummeted During the past 10 years a number of civil engineering organizations have used such hardware as IBM 1130 computers for design and mapping work, at a cost on the order of $3,500 a month. Now, Holguin reports, thanks to far less costly minicomputers, plotters and other units, hardware system cost may soon drop as low as $700 a month for some applications
- For years the hardware suppliers were the leading innovators. Naturally, they invested most in hardware, leaving much of the program-writing task to users. Program development trailed hardware innovation. But in the past year or two or three, software innovations such as Holguins's (and by other companies mentioned in a box elsewhere in this roundup) have been another important cause of the fast growth of computerized mapmaking

Those 10 draftsmen would occupy perhaps 1,000 sq ft (93 m²) of office space, which would also have to be paid for.

Cost savings come also in drafting the alpha-numeric parts of Quinn's maps "We used to have to type dedications, then either fold the plat, which would make it look bad, or send it out on a reproduction material to have it photographed at a cost of perhaps $150 plus the time involved We spent a few thousand dollars sending plans and profile sheets out and having our title blocks professionally done with our company name Now I can take a blank piece of film and prepare a professional looking title block during spare time

"We save more time because the computer stores standard paragraphs, phrases and certifications We only type in data that is unique to the job at hand "

3 *Errors are reduced*, Quinn reports, when maps are made by computer The reason is this that "all information for a given project is stored on magnetic disks We never really 'touch' that data, we simply recall it and rework it when necessary

"When maps are made and calculations done manually, errors are occasionally made in looking up trigonometric functions, or in surveying, or on the drafting table in plotting the wrong grade or elevation Just to calculate the location of each point on a map involves four manual steps, which means four chances for error

"We are turning out work that becomes part of official records If we make errors, we could end up 'buying' a house We have a tremendous liability, and using manual methods we just cannot match the accuracy of the computer. I estimate that computer mapping cuts errors 90% "

4 *Employees are happier* with the computer system, Quinn reports Knowing they have an efficient, trailblazing computerized system, "field teams charge after their work a little more, and they take more pride in their work. Because our people are happier, we have fewer personnel problems with the computer mapping system "

The Wang system can handle peaks in work loads—and is inexpensive enough that the firm can afford to let it set idle during slack periods. The result is less hiring to staff up for work peaks, less laying off during slow periods "A week's overtime of the 10 draftsmen it replaces can equal the system's cost for a month "

Hardware and software

Hardware comes from Wang Laboratories, Lowell, Mass., and includes among key items of equipment minicomputer, CRT (television) display and keyboard, and flat-bed plotter that draws maps up to 31 x 54 in (790 x 1370 mm) All were made by Wang Laboratories, Lowell, Mass

Most of the computer programs for Development Consultants were written by Holguin Associates, El Paso, Tex

That firm's chief, Hector Holguin, reports that his software has been purchased by more than 300 firms, doubtless making him a national leader in marketing civil engineering programs Holguin's software business volume reportedly doubled from 1975 to 1976

In making survey calculations and maps, errors are occasionally made. Making maps by computer may cut these errors 90%.

Texas Highway Department's computerized mapping

Quite different in size and evolution from the mapping system described above is that of the Texas State Department of Highways and Public Transportation

Their system evolved from an in-house aerial mapping operation. While their system is big—because Texas is big—it is adaptable to even small mapping operations, report Tommie Howell and Frank Cooper.

The major items of cost and benefit are:

Annual Costs (5 year amortization)	
Photohead flatbed plotters (highest quality and cost)	$36,000
Digitizers (directly from stereoplotters, 8 of them)	$24,000
Computer (system uses computers required for other purposes)	0
Total cost	$60,000

Annual Benefits	
Eliminate scribe-coat material used in manual mapping	$ 24,000
16 fewer draftsmen (drafting staff reduced from 20 to 4)	$160,000
Total benefit	$184,000

Among factors in the cost benefit equation that Howell uncovered in his pre-purchase analysis

- The big savings come in the need for fewer map draftsmen Only four are needed, compared with 20 a few years ago These labor savings result from the plotter's higher speed than manual mapping Texas uses two kinds of plotters—drum, which is 50 times as fast as manual mapping but sacrifices something in line quality, and the photo-head flatbed, which is 12 times as fast as manual and gives even better quality of linework
- Savings in office space formerly occupied by the draftsmen who are no longer required These 16 people required more than 1,200 sq ft (112 m²), which cost several thousand dollars a year in office space cost

Computer mapping has a benefit: cost ratio of 3:1 for its big operation, says the Texas Highway Dept. Big saving: labor.

- In pre-computer days as today, maps are prepared from photo-pairs in the stereoplotter. In the old days, the stereoplotter operator viewed the projected 3-D image of the aerial photo on a tracing table, and drew the desired line.

Today, instead of drawing lines, the operator moves his digitizer over the projected image, and this automatically transfers to mag tape the X-Y coordinate values of all photo features he traces

The information on mag tape is processed by computer, creating a geographic data base. From this data base, any desired map can be plotted automatically by the computer.

Interactive graphics adopted for mapping and design

Texas views carry weight, if for no other reason than that the state highway department developed the "Roadway Design System" (computerized highway design and mapping program) funded by the Federal Highway Administration The RDS has been furnished to over 50 agencies by FHWA; an RDS user group was recently formed Eleven state transportation agencies have adopted the system.

Who can do such a job for my company?

Talk to more than one system packager. (If you're new to computer mapping, don't try to assemble your own system.)

Be cautious if you're to be a software-supplier's first customer in your type of application. There is a "learning curve" to application of any new technology—in the early months and years of application, mistakes are inevitably made.

Get the best advice you can One good way to judge quality of a consultant or supplier is to talk to those who are using the product or service you're thinking of buying Pin them down as to reliability, cost, speed, quality of line work, etc.

To make or maintain your own maps

Several companies are marketing computer mapping systems, among them (in alphabetical order).

Applicon
154 Middlesex Turnpike
Burlington, Mass 01803

Auto-Trol Corp
5650 N Pecos St.
Denver, Colo.

Calma
707 Kifer Road
Sunnyvale, Calif 94086

Computervision
201 Burlington Rd.
Bedford, Mass. 01730

H Dell Foster
P O Box 32581
San Antonio, Tex. 78216

Information Displays Inc.
150 Clearbrook Rd.
Elmsford, N Y. 10523

M&S Computing
P O Box 5183
Huntsville, Ala. 35805

Synercom Technology
6300 Hillcroft
Houston, Tex 77036

Wang Laboratories
1 Industrial Ave.
Lowell, Mass. 01851

Have a service bureau do the job?

Another option is to have your map data base (and maps) made and/or maintained by a computer service bureau. Such services are offered by, among others:

Aero Service Corp
8100 West Park Drive
Houston, Tex 77063

Autographic Mapping
1200 Milan
Houston, Tex

Auto-Trol
(see listing above)

Computer Graphics Co
899 Logan St.
Denver, Colo.

McDonnell-Douglas Automation
Box 516
St. Louis, Mo. 63166

Utility Data Corp
1111 Fannin
Houston, Tex.

That work was done six years ago. Now being developed at Texas, also under FHWA contract, is an "interactive graphics" system. This will further speed mapping (see photo) In addition, it will permit designing roads and structures directly on the CRT. (See CE cover story, 2/76.)

The new Interactive Graphics System was put together for Texas, and software developed, by M&S Computing, Huntsville, Ala. Hardware includes, at each of two stations, two Tektronics CRT's, and a Digital Equipment Corp. minicomputer (PDP 1135). The system has no separate plotter—output is on tape, which is simply transferred to the state's existing computer-plotter map system described above

(For more information on the Texas system, and the interactive graphics modification, write Tommie Howell, Division of Automation, Texas Department of Highways and Public Transportation; Austin, Tex 78731.)

Engineering education in computer graphics

How widely taught in U.S. engineering schools is computer graphics?

ASCE's Computer Graphics Committee wondered, so conducted a mail survey of 200 CE department chairmen, with these results:

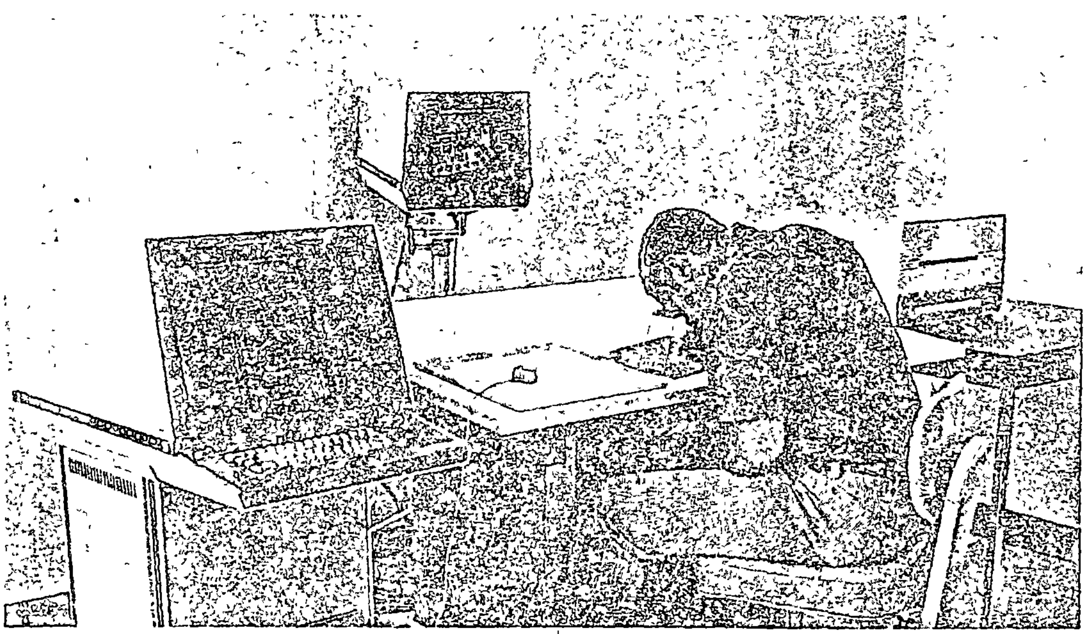
- Of 105 respondents, 13 use graphics in undergraduate education. But virtually all courses on the subject are elective only. And not one school teaches *interactive* graphics (cover story, CE 2/76).

In other words, the old drafting courses have been phased out at most schools, but the new computer graphics isn't yet in.

- How likely is graphics to be taught in the future; and how soon? About 20% of respondents feel more graphics education is needed, and about 30% see unusual design potential in it.
- Why aren't more schools teaching computer graphics? In most cases, because of unavailability of hardware and/or faculty familiar with it.

(Computer graphics has advantages and disadvantages, notes Professor W.R. Spillers, Rensselaer Polytechnic Institute, Troy, N.Y. Advantage: the student can quickly see many cases on the CRT screen. Disadvantage: the computer does the step-by-step computation, so the student may lose some comprehension of design details or steps.)

For a copy of the full study by the ASCE committee, contact its chairman, Barry Flachsbart, McAuto, P.O. Box 516, St. Louis, Mo. 63166.



Texas Highway Department's computer mapping system is being updated with this "interactive graphics" system This work station has two CRT's—small one shows map of large area larger screen zooms in to show small area in great detail Large table is digitizing table, for inputting map lines to computer Smaller board is tablet with many mapmaking instructions written on it's surface (and pre-programmed into computer), tablet speeds mapmaking At right is printer that makes paper print of map shown on CRT Operator is using magnifying glass to study print of map shown on CRT, will make corrections if necessary

Do You Know That

. . . someday homeowners might be taking their dogs out for a swim? Well, not quite, but homeowners of the future may be investing in a nice little undersea habitat with a front yard of well-trimmed sea grass—at least if the research being conducted by Wolf Hibertz, associate professor of architecture at the University of Texas, is successful For several years he has been working on the development of a technology for producing materials to build undersea habitats for humans an architecture which might be located under water, float on the surface or be transported onto land His process involves the electrical deposition of minerals on the surface of submerged metal forms—such as wire mesh An electrical current passed through the metal sets off a chemical reaction which causes the minerals to adhere to the wire and eventually cover the mesh completely

. . . steel drums holding paint for highway surface markings could become trash, but instead hold it? Steel drum producers surveyed the disposal practices of a major user, the highway department of the 50 states which buy traffic paint in 55 gal (206 l) drums for road and highway surface markings A high response—33 of the 50 states—showed that the overwhelming usage of emptied drum is conversion by the departments to trash receptacles There are other secondary applications such as construction markers which are also efficient and conservation-minded But because more drums are emptied than needed respondents reported that the surplus is sold to drum reconditioners who refurbish them for use as shipping containers Drums in poor condition are sold to scrap dealers for recycling through the steel-making process

Comput. & Graphics, Vol. 1, pp. 361–374 Pergamon Press 1975 Printed in Great Britain

THE USE OF INTERACTIVE COMPUTER GRAPHICS IN THE CONFORMAL MAPPING AREA*

JOHN C. BRUCH, JR.
Department of Mechanical and Environmental Engineering, University of California, Santa Barbara, CA 93106, U.S.A.

Abstract—The conformal mapping area lends itself very readily and easily to the use of interactive computer graphics. This is demonstrated by investigating with an interactive computer graphics system (the University of California at Santa Barbara On-Line Computer System) the conformal transformations applied to fluid flow problems. A description of the design philosophy of the system is given along with a variety of flow examples.

INTRODUCTION

In the study of hydro- and aerodynamics[8, 13–15], flow visualization is an important tool for the complete understanding of fluid flowfields. Being able to visually see the streamlines enables one to study and analyze the flowfield with more exactness and confidence. The shape of the flowfield, the location of critical points, and other fluid dynamics quantities can be obtained from accurate flowfield patterns.

There are many ways of obtaining flow visualization of fluid flow patterns. Some of these are: qualitative blackboard sketches, graphical point-by-point pictures (one approach here would be to add streamlines of simple flow patterns), analogs, movies[11], and other experimental techniques such as the hydrogen bubble method[11] and smoke tunnels. However, some of these approaches are limited in the number and types of flowfields that can be shown. These approaches also yield little desirable information other than the shapes of the flowfields.

The method of flow visualization described herein uses conformal mapping techniques and the University of California at Santa Barbara (UCSB) On-Line Computer System[6, 7]. A description of the design philosphy of the system is given in the next section. Since this system has the capability of operating on vectors, it is ideally suited to perform complex operations, especially conformal transformations[5, 12]. This feature in addition to its versatility and speed makes the system unusually outstanding in presenting flow visualizations. It also has the capability of printing out on the oscilloscope the coördinates of any point it has displayed or any other value it has calculated.

During the past several years, the On-Line System (OLS) has been used extensively for computer-aided instruction in Engineering at UCSB[9]. Typical uses of the system have included instruction in complex variables[16], circuits and networks, control systems and hydrodynamics[4], among others. The ways in which the system has been employed to enhance the impact of the engineering curriculum are the on-line generation of displays for the lecture hall[4], laboratory sessions in a computerized classroom[17], and as a vehicle for the preparation of audio-visual aids[2, 3].

The paper is directed toward use of the UCSB OLS as a computer-aided research and design tool in the conformal mapping area with emphasis on applications to fluid flowfields. Herein, the OLS' capabilities and potential in the study of some classical flowfields without free surfaces is discussed. However, flow problems with free surfaces can also be handled very easily[1]. A later section demonstrates a variety of these classical two-dimensional irrotational flow patterns[8, 13–15] that can be mapped and displayed and the pertinent fluid dynamics information obtainable from the displays.

THE COMPUTER SYSTEM

The UCSB On-Line System[6, 10] is an extended version of the so-called Culler–Fried system[6]. It is implemented on an IBM 360 Model 75 computer and employs a storage oscilloscope for output and a dual keyboard for input, as illustrated in Fig. 1. The duality of the keyboard arises from the operation–operand orientation of the On-Line System. The system design philosophy provided for the isolation of the controlling processes (the operators) and their effects (the transformation of the operands). Thus the system structure consists of a set of operators, a set of appropriate operands on which the operations are defined, and a means of sequencing the application of the operators to the appropriate operands. The buttons of the lower keyboard address operands, which currently can consist of single numbers, of lists of numbers (vectors), or of arrays (two-dimensional lists of numbers) all composed of either real or complex elements. These operands provide temporary storage for intermediate results, and generally play the role of variables in an algebraic expression. The variable structure for a given problem is specified by the top row of buttons on the upper keyboard (see Fig. 1). The first six buttons (I–VI) are operator category keys and specify the form of both operators and operands. Hence they essentially act as shift keys with I specifying a context of single numbers, II specifying vectors, and III arrays. The next two keys (REAL, CMPLX) specify whether the operators (and their operands) are defined as real or complex. For example, if II CMPLX is specified,

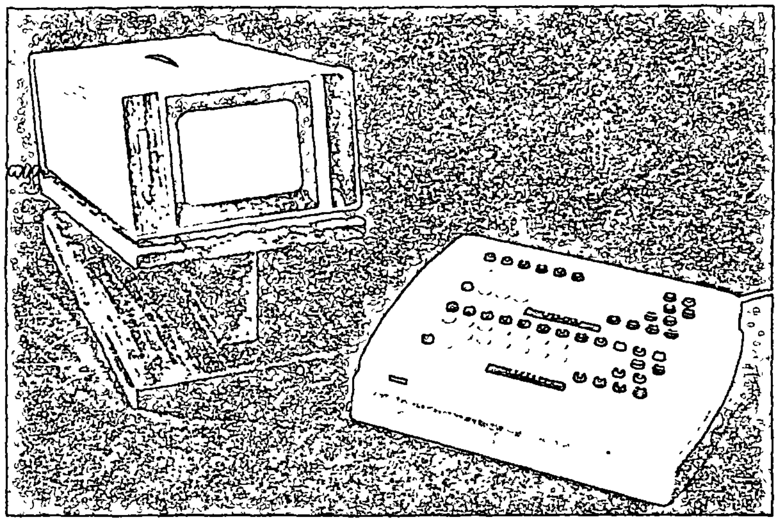


Fig 1 An on-line console.

the operators will be transformations on complex vectors (of 51 elements, unless otherwise specified). Since this paper deals with complex (conformal) transformations, the appropriate operations are primarily those of II CMPLX and the nature of the operators on that level will be described.

The available operations, which can be inferred from an observation of the keys shown in Fig. 1, include arithmetic operations such as addition, subtraction, multiplication, division, square root and conjugation, elementary function operations such as sine, cosine, logarithm, exponentiation and arctangent, data manipulation operations such as load and store; input/output operations of enter and display; and program control operations such as reset, test and repeat Moreover, the system provides for the combination (concatenation) of operators to make compound operators. By means of the LIST operator, it is possible to specify a sequence of button pushes (representing a particular sequence of operations upon specified operands), and store the sequence away for future use Six levels of user operator storage are provided, which are called by keying USER and then one of the buttons I–VI Thus, for example, a given list of operations could be stored 'under' the name USER I EXP The subsequent pushing of those three buttons would then call up the original operator sequence This technique provides an essential program writing capability which was used extensively in the generation of the illustrative displays described in this paper. Appendix contains listings of a number of the programs employed in this study, together with a discussion of the mathematical operations being performed, and provides an idea of the power of the system and of the ease with which it can be used.

In applying the system to the mathematics of fluid flow, first note that a complex vector can be considered to be a parametric representation of a complex function Thus by generating a vector with elements whose real parts increase by a fixed amount from element to element and whose imaginary parts are all equal, a (discrete) parametric representation of a straight line can be produced parallel to the x-axis and at a specified distance from it. Facility for the generation of such a vector (along with many other convenient parametric forms such as the unit circle) is provided by means of an operator designated ID. Once generated, a function can be viewed by pushing DISPLAY RETURN which gives a straight-line segment representation of the function) or either DISPLAY· RETURN or DISPLAY··RETURN (which give pointwise displays of the function, the only difference being in the intensity and/or shape of the dots). The scale of the function is automatically adjusted so that the entire display appears on the face of the tube, and that scale can be displayed in numerical form by keying DISPLAY O RETURN. A component of the vector (point of the function) can be displayed numerically by indicating DISPLAY n RETURN, where n is the number of the desired component.

COMPUTER GENERATED FLOW EXAMPLES

First, a parametric representation of a straight line, parallel to the ϕ-axis and a specified distance from it can be produced by the generation of a vector with elements whose real parts increase by a fixed amount from element to element and whose imaginary parts are all equal A series of vectors can then be generated This can be considered the W-plane (Fig 2) or a uniform flow defined by

$$W = \phi + i\psi, \tag{1}$$

where W is the complex potential, ϕ is the velocity potential, and ψ is the streamfunction.

The W-plane (Fig 2) or uniform flow can be mapped into a flow at a wall angle θ by the transformation

$$W = Az^n, \tag{2}$$

where $n = \pi/\theta$, A is a real and positive number, $z = x + iy$ is a complex number, and x, y are Cartesian coordinates. Figures 3 and 4 are examples of using equation (2) for different wall angles.

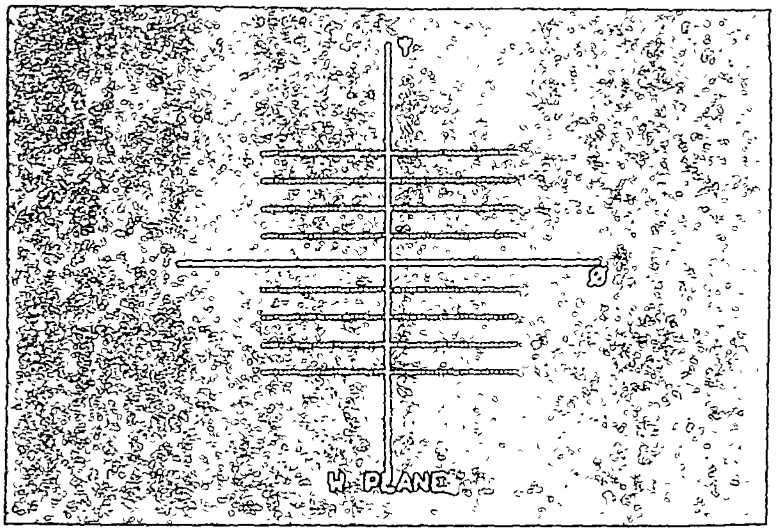


Fig. 2. W-plane Streamlines representing uniform flow from left to right.

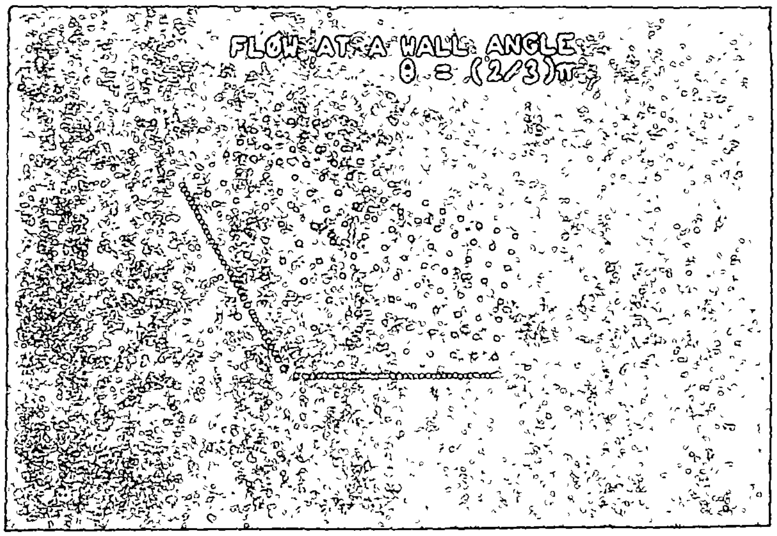


Fig. 3. Flow at a wall angle $\theta=(2/3)\pi$, $A = 1.0$.

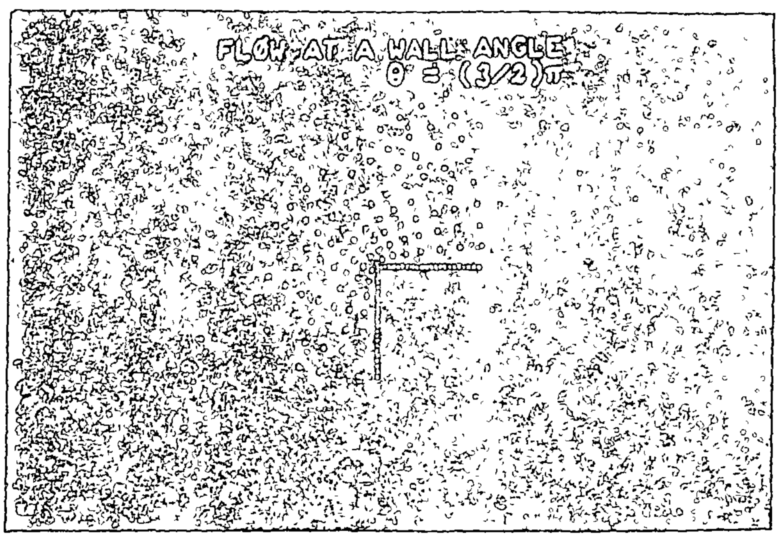


Fig 4 Flow at a wall angle $\theta = (3/2)\pi$, $A = 1.0$.

More complicated flows can be obtained by combining the complex potentials of simple flow patterns. For example, the combination of a doublet and uniform flow which yields the following equation

$$W = U_\infty\left(z + \frac{a^2}{z}\right) \tag{3}$$

where U_∞ is the velocity of the uniform flow and a is the radius of the circular cylinder or circle, transforms the W-plane (Fig 2) into a steady pattern of irrotational flow from left to right past a circular cylinder or circle with radius a (Fig. 5). (Note: in Fig. 5 as well as in all the following displays and equations in which U_∞ and a are used, they are set equal to 1.0 ft/s and 1.0 ft, respectively). If a uniform velocity of $-U_\infty$ is superimposed on this steady pattern, the result is the unsteady flow pattern caused by a cylinder moving to the left through a stationary fluid (Fig. 6). This would be the pattern seen by an observer at rest.

Adding a clockwise vortex flow of strength K to equation (3) yields

$$W = U_\infty\left(z + \frac{a^2}{z}\right) + i\frac{K}{2\pi}\ln z \quad (K > 0), \tag{4}$$

where K is the circulation or the strength of the vortex. Equation (4) transforms the W-plane (Fig. 2) into the pattern of irrotational flow past a circular cylinder of radius a with a clockwise circulation. Figure 5 is a special case of equation (4) with $K = 0$.

Figures 7–9 are examples of irrotational flow past a cylinder for different values of a clockwise circulation. In Fig 7 since $K < 4\pi a U_\infty$, there are two stagnation points on the surface of the cylinder located at

$$\theta_1 = \sin^{-1}\frac{-K}{4\pi a U_\infty}, \tag{5}$$

where θ_1 is the argument of the stagnation point's location

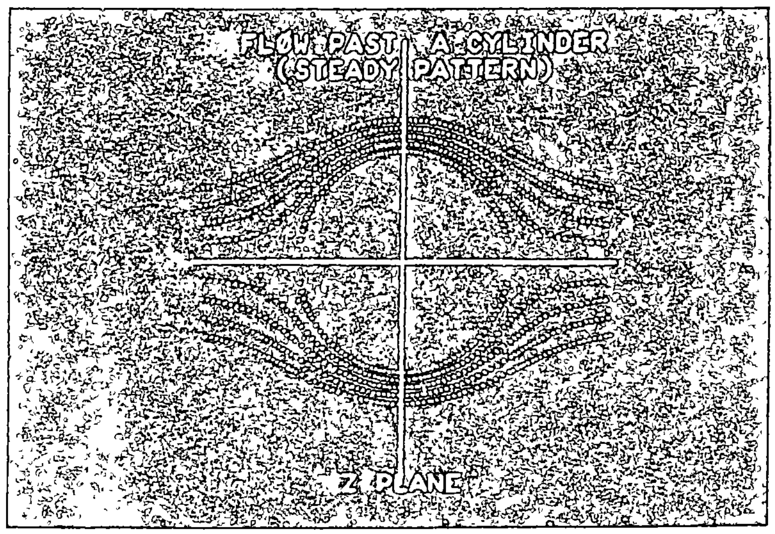


Fig 5 Flow past a cylinder (steady pattern) $U_\infty = 1.0$ ft/s, $a = 1.0$ ft

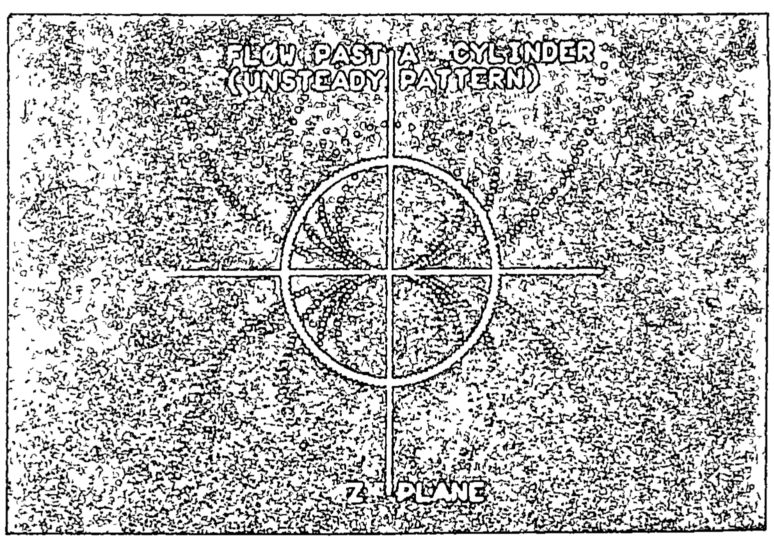


Fig 6 Flow past a cylinder (unsteady pattern) $U_\infty = 1.0$ ft/s, $a = 1.0$ ft

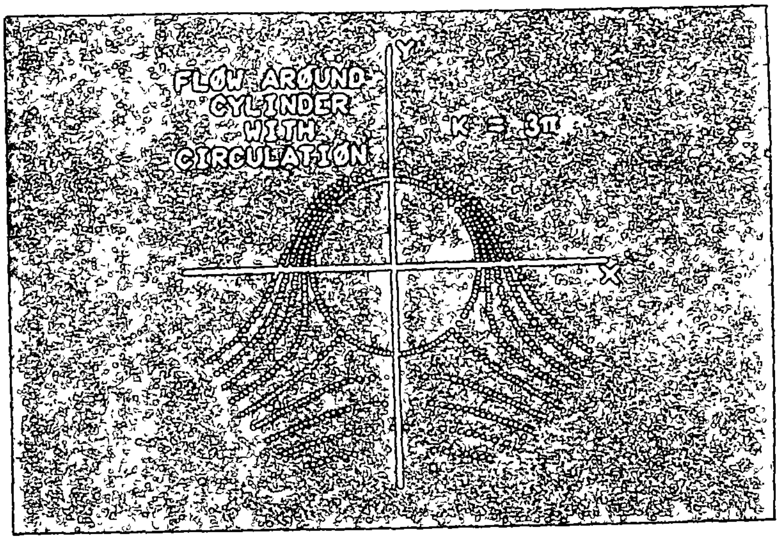


Fig 7. Flow around a cylinder with a clockwise circulation $U_\infty = 1\ 0$ ft/s, $a = 1\ 0$ ft, $K = 3\pi$ ft²/s.

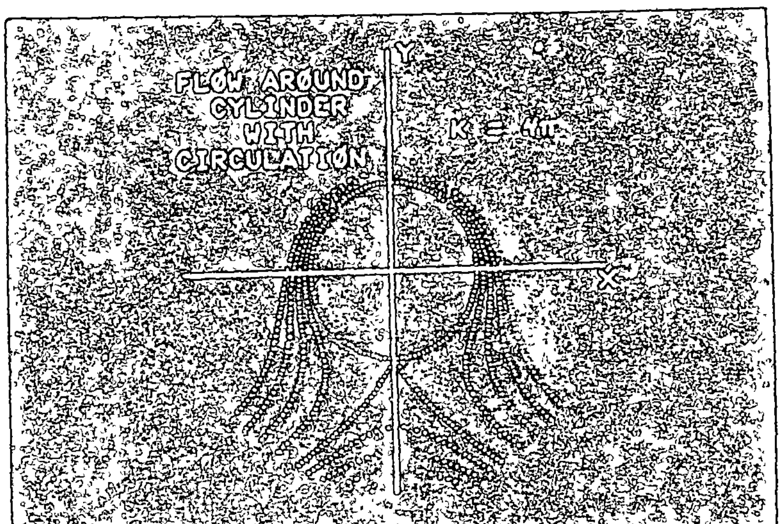


Fig 8 Flow around a cylinder with a clockwise circulation $U_\infty = 1\ 0$ ft/s, $a = 1{\cdot}0$ ft, $K = 4\pi$ ft²/s.

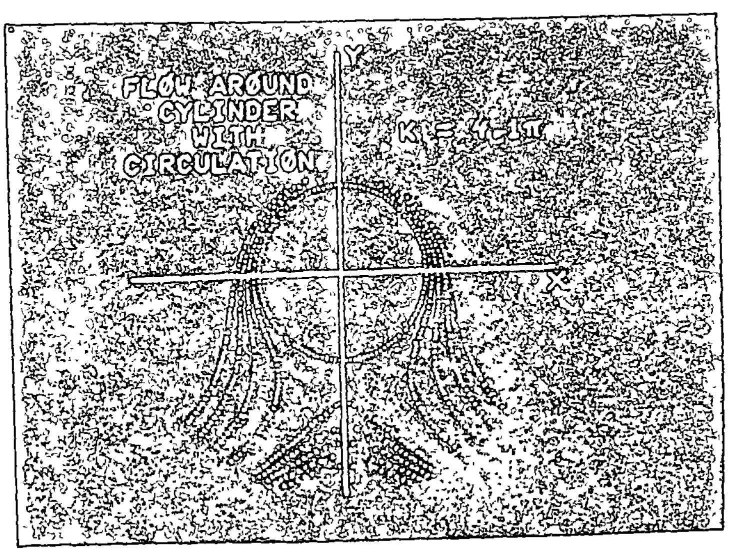


Fig. 9. Flow around a cylinder with a clockwise circulation. $U_\infty = 1\ 0$ ft/s, $a = 1\ 0$ ft, $K = 4\ 1\pi$ ft²/s

measured from the positive x-axis. For $K = 3\pi$ ft^2/s, $\theta_1 = -48°35{\cdot}4'$ and $-131°24{\cdot}6'$ The On-Line display gives the locations of the stagnation points to be $\theta_1 = -48°35{\cdot}3'$ and $-131°24{\cdot}7'$ Figure 8 demonstrates $K = 4\pi a U_\infty$ For this case there is only one stagnation point on the surface as shown in the figure and demonstrated mathematically in equation (5). Figure 9 illustrates what happens if $K > 4\pi a U_\infty$. There is no stagnation point on the cylinder. It has moved out into the flow to $z = -1{\cdot}4824i$ which is obtained directly from the display To calculate the point mathematically would require the solution of a transcendental equation.

Another example of adding simple flows and obtaining a more complicated flow pattern is the flow past a half body. Here a source flow of strength $Q_2 = 2\pi$ ft^3/s is added to a uniform flow so that

$$W = \frac{Q_2}{2\pi} \ln z + U_\infty z \qquad (6)$$

to yield Fig. 10. The location of the stagnation point from the display is $z = -1{\cdot}0$ ft Mathematically,

$$x_{stag} = -\frac{Q_2}{2\pi U_\infty}, \qquad (7)$$

where x_{stag} is the location of the half body's stagnation point, yields the location as $x_{stag} = -1{\cdot}0$ ft. Since the half body streamline ends as it does in Fig. 10, the maximum width cannot be obtained from the display However, the width to any point along the body as shown can be obtained. For example, the last point shown on the upper half of the body is located at $z = 2{\cdot}3562 + 2{\cdot}3562i$ which gives a half width equal to 2·3562 ft. Using the following equation

$$h = \frac{Q_2(\pi - \theta')}{2\pi U_\infty}, \qquad (8)$$

where θ' is the argument to the location of a point on the surface of the half body measured from the positive x-axis, gives the mathematical value at that point which is 2·3561 ft.

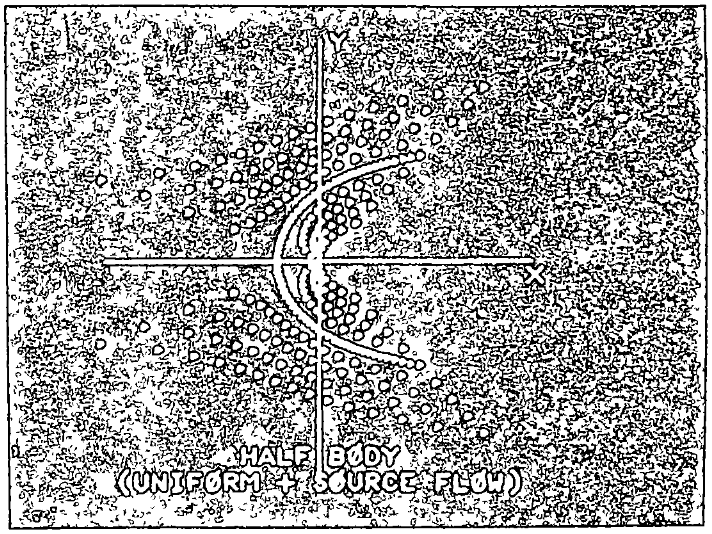


Fig 10. Flow past a half body. $U_\infty = 1{\cdot}0$ ft/s, $Q_2 = 2\pi$ ft^3/s.

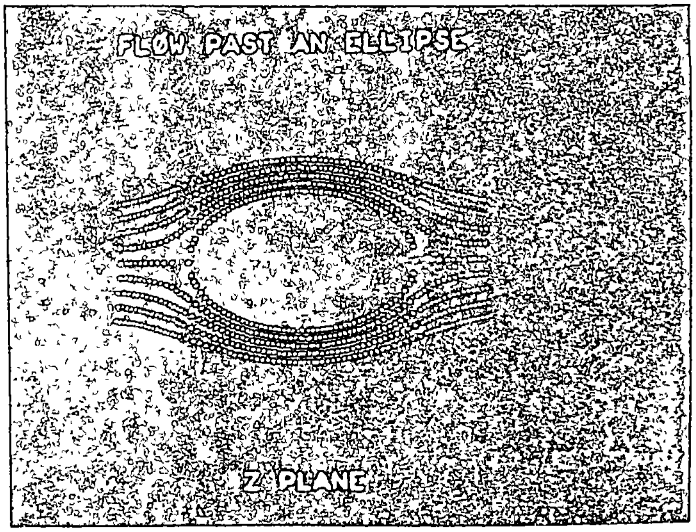


Fig 11 Flow past an ellipse $U_\infty = 1{\cdot}0$ ft/s, $a = 1{\cdot}0$ ft, $b = 0{\cdot}5$ ft

Figure 11 shows the flow past an ellipse This display was obtained from the steady pattern of flow past a cylinder by replacing z by z_1 in equation (3) and then using the transformation

$$z = z_1 + \frac{b^2}{z_1} \quad (b < a), \tag{9}$$

where b is the radius of a circle inside the a-circle (here $b = 0{\cdot}5$ ft). The a-circle is transformed into the ellipse Although not shown, the b-circle would transform into a line of length $4b$ inside the ellipse along the x-axis symmetric around the y-axis.

In all the above computer generated displays except for the uniform flow, it is noted that the dots in the figures are not uniformly spaced along every streamline, e.g. the streamlines passing through the stagnation points The spacing of dots is related to the velocity distribution along the streamlines The velocity is inversely proportional to the dot spacing The complex velocity relationship

$$\frac{dW}{dz} = u - iv, \tag{10}$$

where u is the horizontal velocity component and v is the vertical velocity component, shows how the dot spacing is related to the velocity. Taking the differences in the dot spacing of a streamline and dividing by these differences yields the velocity distribution of the streamline

Figure 12 is an example of applying the above procedure to the streamline along the upper boundary of flow past an ellipse. The velocity at the top of the ellipse from the display is 1·6 ft/s Using the following equation for the velocity at the same point on the ellipse

$$\frac{dW}{dz} = \frac{2U_\infty}{1 + (b^2/a^2)} \tag{11}$$

gives a value of 1·6 ft/s

Then, applying the following relationship to the above velocity profile,

$$\frac{\Delta p}{1/2 \rho U_\infty^2} = 1 - \left(\frac{V}{U_\infty}\right)^2, \tag{12}$$

where Δp is the pressure difference between a reference pressure, p_∞, where the velocity is U_∞, and the pressure at any point in the flowfield, p, where the velocity is V, and ρ is the density of the fluid, the pressure distribution of the upper streamline is obtained (Fig. 13).

If instead of equation (9) the following equation

$$z = z_1 + \frac{a^2 e^{2i\eta}}{z_1}, \tag{13}$$

is used, the steady flow past a cylinder is transformed into flow past a flat plate at some angle η. Figure 14 shows the flow past a flat plate at a 45° angle. Using equations (10) and (12) on the upper streamline of flow past the plate yields Figs. 15 and 16. The On-Line displayed velocities at the points $z = 0$ and $z = 0{\cdot}707 + 0{\cdot}707i$ are 0·707 ft/s and 1·112 ft/s, respectively. Checking these with the following equation

$$\frac{dW}{dz} = U_\infty \frac{\sin \Phi}{\sin(\Phi - \alpha)}, \tag{14a}$$

where

$$\Phi = \alpha + \cos^{-1}\frac{|z|}{2a}, \tag{14b}$$

and α is the angle of attack, yields velocity values of 0·707 ft/s and 1·115 ft/s for the same points

Finally, if the following set of transformations,

$$z_2 = z_1 e^{-i\alpha} \tag{15}$$

which rotates the pattern of flow through an angle of attack, $-\alpha$,

$$z_3 = z_2 + m e^{i\delta} \tag{16}$$

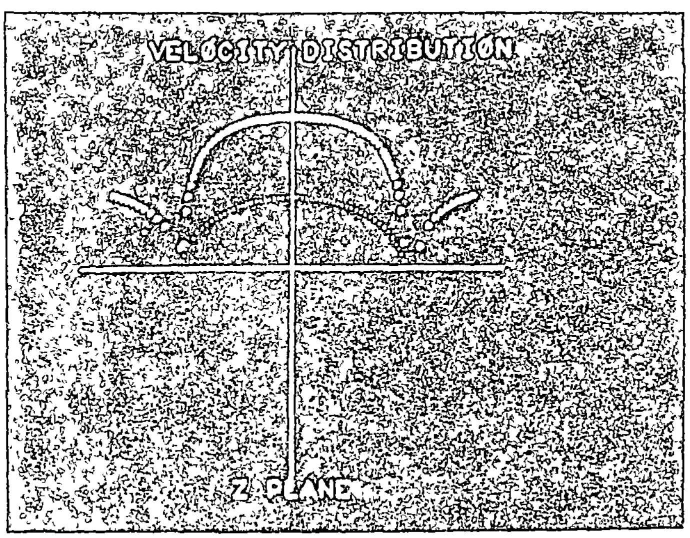


Fig. 12 Velocity profile of streamline along the upper boundary of flow past an ellipse. $U_\infty = 1{\cdot}0$ ft/s, $a = 1{\cdot}0$ ft, $b = 0{\cdot}5$ ft

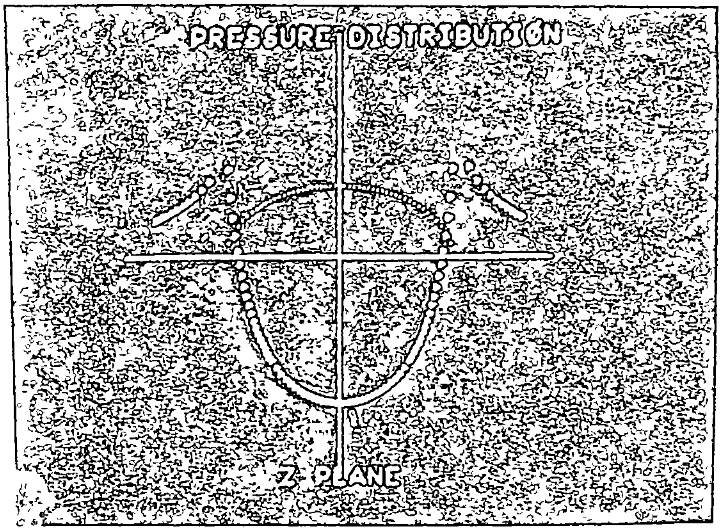


Fig 13. Pressure profile of streamline along the upper boundary of flow past an ellipse. $U_\infty = 1{\cdot}0$ ft/s, $a = 1{\cdot}0$ ft, $b = 0{\cdot}5$ ft.

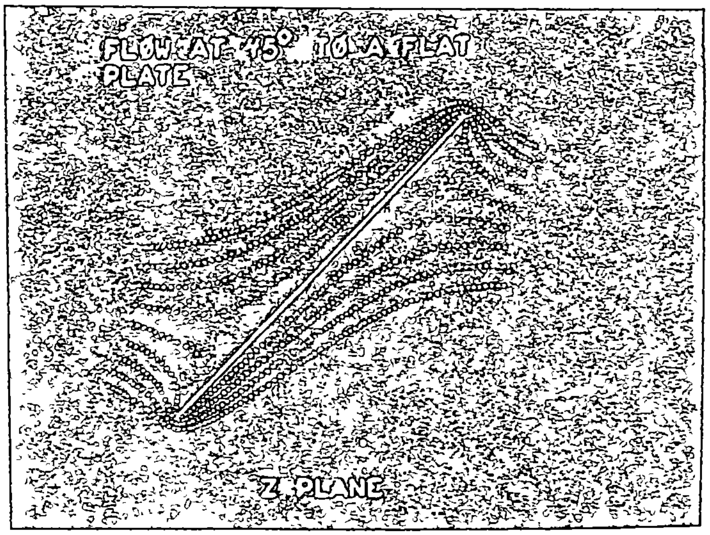


Fig 14 Flow past a flat plate. $U_\infty = 1{\cdot}0$ ft/s, $a = 1{\cdot}0$ ft, $\eta = 45°$.

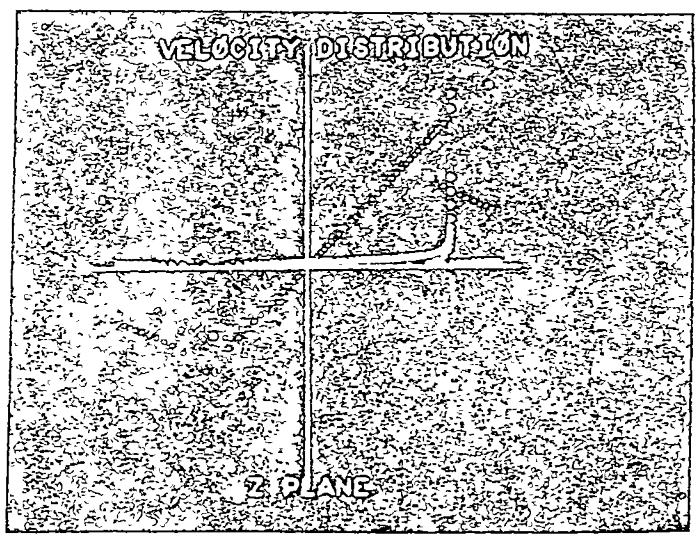


Fig 15. Velocity profile of streamline along the upper boundary of flow at 45° to a flat plate. $U_\infty = 1{\cdot}0$ ft/s, $a = 1{\cdot}0$ ft, $\eta = 45°$.

wh
th
wh

w
is
re
th
al
18
o
p
d
e

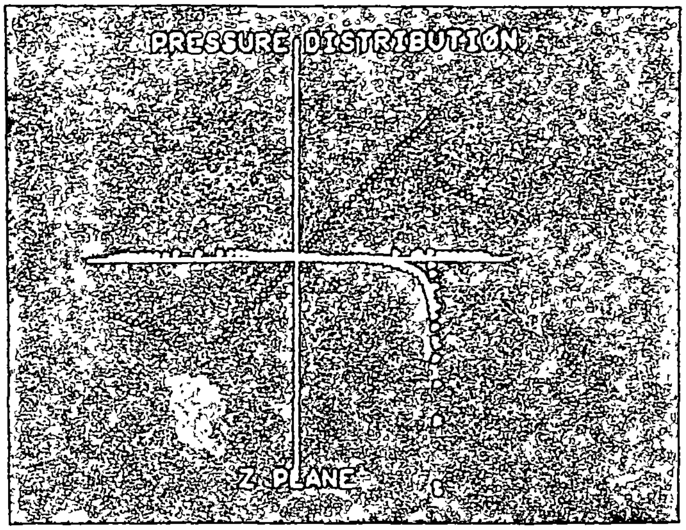


Fig. 16 Pressure profile of streamline along the upper boundary of flow at 45° to a flat plate $U_\infty = 1{\cdot}0$ ft/s, $a = 1{\cdot}0$ ft, $\eta = 45°$.

which removes the a-circle a distance m from the origin in the direction δ, (Note: if $\delta = 0$, the airfoil has no camber, while if $\delta \neq 0$, there is camber in the airfoil), and

$$z = z_3 + \frac{B^2}{z_3}, \qquad (17)$$

where $B = a\cos\beta - m\cos\delta$ and $\beta = \sin^{-1}(m\sin\delta/a)$, is used in place of equation (9), the flow past an airfoil results (Fig 17). Equations (10) and (12) are used to obtain the velocity and pressure distributions of the streamline along the upper boundary of flow past this airfoil Figures 18 and 19 are the corresponding displays As an example of a comparison of the velocity values (Fig 18), take the point $z = 0{\cdot}5649 + 0{\cdot}4044i$, where the value from the display is 1·28613 ft/s Using the following mathematical equation

$$\frac{dW}{dz} = U_\infty[2i\sin\Phi_1]\frac{e^{i(\alpha-\Phi_1)}}{1-(B^2/z_3^2)}, \qquad (18)$$

where Φ_1 is the argument of the location of the point in the z_1-plane, gives a value of 1·28613 ft/s for the velocity.

If in place of equation (17) the Karman–Trefftz transformation

$$\frac{z+NB}{z-NB} = \left(\frac{z_3+B}{z_3-B}\right)^N, \qquad (19)$$

where N is some constant a little less than 2, is used, the resulting flow is that of flow past an airfoil with a finite angle at the trailing edge (Fig 20) If $N = 2$, equation (19) reduces to equation (17).

In Figs 17 and 20 there was no circulation present in the flows If circulation is added and increased to move the rear stagnation point to the trailing edge, Figs. 21 and 22 are the resulting flowfields.

Some other physical flows that can be obtained from the W-plane (Fig 2) by the appropriate transformations are the flow past a cylindrical log (Fig. 23) obtained from

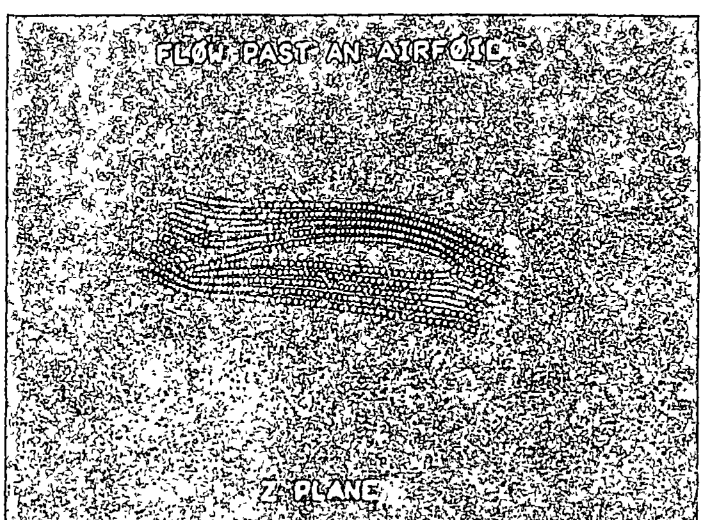


Fig 17. Flow past an airfoil with camber. $U_\infty = 1{\cdot}0$ ft/s, $a = 1{\cdot}0$ ft, $\alpha = 15°$, $m = 0{\cdot}15$ ft, $\delta = 54°$

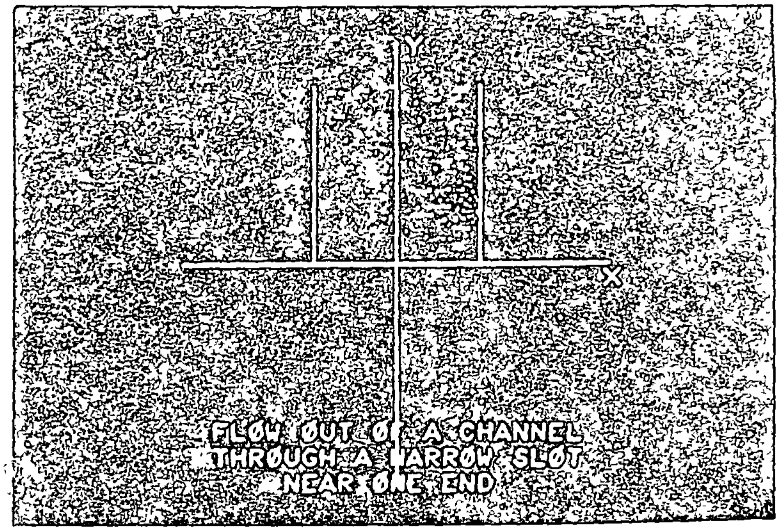


Fig 24. Flow out of a channel through a narrow slot near one end. $Q = \pi$ ft³/s, $l = \pi$ ft, $c = 0.962$ ft

the transformation

$$W = ir\pi U_\infty \cot \frac{ir\pi}{z}, \tag{20}$$

where r is the radius of the log, and the flow out of a channel through a narrow slot near one end (Fig. 24) obtained from

$$W = -\frac{Q}{\pi} \ln \left(\sin \frac{\pi}{l} z - \cosh \frac{\pi}{l} c \right), \tag{21}$$

where Q is the discharge through the slot, c is the height of the slot above the x-axis, and l is the width of the channel.

SUMMARY AND CONCLUSIONS

The above examples are just a few of the many fluid flowfields from the area of hydro- and aerodynamics that can be mapped, displayed, and analyzed on the OLS The intent of showing the above examples was to illustrate what the interactive graphics is capable of demonstrating on flowfields that have known solutions. Once having confidence in the approach, one can use it to verify a mathematical model or to construct a desired result by trial and error Also, the examples demonstrate the wide variety of flows that can be studied. This feature points out the system's versatility and broader usefulness over the other flow visualization techniques mentioned in the Introduction.

The response of the system on a typical operating day depends upon the number of users simultaneously operating on other OLS consoles tied into the UCSB computer However, the displays shown in the above figures took from a minimum of about 5 s to a maximum of 45 s, including the entire computational process The amount of actual computer time per display is at most a few hundredths of a second. Furthermore, the comparisons between the theoretical results and the OLS display results were exact in all the numerical examples Therefore, because of the OLS' rapid response, accuracy, and display capabilities, it is an excellent research and design tool.

The costs of using this system are not great in terms of time and money. The present operating cost of an On-Line console varies from \$4.00 to \$20.00 per console hr depending upon the priority the user selects.

The use of OLS is not restricted to the UCSB campus alone. Currently, there are a considerable number of remote users across the nation, for example, those stations which are tied into USCB via the ARPA network. However, this type of operation involves additional costs other than the charges for the use of the OLS itself. Here the main additional cost that is incurred is the communication line cost As an alternative to the remote use of the OLS, the system itself is exportable. The UCSB version, discussed in this paper, was designed to work on any IBM 360, Model 50 or above, with sufficient core and peripherals.

Acknowledgements—The author expresses his appreciation to Mr Edward C Keeley for writing the On-Line System computer programs for the examples and for taking the photographs for the figures

REFERENCES

1. J C Bruch, Jr, Free streamline theory and computer generated displays TR UCSB ME-71-1, University of California, Santa Barbara, CA (Jan 1971)
2. J. C Bruch, Jr & J A Howard, Flow visualization in hydrodynamics using On-Line computer-generated displays 16 mm movie, Learning Resources Motion Picture Production Section, University of California, Santa Barbara, CA (1969)
3. J C Bruch, Jr & J A Howard, Flow visualization in hydrodynamics using On-Line computer-generated displays Paper presented at the Pacific Southwest ASEE Annual Meeting, University of California, Davis, (29-30 Dec 1969)
4. J C Bruch, Jr & R C Wood, The teaching of hydrodynamics using computer-generated displays *Bull Mech Engng Educ* 9 (May 1970)
5. R V Churchill, *Complex Variables and Applications* 2nd edn McGraw-Hill, New York (1960)
6. *UCSB On-Line System Manual*, University of California, Santa Barbara, CA (Jan 1971)
7. G. J Culler, Mathematical laboratories a new power for the physical and social sciences *Proc Symp Interactive Systems Expl appl Math*, Assoc for Computing Machinery, Washington, D C (26–28 August 1967)

8. W. F. Durand, *Aerodynamic Theory* Dover, New York (1963)
9 J. A. Howard & R. C Wood, An On-Line system for computer assisted instruction, *J. Engng Educ* 61, 343–344 (Jan 1971).
10. W J. Karplus, Ed, *On-Line Computing*, pp 303–324 McGraw-Hill, New York (1968).
11. S J Kline, Film Principal, Flow visualization, Stanford University, 16 mm sound demonstration/experiment film distributed by Encyclopedia Britannica Educational Corp Chicago, Ill
12 H Kober, *Dictionary of Conformal Representations* Dover, New York (1957).
13 L. M. Milne-Thomson, *Theoretical Hydrodynamics*, 4th edn MacMillan, New York (1962)
14 J M Robertson, *Hydrodynamics in Theory and Application* Prentice-Hall, Englewood Cliffs (1965)
15. H R. Vallentine, *Applied Hydrodynamics*, 2nd edn Plenum Press, New York (1967)
16 R. C. Wood & J C Bruch, Jr, Teaching complex variables with an interactive computer system, *IEEE Trans. Education*, E-15, 73–80 (Feb. 1972).
17. R. C. Wood & J A. Howard, An interactive computer classroom, *Educ. Res. Methods J.* 2(4), 29–31 (June 1970).

APPENDIX I

Programming statements with explanations for the case of flow past an airfoil:

USER I EXP

```
USER I SQ USER ↑ ENTER ERASE USER I LS USER I MOD ERASE
USER I RS USER I MOD ERASE   USER I REFL USER I MOD ERASE
LOAD L Ø 57 3, 0 COS ○ M STORE N   LOAD L Ø 57·3 SIN ○ M Ø A
SQ ○ (−1) ⊕ 1 SQRT ⊙ A ⊖ N STORE B   USER I COS USER I MOD
```

Explanation Transforms W to z_1 (USER I SQ) DISPLAY, reverses flow (USER I LS) DISPLAY, transforms z_1 to z_2 (USER I RS) DISPLAY; transforms z_2 to z_3 (USER I REFL) DISPLAY, computes b from $b = a\cos\beta - m\cos\delta$, where $\cos\beta = \sqrt{(1-[m(\sin\delta)/a]^2)}$; transforms z_3 to z (USER I COS) DISPLAY.

USER I COS

```
LOAD Q USER I SIN Q   LOAD W USER I SIN W   LOAD E USER I SIN E
LOAD R USER I SIN R   LOAD T USER I SIN T   LOAD Y USER I SIN Y
LOAD U USER I SIN U LOAD I USER I SIN I LOAD O USER I SIN O
LOAD P USER I SIN P
```

Explanation Performs transformation from z_3 to z

USER I SIN

```
STORE Z LOAD B SQ Ø Z ⊕ Z STORE
```

Explanation. Enter vector z and perform operation $z + b^2/z$, then prepares to store it

USER I SQ

```
USER I Ø ERASE   LOAD Q USER I ○ Q   LOAD W USER I ○ W
LOAD E USER I ○ E   LOAD R USER I ○ R   LOAD T USER I ○ T
LOAD Y USER I ○ Y   LOAD U USER I ○ U   LOAD I USER I ○ I
LOAD O USER I ○ O   LOAD P USER I ○ P   LOAD A STORE B
```

Explanation Generates W lines (USER I Ø) DISPLAYS, transforms W to z_1 DISPLAYS)

USER I Ø

```
II REAL ID LOAD O CMPLX ○ 2 5, 0⊕0, 5↑STORE Q⊖0, 125
↑STORE W⊖0, 125↑STORE E⊖0, 125↑STORE R
⊖ 0, 125↑STORE T   ⊖ 0, 001↑STORE Y   ⊖ 0, 1249↑STORE U
⊖ 0, 125↑STORE I   ⊖ 0, .125↑STORE O   ⊖ 0, 125↑STORE P
USER I MOD
```

Explanation Generates W lines, 5 units in length equally spread from +0 5i to −0 5i centered on origin, 10 lines (2 on horizontal axis)

USER I ○

```
STORE X SQ STORE S LOAD A SQ ○ −4, 0 ⊕ S SQRT ⊖ X Ø −2, 0 DISPLAY
RETURN STORE
```

Explanation Enter a vector x and perform operation $(x - \sqrt{(x^2-4a^2)})/2$, displays resultant vector and prepares to store it

USER I ↑

```
II REAL ID LOAD O DISPLAY RETURN CMPLX ○ 0, 1 DISPLAY RETURN
```

Explanation Graphs vertical and horizontal axes

USER I LS

```
LOAD Q REFL NEG REFL STORE Q   LOAD W REFL NEG REFL STORE W
LOAD E REFL NEG REFL STORE E   LOAD R REFL NEG REFL STORE R
LOAD T REFL NEG REFL STORE T   LOAD Y REFL NEG REFL STORE Y
LOAD U REFL NEG REFL STORE U   LOAD I REFL NEG REFL STORE I
```

LOAD O REFL NEG REFL STORE O LOAD P REFL NEG REFL STORE P

Explanation Reverses direction of flow

USER I MOD

DISPLAY Q W E R T Y U I O P USER I↑ ENTER

Explanation Displays flow lines, puts in axes (USER ↑), halts program.

USER I RS

USER I INV Q STORE Q USER I INV W STORE W USER I INV E STORE E
USER I INV R STORE R USER I INV T STORE T USER I INV Y STORE Y
USER I INV U STORE U USER I INV I STORE I USER I INV O STORE O
USER I INV P STORE P

Explanation Transforms z_1 to z_2

USER I INV

LOAD J ○ 0, −1 ∅ 57 3, 0 EXP ○

Explanation. Performs operation $e^{-i\alpha/57.3}$.

USER I REFL

USER I DIFF Q STORE Q USER I DIFF W STORE W USER I DIFF E STORE E
USER I DIFF R STORE R USER I DIFF T STORE T USER I DIFF Y STORE Y
USER I DIFF U STORE U USER I DIFF I STORE I USER I DIFF O STORE O
USER I DIFF P STORE P

Explanation Transforms z_2 to z_3

USER I DIFF

LOAD L ⊙ 0, 1 ∅ 57·3, 0 EXP ○ M ⊕

Explanation Performs operation $m e^{i\beta/57.3}$.

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

MANEJO DE SISTEMAS DE INFORMACION GEOGRAFICA EN PLANEACION

SISTEMA PARA EL MANEJO DE MALLAS DE INFORMACION EN EL ANALISIS DE LOS USOS DEL SUELO Y RECURSOS NATURALES.

ARQ. ALEJANDRO VILLANUEVA EGAN

NOVIEMBRE DE 1977.

DIVISION DE ESTUDIOS SUPERIORES

FACULTAD DE INGENIERIA/UNAM.

SECCION DE PLANEACION.

MANEJO DE SISTEMAS DE INFORMACION GEOGRAFICA EN PLANEACION

ARQ. ALEJANDRO VILLANUEVA EGAN

IMGRID

SISTEMA PARA EL MANEJO DE MALLAS DE INFORMACION EN EL ANALISIS DE LOS USOS DEL SUELO Y LOS RECURSOS NATURALES.

LABORATORIO DE PLANEACION.

BANCO DE DATOS

Se pueden identificar 6 pasos para desarrollar cualquier *BANCO DE DATOS*.

Paso 1 *DETERMINACION DEL AREA DE ESTUDIO.*

El paso inicial consiste en encuadrar el area bajo estudio. Este encuadramiento puede establecerse de diferentes maneras, pero usualmente se determina por condiciones geográficas o mediante las jurisdicciones político administrativas.

Paso 2 *DETERMINACION DE LOS USOS DEL SUELO.*

La identificación de los usos del suelo que serán considerados en el proyecto puede lograrse mediante *la interacción interdisciplinaria de personas informadas en el estado actual y las potencialidades del área o región bajo estudio.* La definición de estos usos del suelo constituye la base para seleccionar la lista de variables y subvariables.

Paso 3 *DETERMINACION DE LA LISTA DE DATOS.*

En este paso se realiza la lista de los datos que estarán contenídos en el *BANCO*. Debe procurarse que los datos sean de la misma escala, exactitud y tiempo.*No existe una lista tipo que sirva para cualquier area.* Los datos pueden obtenerse a partir de la interpretación de fotografias aereas o de satélite, utilización de mapas de levantamiento geográfico, investigación de campo y consultas con dependencias federales o estatales.

Paso 4 *DETERMINACION DEL TAMANO DE LA CELULA.*

El tamaño de la célula, es decir, la unidad básica para el análisis espacial, depende de los siguientes factores:

a.-) Exactitud y tipo de los datos disponibles.

b.-) Propósito para el cual van a usarse los datos.

c.-) Tamaño del area de estudio.

d.-) Limitaciones en los recursos para codificar la información.

El tamaño de la célula debe permanecer constante en toda el área de estudio durante el proceso. (ver figura 1)

Paso 5 *CODIFICACION DE LOS DATOS.*

Este paso consiste en organizar los datos disponibles en un formato compatible con la computadora. El programa limita el número de variables a 50 y el de subvariables a 10, numerando estas del 0 al 9. Una tarjeta de cualquier archivo específico representa un renglón de datos en el mapa fuente; cada columna en la tarjeta representa una célula en el mapa. Si se necesitan mas de 80 columnas puede perforarse una segunda tarjeta por renglón. Los datos pueden ser codificados como:

a.-) Datos de punto (vgr. una cascada, un pozo etc.)

b.-) Porcentaje de la célula con una actividad determinada.

c.-) Tipo predominante de uso del suelo.

d.-) Datos de linea (vgr. una carretera, un rio etc.)

La codificacion del contorno debe realizarse especificando las fronteras de este en formato (3I5), dando en el primer campo el número de renglones que tienen el mismo formato y en los dos restantes el desfasamiento en ambos lados del area de estudio.

Paso 6 *TRANSFERIR LOS DATOS A LA COMPUTADORA.*

La forma usual de proporcionar los datos a la computadora es mediante tarjetas perforadas cuyo contenido es guardado en cintas o discos magnéticos para su posterior utilización en los *MODELOS*, este paso se realiza utilizando el programa *ARCHIVOS/IMGRID*, el cual es preparado como se indica en la figura 2.

Paso 7 *MAPEAR EL BANCO DE DATOS.*

Esto es realizado mediante la utilización del programa *MAPAS/IMGRID*, cuya descripción es proporcionada en el capítulo VI.

Una vez impresos, revisados y corregidos los mapas es posible pasar a la siguiente etapa del proceso en *IMGRID*.

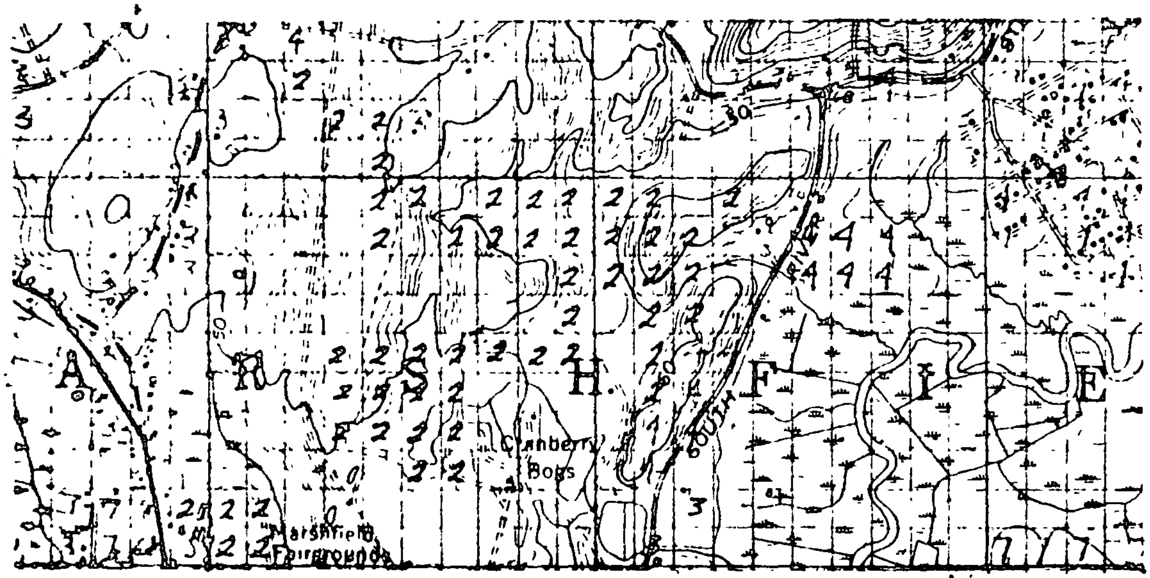
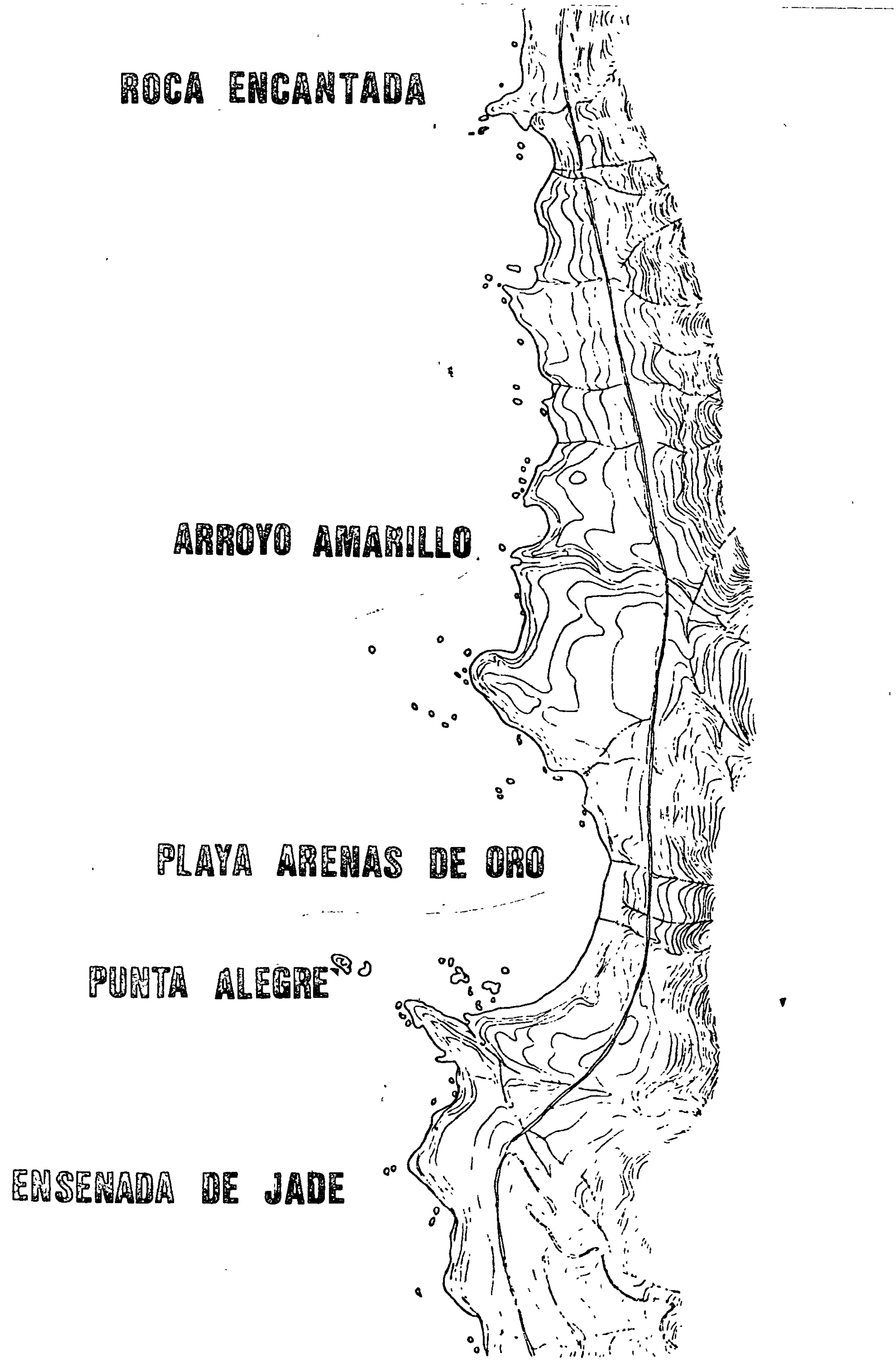


Figura 1. Ejemplo de retícula sobre mapa base.


ROCA ENCANTADA
ARROYO AMARILLO
PLAYA ARENAS DE ORO
PUNTA ALEGRE
ENSENADA DE JADE

VALLE DEL PACIFICO

ARCHIVOS

```
RUN ARCHIVOS

NUMERO DE RENGLONES   52   NUMERO DE COLUMNAS   16
    VARIABLE DATO NUMERO  1
         EL ARCHIVO DE DATOS ES LAEP/DATA 1
    VARIABLE DATO NUMERO  2
         EL ARCHIVO DE DATOS ES LAEP/DATA 2
    VARIABLE DATO NUMERO  3
         EL ARCHIVO DE DATOS ES LAEP/DATA 3
    VARIABLE DATO NUMERO  4
         EL ARCHIVO DE DATOS ES LAEP/DATA 4
    VARIABLE DATO NUMERO  5
         EL ARCHIVO DE DATOS ES LAEP/DATA 5
    VARIABLE DATO NUMERO  6
         EL ARCHIVO DE DATOS ES LAEP/DATA 6
    VARIABLE DATO NUMERO  7
         EL ARCHIVO DE DATOS ES LAEP/DATA 7
    VARIABLE DATO NUMERO  8
         EL ARCHIVO DE DATOS ES LAEP/DATA 8
    VARIABLE DATO NUMERO  9
         EL ARCHIVO DE DATOS ES LAEP/DATA 9
    VARIABLE DATO NUMERO 10
         EL ARCHIVO DE DATOS ES LAEP/DATA10

 TERMINADO
```

MAPAS

```
RUN MAPAS


CONTORNO IRREGULAR
==================


TITULO DEL MAPA
===============
MAPA DE LA VARIABLE 1 DEL INVENTARIO DE DATOS    USOS DEL SUELO ACTUALES
AREA DE ESTUDIO    VALLE PACIFICO
LABORATORIO DE PLANEACION URBANA


OPCIONES USADAS PARA ESTE MAPA
==============================

 1     EL TAMANO DE LA MALLA ES  52 RENGLONES Y  16 COLUMNAS
       EL TAMANO DE LA CELULA ES   1 CARACTERES EN SENTIDO VERTICAL
 7     LOS SIMBOLOS SON
        ...◇X000000123456789        1 CARACTERES EN SENTIDO HORIZONTAL
          .◇X0000A
              /X*X
                ◇V

10     EL TEXTO DEL MAPA ES         ===========
                  SUBVARIABLES DATO

                        0 = NO HAY DATOS
                        1 = CAMPAMENTOS
                        2 = USO DIURNO
                        3 = ADMINISTRACION DE SERVICIO FORESTAL
                        4 = ESCUELA
                        6 = PASTO
                        7 = CARRETERAS
                        8 = RESERVA FORESTAL
                        9 = OCEANO

                  EL TAMANO DE LA CELULA DE LA MALLA ES DE 2.5 ACRES
13    LA NUMERACION DE LA MALLA COMIENZA EN      1     52
14    SE SUPONE QUE LOS DATOS ESTAN PRE-ESCALADOS
```

TITULO DEL MAPA

MAPA DE LA VARIABLE 2 DEL INVENTARIO DE DATOS   PORCIENTO DE PENDIENTE

AREA DE ESTUDIO   VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

OPCIONES USADAS PARA ESTE MAPA

10   EL TEXTO DEL MAPA ES
SUBVARIABLES DATO

0 = NO HAY DATOS
1 = 100% AGUA
3 = 0-9%
5 = 10-15%
7 = 16-25%
8 = 25%+
9 = OCEANO

EL TAMANO DE LA CELULA DE LA MALLA ES DE 2.5 ACRES

TITULO DEL MAPA

MAPA DE LA VARIABLE 3 DEL INVENTARIO DE DATOS ORIENTACION PENDIENTE

AREA DE ESTUDIO   VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

OPCIONES USADAS PARA ESTE MAPA

10   EL TEXTO DEL MAPA ES
SUBVARIABLES DATO

0 = NO HAY DATOS
1 = LLANURA
2 = NORTE
3 = NOROESTE
4 = NORESTE
5 = SUR
6 = SUROESTE
7 = SURESTE
8 = OESTE
9 = OCEANO

EL TAMANO DE LA CELULA DE LA MALLA ES DE 2.5 ACRES

TITULO DEL MAPA
==================

MAPA DE LA VARIABLE 4 DEL INVENTARIO DE DATOS VEGETACION POR ZONA

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

OPCIONES USADAS PARA ESTE MAPA
==============================

10 EL TEXTO DEL MAPA ES ===========

SUBVARIABLES DATO

0 = NO HAY DATOS
2 = ZONA COSTERA
3 = SUELO PASTADO
4 = VEGETACION CHAPARRAL
5 = ZONA DE BOSQUE
9 = OCEANO

EL TAMANO DE LA CELULA DE LA MALLA ES DE 2.5 ACRES

TITULO DEL MAPA
==================

MAPA DE LA VARIABLE 5 DEL INVENTARIO DE DATOS DENSIDAD DE ARBOLES

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

OPCIONES USADAS PARA ESTE MAPA
==============================

10 EL TEXTO DEL MAPA ES ===========

SUBVARIABLES DATO

0 = NO HAY DATOS
2 = NADA
3 = 1-25%
4 = 26-50%
5 = 51-75%
7 = 76-100%
9 = OCEANO

EL TAMANO DE LA CELULA DE LA MALLA ES DE 2.5 ACRES

TITULO DEL MAPA

MAPA DE LA VARIABLE 6 DEL INVENTARIO DE DATOS ZONAS LLANAS

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

OPCIONES USADAS PARA ESTE MAPA

10 EL TEXTO DEL MAPA ES SUBVARIABLES DATO

0 = NO HAY DATOS
2 = NO ZONAS LLANAS
3 = LLANOS SOBRE LA COSTA > DE 2000I
4 = LLANOS SOBRE LA COSTA < DE 2000I
5 = LLANOS TERRESTRES < DE 200I
7 = LLANOS TERRESTRES > DE 200I
9 = OCEANO

EL TAMANO DE LA CELULA DE LA MALLA ES DE 2.5 ACRES

TITULO DEL MAPA

MAPA DE LA VARIABLE 7 DEL INVENTARIO DE DATOS ACCIDENTES GEOLOGICOS

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

OPCIONES USADAS PARA ESTE MAPA

10 EL TEXTO DEL MAPA ES SUBVARIABLES DATO

0 = NO HAY DATOS
2 = AREA ESTABLE
5 = DESLIZAMIENTOS ACTIVOS NO NATURALES
7 = DESLIZAMIENTOS ACTIVOS DE TIERRA
9 = OCEANO

EL TAMANO DE LA CELULA DE LA MALLA ES DE 2.5 ACRES

TITULO DEL MAPA

MAPA DE LA VARIABLE 8 DEL INVENTARIO DE DATOS SUELOS

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

OPCIONES USADAS PARA ESTE MAPA

10 EL TEXTO DEL MAPA ES
SUBVARIABLES DATO

0 = NO HAY DATOS
1 = NGFG2
2 = ST
3 = LHBC
4 = CHE
5 = RHGH
6 = GZF
7 = GRGH
8 = GZE
9 = OCEANO

EL TAMANO DE LA CELULA DE LA MALLA ES DE 2.5 ACRES

TITULO DEL MAPA

MAPA DE LA VARIABLE 9 DEL INVENTARIO DE DATOS PROXIMIDAD AL AGUA

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

OPCIONES USADAS PARA ESTE MAPA

10 EL TEXTO DEL MAPA ES
SUBVARIABLES DATO

0 = NO HAY DATOS
2 = CORRIENTES DENTRO DE LA CELULA
3 = AGUA EN LA CELULAS ADYACENTES
4 = AGUA A UNA CELULA DE PROXIMIDAD
5 = AGUA A DOS CELULAS DE PROXIMIDAD
6 = ZONA DE MAREA
9 = OCEANO

EL TAMANO DE LA CELULA DE LA MALLA ES DE 2.5 ACRES

MAPA DE LA VARIABLE 1 DEL INVENTARIO DE DATOS USOS DEL SUELO ACTUALES

AREA DE ESTUDIO VALLE PACIFICO LABORATORIO DE PLANEACION URBANA

SUBVARIABLES DATO

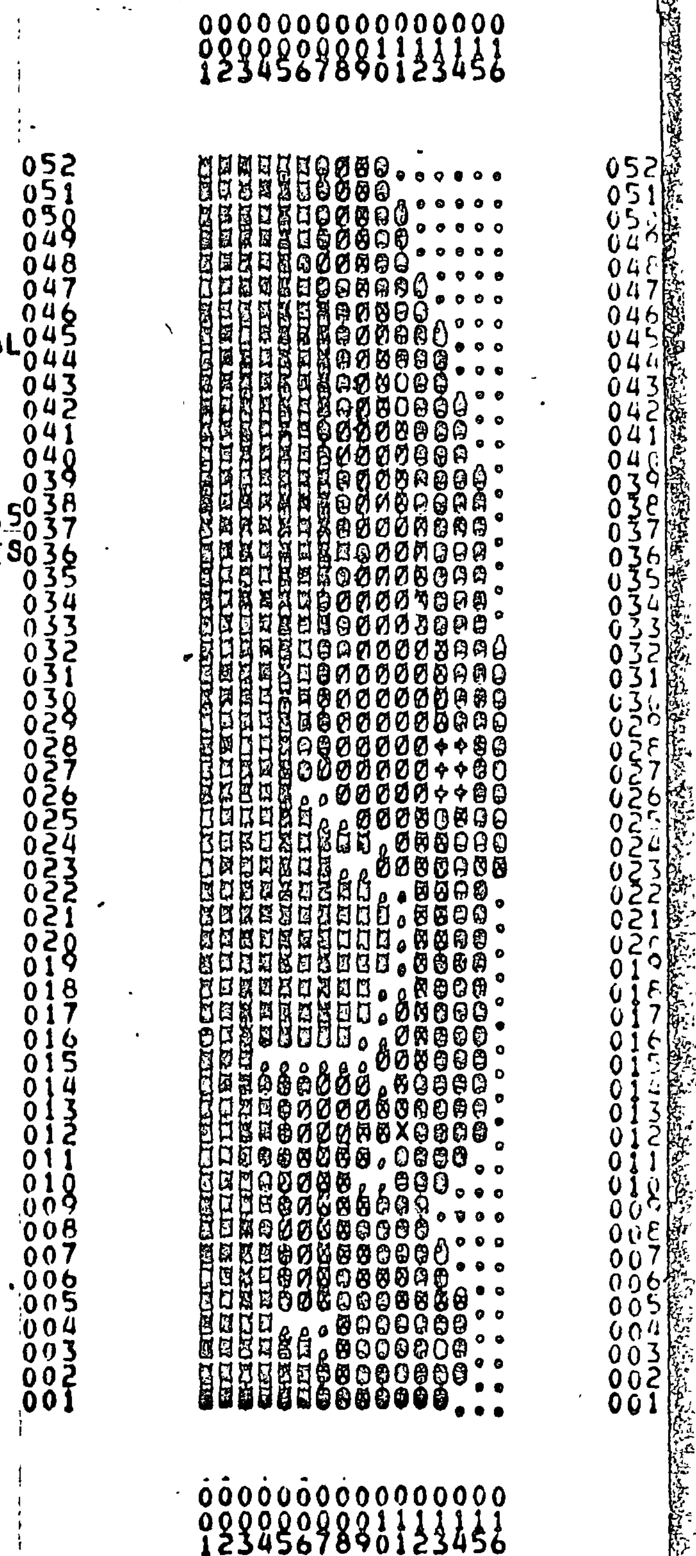
0 = NO HAY DATOS
1 = CAMPAMENTOS
2 = USO DIURNO
3 = ADMINISTRACION DE SERVICIO FORESTAL
4 = ESCUELA
6 = PASTO
7 = CARRETERAS
8 = RESERVA FORESTAL
9 = OCEANO

EL TAMANO DE LA CELULA DE LA MALLA ES DE 2.5 ACRES

NIVELES	0	1	2
SIMBOLOS	0	1	2
FRECUENCIA	98	3	28

NIVELES	3	4	5	6
SIMBOLOS	3	4	5	6
FRECUENCIA	6	1	2	98

NIVELES	7	8	9
SIMBOLOS	7	8	9
FRECUENCIA	76	203	317



MAPA DE LA VARIABLE 2 DEL INVENTARIO DE DATOS PORCIENTO DE PENDIENTE

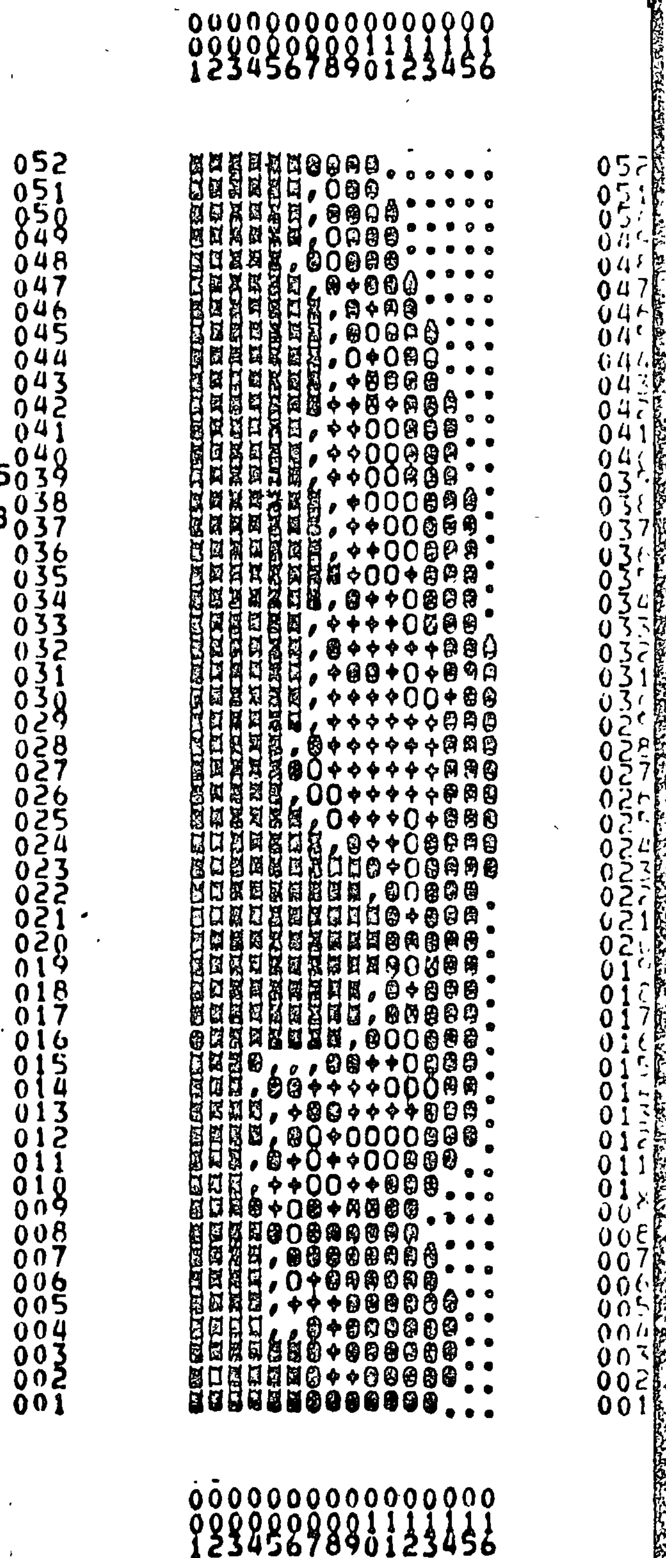
AREA DE ESTUDIO VALLE DEL PACIFICO LABORATORIO DE PLANEACION URBANA

SUBVARIABLES DATO

0 = NO HAY DATOS
1 = 100% AGUA
3 = 0-9%
5 = 10-15%
7 = 16-25%
8 = 25%+
9 = OCEANO

EL TAMANO DE LA CELULA DE LA MALLA ES DE 2.5 ACRES

NIVELES	0	1	2	3	4	5	6	7	8	9
FRECUENCIA	99	42	0	100	0	54	0	7	214	316



MAPA DE LA VARIABLE 3 DEL INVENTARIO DE DATOS ORIENTACION PENDIENTE

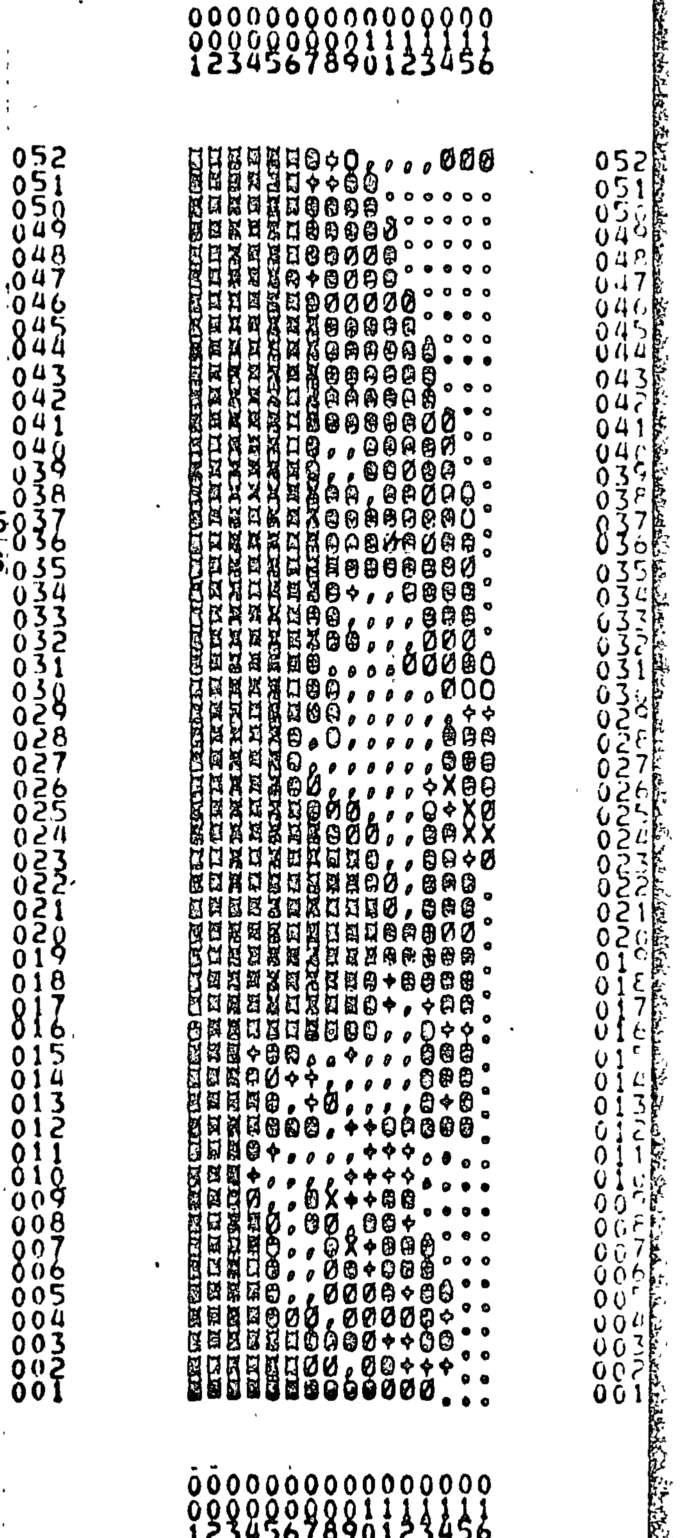
AREA DE ESTUDIO VALLE DEL PACIFICO LABORATORIO DE PLANEACION URBANA

SUBVARIABLES DATO

0 = NO HAY DATOS
1 = LLANURA
2 = NORTE
3 = NORDESTE
4 = NORESTE
5 = SUR
6 = SUROESTE
7 = SURESTE
8 = OESTE
9 = OCEANO

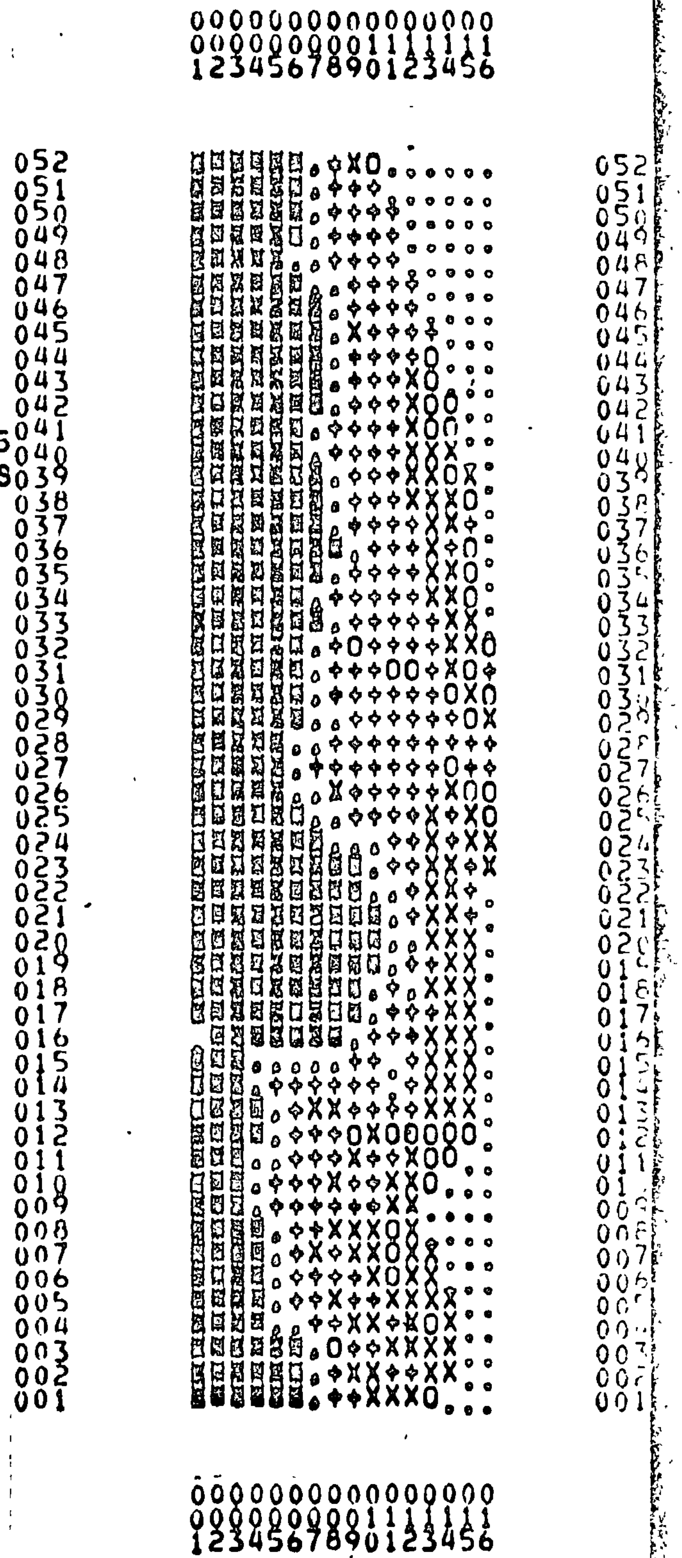
EL TAMANO DE LA CELULA DE LA MALLA ES DE 2.5 ACRES

NIVELES	0	1	2	3	4	5	6	7	8	9
FRECUENCIA	98	88	14	44	6	5	60	0	200	317



MAPA DE LA VARIABLE 4 DEL INVENTARIO DE DATOS VEGETACION POR ZONA

AREA DE ESTUDIO VALLE DEL PACIFICO LABORATORIO DE PLANEACION URBANA



SUBVARIABLES DATO

- 0 = NO HAY DATOS
- 2 = ZONA COSTERA
- 3 = SUELO PASTADO
- 4 = VEGETACION CHAPARRAL
- 5 = ZONA DE BOSQUE
- 9 = OCEANO

EL TAMANO DE LA CELULA DE LA MALLA ES DE 2.5 ACRES

NIVELES	0	1	2	3	4	5	6	7	8	9
FRECUENCIA	101	0	69	200	106	39	0	0	0	317

MAPA DE LA VARIABLE 5 DEL INVENTARIO DE DATOS DENSIDAD DE ARBOLES

AREA DE ESTUDIO VALLE DEL PACIFICO. LABORATORIO DE PLANEACION URBANA

SUBVARIABLES DATO

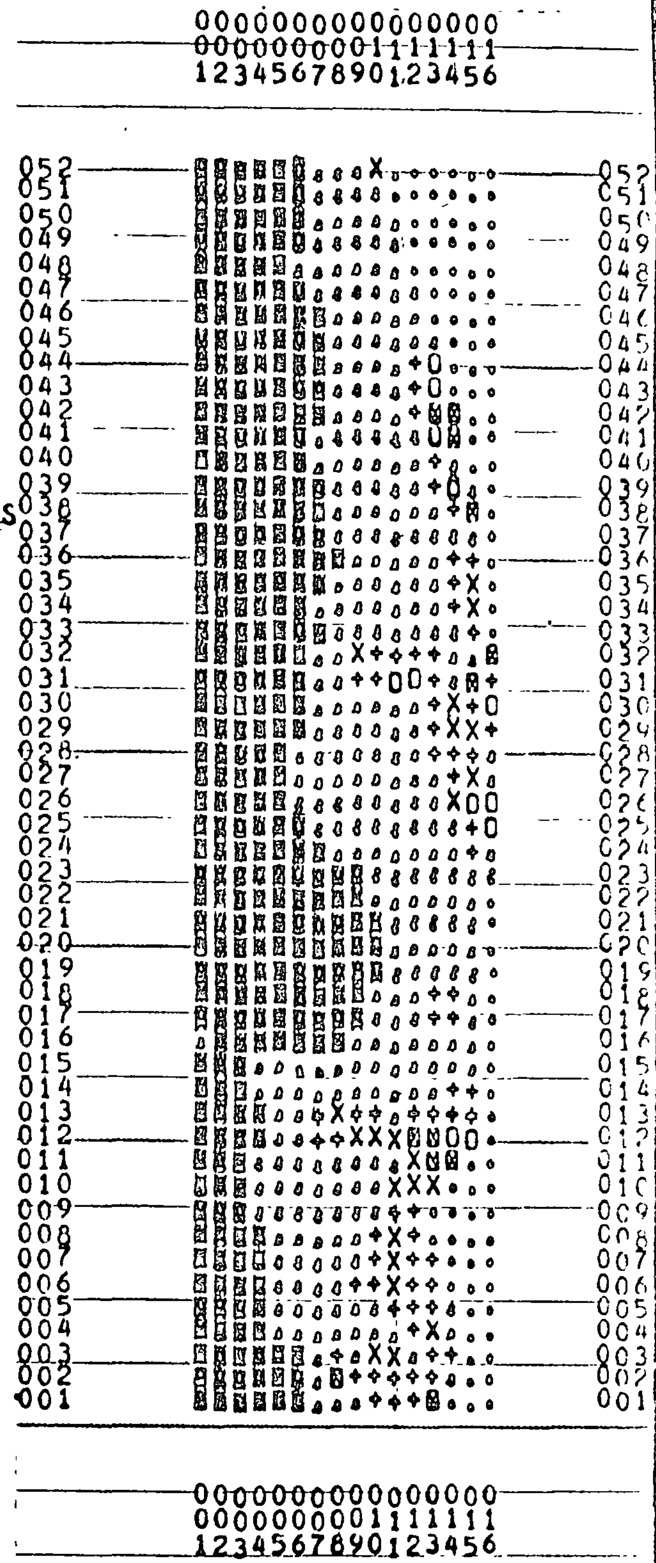
0 = NO HAY DATOS
2 = NADA
3 = 1-25%
4 = 26-50%
5 = 51-75%
7 = 76-100%
9 = OCEANO

EL TAMANO DE LA CELULA DE LA MALLA ES DE 2. ACRES

NIVELES	0	1	2
SIMBOLOS	0	1	2
FRECUENCIA	98	0	300

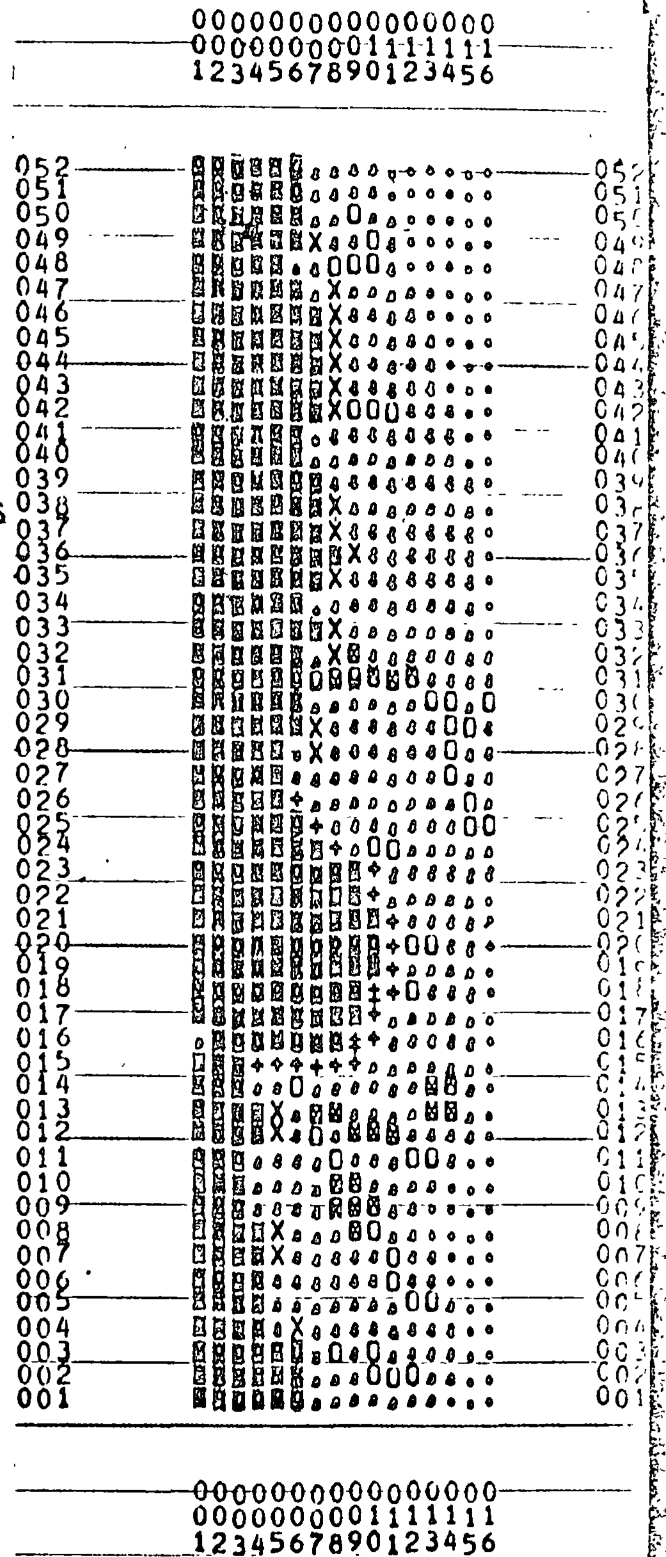
NIVELES	3	4	5	6
SIMBOLOS	3	4	5	6
FRECUENCIA	70	23	11	0

NIVELES	7	8	9
SIMBOLOS	7	8	9
FRECUENCIA	13	0	317



MAPA DE LA VARIABLE 6 DEL INVENTARIO DE DATOS ZONAS LLANAS

AREA DE ESTUDIO VALLE DEL PACIFICO LABORATORIO DE PLANEACION URBANA



SUBVARIABLES DATO

0 = NO HAY DATOS
2 = NO ZONAS LLANAS
3 = LLANOS SOBRE LA COSTA > DE 2000
4 = LLANOS SOBRE LA COSTA < DE 2000
5 = LLANOS TERRESTRES < DE 200
7 = LLANOS TERRESTRES > DE 200
9 = OCEANO

EL TAMANO DE LA CELULA DE LA MALLA ES DE 2.5 ACRES

NIVELES	0	1	2	3	4	5	6	7	8	9
FRECUENCIA	97	1	318	19	20	38	0	22	0	317

MAPA DE LA VARIABLE 7 DEL INVENTARIO DE DATOS ACCIDENTES GEOLOGICOS

AREA DE ESTUDIO VALLE DEL PACIFICO LABORATORIO DE PLANEACION URBANA

SUBVARIABLES DATO

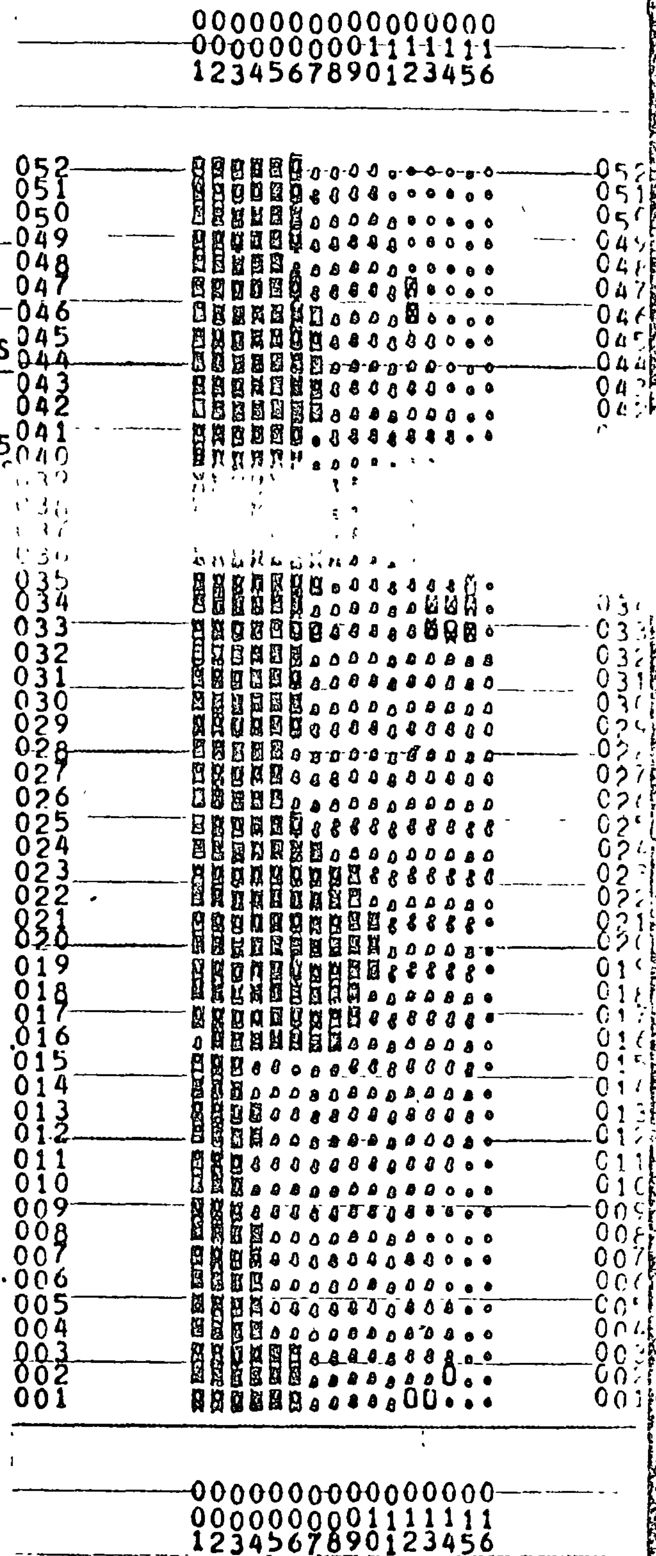
0 = NO HAY DATOS
2 = AREA ESTABLE
5 = DESLIZAMIENTOS ACTIVOS NO NATURALES
7 = DESLIZAMIENTOS ACTIVOS DE TIERRA
9 = OCEANO

EL TAMAÑO DE LA CELULA DE LA MALLA ES DE 2.5 ACRES

NIVELES	0	1	2
SIMBOLOS	0	1	2
FRECUENCIA	98	0	405

NIVELES	3	4	5	6
SIMBOLOS	3	4	5	6
FRECUENCIA	0	0	3	0

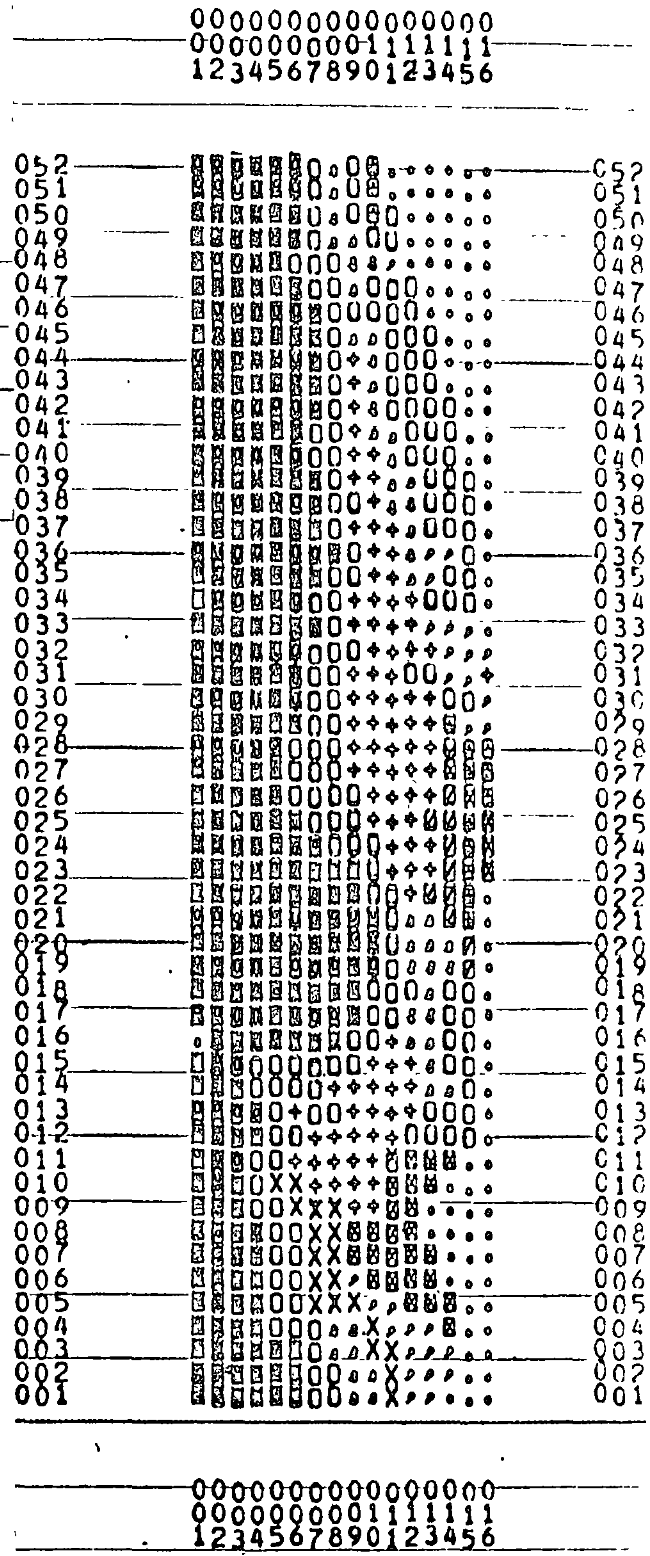
NIVELES	7	8	9
SIMBOLOS	7	8	9
FRECUENCIA	9	0	317



MAPA DE LA VARIABLE 8 DEL INVENTARIO DE DATOS SUELOS

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

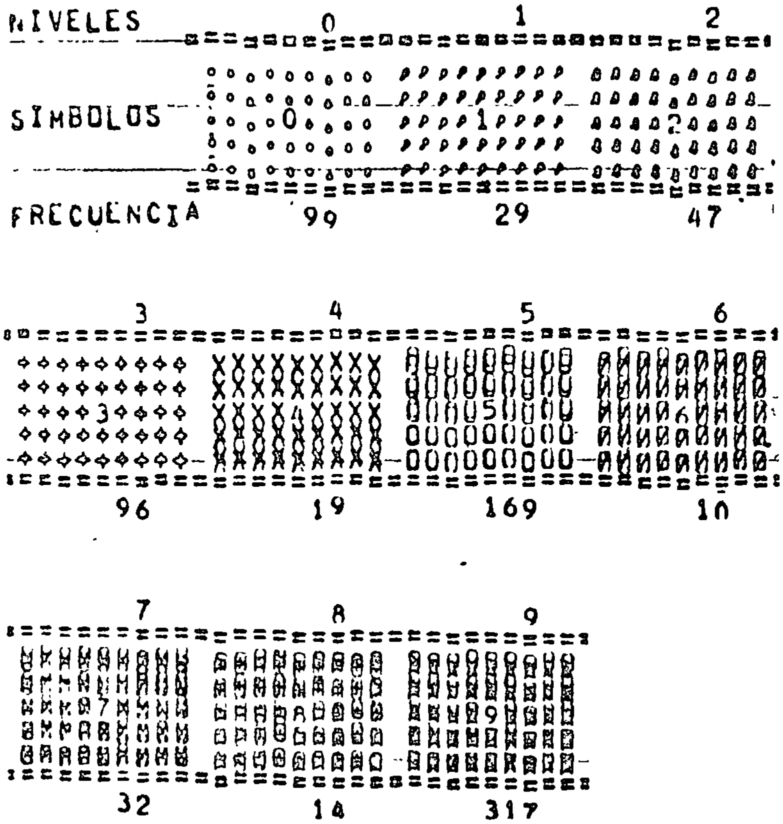


SUBVARIABLES DATO

0 = NO HAY DATOS
1 = HGFG2
2 = ST
3 = LHBC
4 = CHE
5 = RHGH
6 = GZF
7 = GRGH
8 = GZE
9 = OCEANO

EL TAMANO DE LA CELULA DE LA MALLA ES DE 2.5 ACRES

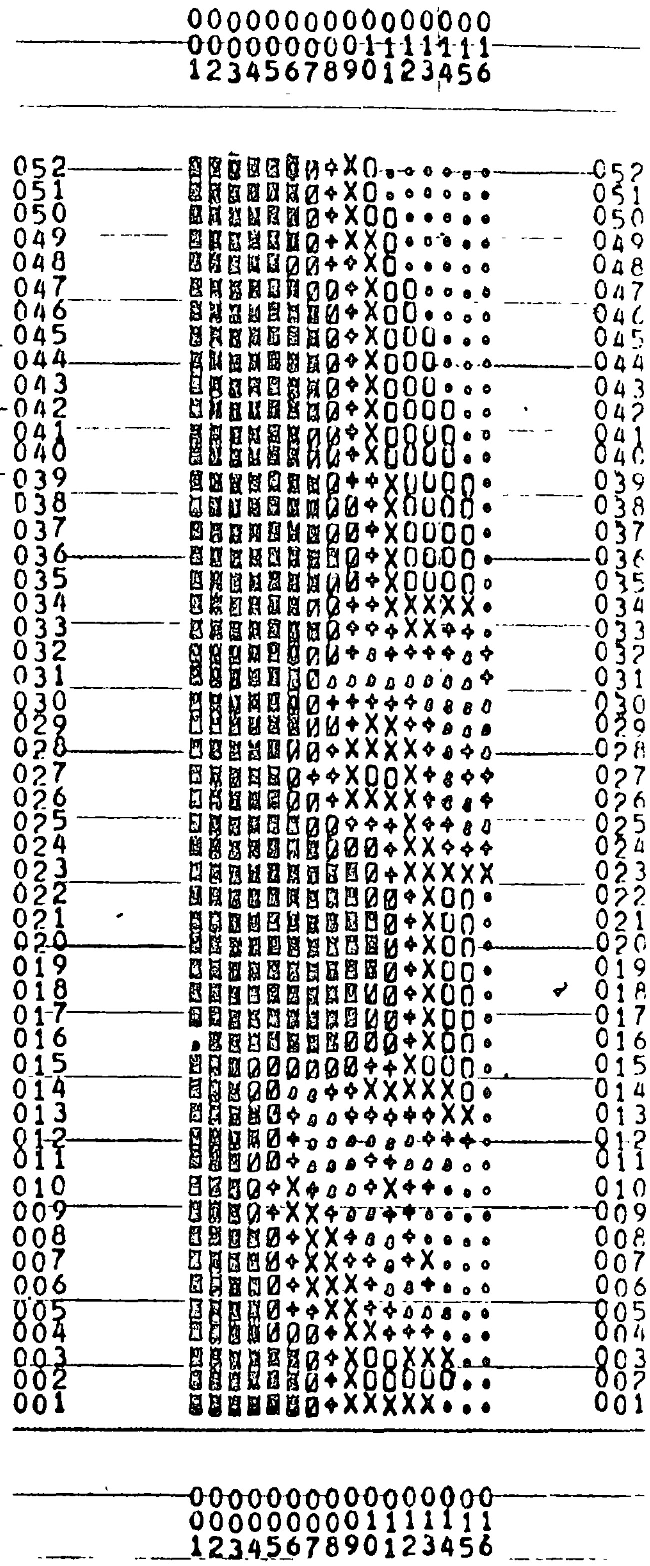
NIVELES	0	1	2	3	4	5	6	7	8	9
FRECUENCIA	99	29	47	96	19	169	10	32	14	317



MAPA DE LA VARIABLE 9 DEL INVENTARIO DE DATOS PROXIMIDAD AL AGUA

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA



SUBVARIABLES DATO

0 = NO HAY DATOS
2 = CORRIENTES DENTRO DE LA CELULA
3 = AGUA EN LAS CELULAS ADYACENTES
4 = AGUA A UNA CELULA DE PROXIMIDAD
5 = AGUA A DOS CELULAS DE PROXIMIDAD
6 = ZONA DE MAREA
9 = OCEANO

EL TAMANO DE LA CELULA DE LA MALLA ES DE

2.5 ACRES

NIVELES	0	1	2	3	4	5	6	7	8	9
SIMBOLOS	0	1	2	3	4	5	6	7	8	9
FRECUENCIA	99	0	53	117	87	78	81	0	0	317

MAPA DE LA VARIABLE 10 DEL INVENTARIO DE DATOS VISTAS

AREA DE ESTUDIO VALLE DEL PACIFICO

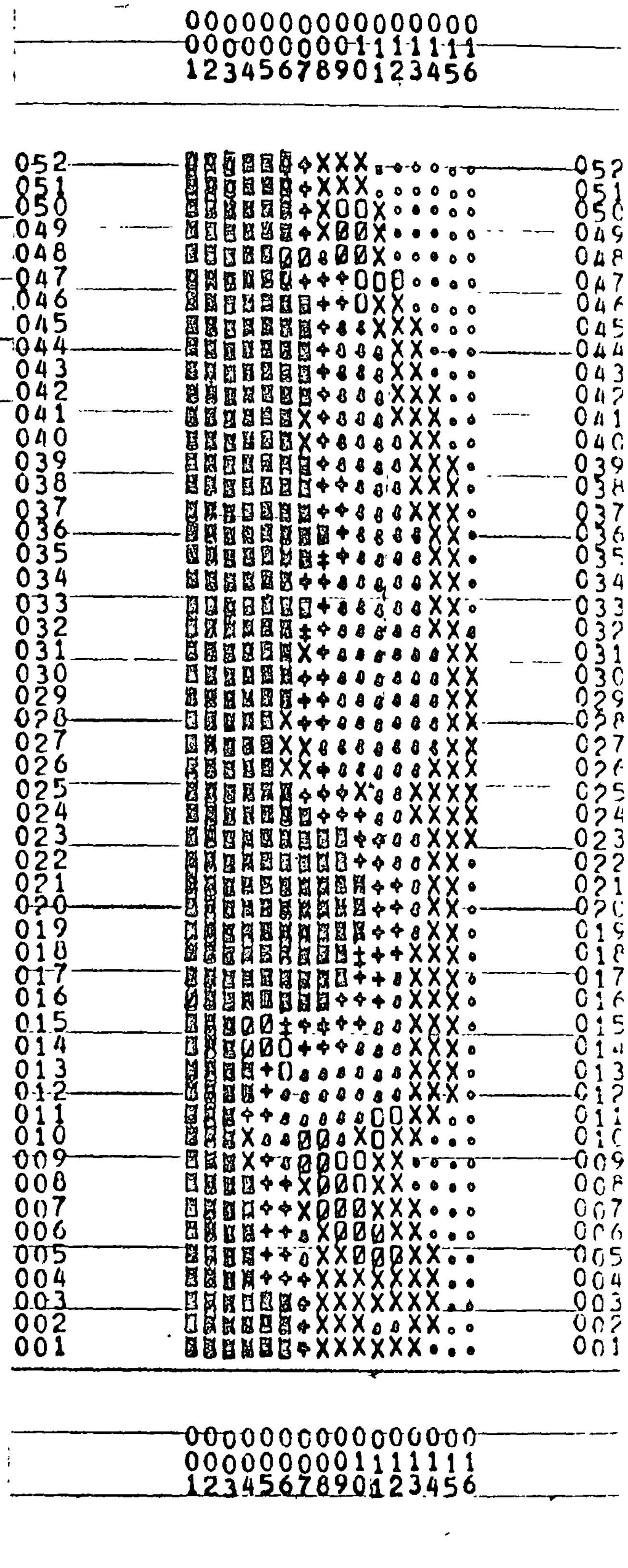
LABORATORIO DE PLANEACION URBANA

SUBVARIABLES DATO

- 0 = NO HAY DATOS
- 2 = NO HAY VISTAS IMPORTANTES
- 3 = BUENAS VISTAS AL OCEANO
- 4 = EXCELENTES VISTAS AL OCEANO
- 5 = BUENAS PANORAMICAS DEL LUGAR
- 6 = EXCELENTES PANORAMICAS DEL LUGAR
- 9 = OCEANO

EL TAMANO DE LA CELULA DE LA MALLA ES DE 2.5 ACRES

NIVELES	0	1	2	3	4	5	6	7	8	9
SIMBOLOS	0	1	2	3	4	5	6	7	8	9
FRECUENCIA	98	0	133	87	157	14	26	0	0	317



(Tarjetas del programa MAPAS/IMGRID ver capítulo VI)

```
52        0        0
IRREGULAR OUTLINE
@DATA
@FILE FILE40(TITLE=LAEP/ATTR10)
@FILE FILE39(TITLE=LAEP/ATTR09)
@FILE FILE38(TITLE=LAEP/ATTR08)
@FILE FILE37(TITLE=LAEP/ATTR07)
@FILE FILE36(TITLE=LAEP/ATTR06)
@FILE FILE35(TITLE=LAEP/ATTR05)
@FILE FILE34(TITLE=LAEP/ATTR04)
@FILE FILE33(TITLE=LAEP/ATTR03)
@FILE FILE32(TITLE=LAEP/ATTR02)
@FILE FILE31(TITLE=LAEP/ATTR01)
@EXECUTE MAPAS/IMGRID
@JOB IMGRID; USER AA80/AA;CLASS 3;BEGIN;
```

Centro de Servicios de Cómputo

Figura 6.- Forma de obtener los mapas de los "modelos de atractividad" con el programa MAPAS/IMGRID. Este ejemplo muestra la manera de obtener 10 mapas (FILE31,..,40) a partir de los archivos (LAEP/ATTR01,...,10) que fueron generados por el programa ATRACTIVO/IMGRID.

*** MODELO DE ATRACTIVIDAD NUMERO 1***

LOS INDICES DE ATRACTIVIDAD SON PUESTOS EN EL ARCHIVO EN DISCOLAEP/ATTR 1

EL NUMERO DE VARIABLES EN EL MODELO ES 4

ESTAS SON LAS TARJETAS DE LA HOJA DE CODIFICACION USADA EN EL MODELO :

10	0 0 0 6 8 7 9 0 0 0	2
6	0 0 0 6 6 9 0 9 0 0	2
9	0 0 9 7 4 0 5 0 0 0	2
2	0 0 0 9 0 8 0 5 0 0	1

TITULO DEL MAPA

MODELO DE ATRACTIVIDAD # 1: ACTIVIDADES DE ESPARCIMIENTO

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

OPCIONES USADAS PARA ESTE MAPA

1 EL TAMAÑO DE LA MALLA ES 52 RENGLONES Y 16 COLUMNAS
EL TAMAÑO DE LA CELULA ES 1 CARACTERES EN SENTIDO VERTICAL
Y 1 CARACTERES EN SENTIDO HORIZONTAL

7 LOS SIMBOLOS SON
.,,+XOOOOO0123456789
.+XOOOOA
7X*X
+V

10 EL TEXTO DEL MAPA ES
VARIABLES DATO USADAS EN EL MODELO

NO	NOMBRE	PESO
10	VISTAS	2
6	ZONAS LLANAS	2
9	VISTAS	2
2	PORCIENTO DE PENDIENTE	1

13 LA NUMERACION DE LA MALLA COMIENZA EN 1 52

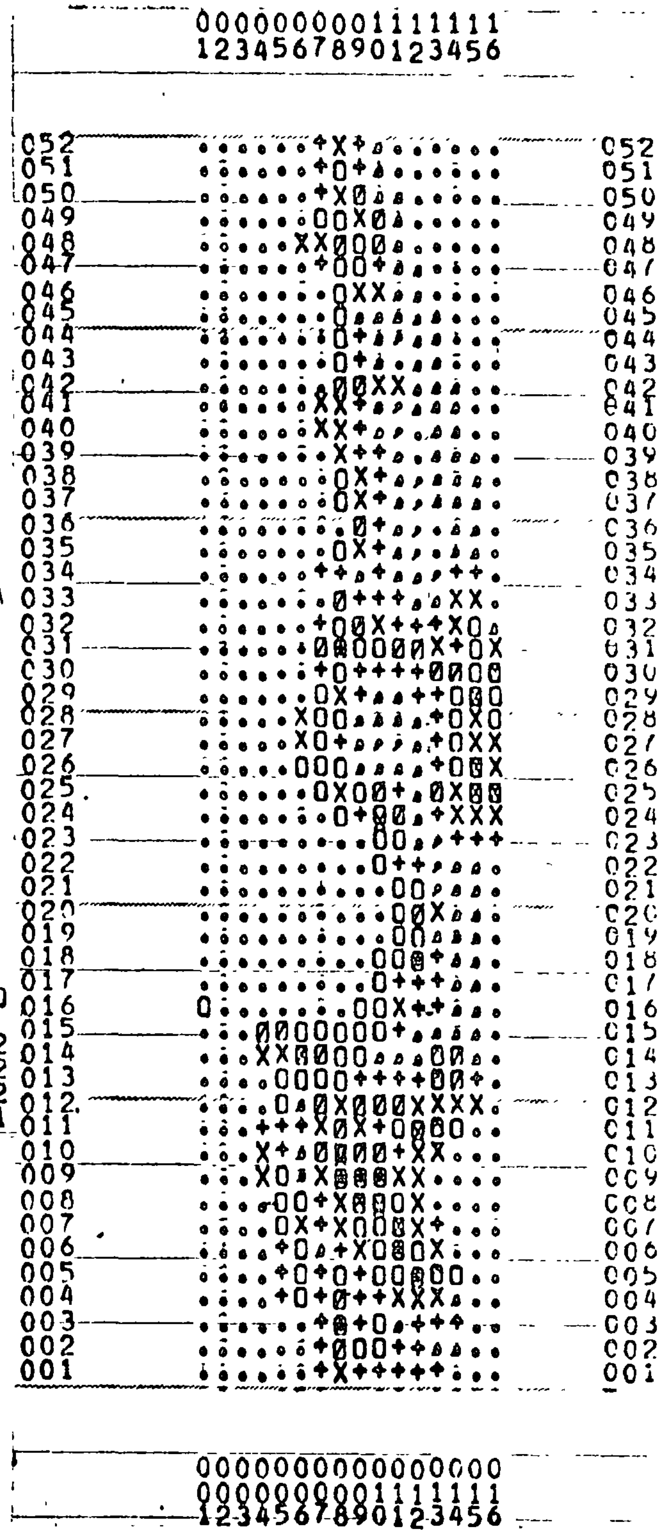
14 SE SUPONE QUE LOS DATOS ESTAN PRE-ESCALADOS

MODELO DE ATRACTIVIDAD # 1

ACTIVIDADES DE ESPARCIMIENTO

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA



VARIABLES DATO USADAS EN EL MODELO

NO	NOMBRE	PESO
10	VISTAS	2
6	ZONAS LLANAS	2
9	VISTAS	2
2	PORCIENTO DE PENDIENTE	1

NIVELES	0	1	2	3	4	5	6	7	8	9
FRECUENCIA	420	13	101	94	64	82	33	16	8	1

*** MODELO DE ATRACTIVIDAD NUMERO 2***

LOS INDICES DE ATRACTIVIDAD SON PUESTOS EN EL ARCHIVO EN DISCOLAEP/ATTR 2
EL NUMERO DE VARIABLES EN EL MODELO ES 4
ESTAS SON LAS TARJETAS DE LA HOJA DE CODIFICACION USADA EN EL MODELO :

2	C 0 0 9 0 2 0 0 0 0	2
5	C 0 0 7 9 6 0 5 0 0	1
6	C 0 3 0 0 6 0 9 0 0	1
8	0 9 0 9 0 9 0 9 0 0	1

TITULO DEL MAPA

MODELO DE ATRACTIVIDAD # 2: ESTACIONAMIENTOS
AREA DE ESTUDIO VALLE DEL PACIFICO
LABORATORIO DE PLANEACION URBANA

OPCIONES USADAS PARA ESTE MAPA

10 EL TEXTO DEL MAPA ES

VARIABLES DATO USADAS EN EL MODELO

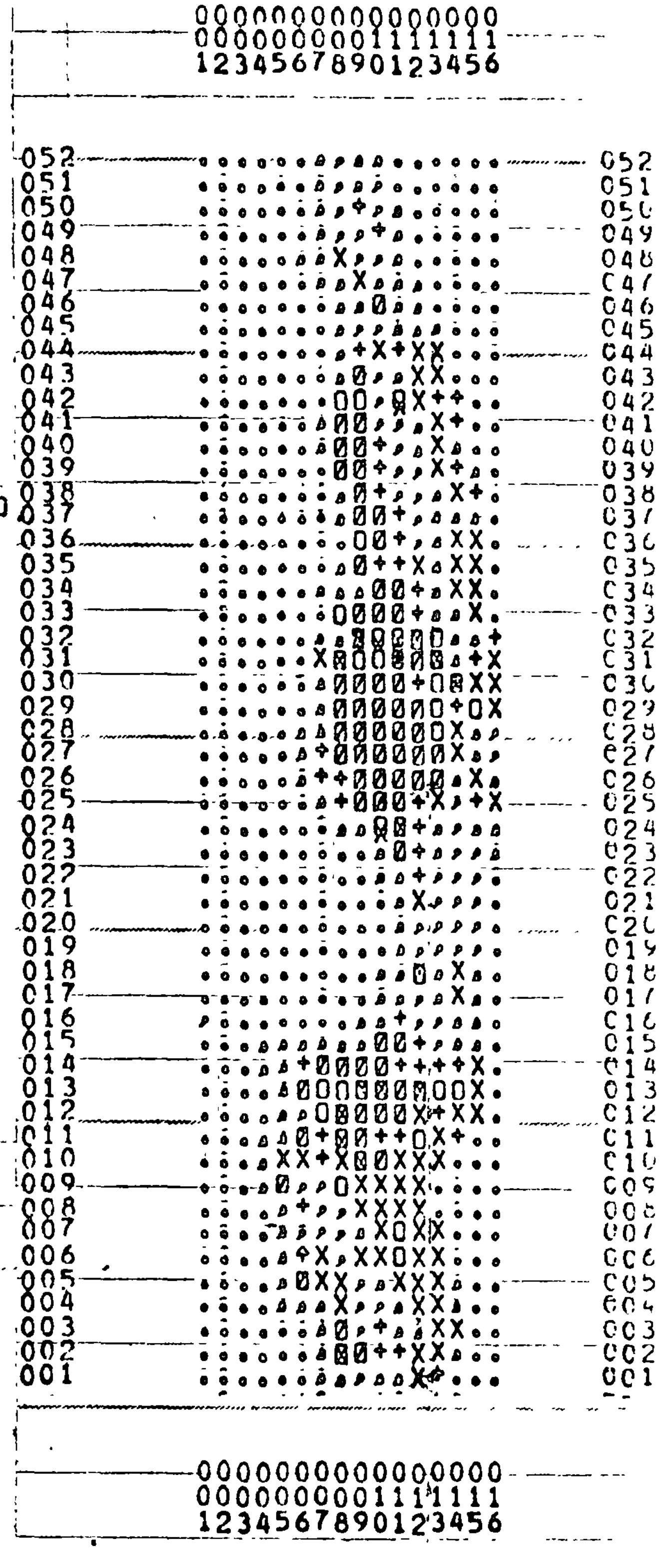
NO	NOMBRE	PESO
3	PORCIENTO DE PENDIENTE	2
5	PORCIENTO DENSIDAD DE ARBOLES	1
6	ZONAS LLANAS	1
8	SOILS	1

MODELO DE ATRACTIVIDAD # 2

ESTACIONAMIENTOS

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA



VARIABLES DATO USADAS EN EL MODELO

NO	NOMBRE	PESO
3	PORCIENTO DE PENDIENTE	2
5	PORCIENTO DENSIDAD DE ARBOLES	1
6	ZONAS LLANAS	1
8	SOILS	1

NIVELES	0	1	2	3	4	5	6	7	8	9
SIMBOLOS										
FRECUENCIA	415	59	129	50	74	16	66	20	2	1

MODELO DE ATRACTIVIDAD NUMERO 3

LOS INDICES DE ATRACTIVIDAD SON PUESTOS EN EL ARCHIVO EN DISCOLAEP/ATTR 3

EL NUMERO DE VARIABLES EN EL MODELO ES 4

ESTAS SON LAS TARJETAS DE LA HOJA DE CODIFICACION USADA EN EL MODELO :

2	0 0 0 9 0 7 0 0 0 0	1
5	0 0 0 6 8 9 0 7 0 0	1
6	0 0 0 0 0 9 0 7 0 0	1
7	0 0 9 0 0 0 0 0 0 0	1

TITULO DEL MAPA

MODELO DE ATRACTIVIDAD # 3! ESTRUCTURAS

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

OPCIONES USADAS PARA ESTE MAPA

10. EL TEXTO DEL MAPA ES

VARIABLES DATO USADAS EN EL MODELO

NO	NOMBRE	PESO
2	PORCIENTO DE PENDIENTE	1
5	PORCIENTO DENSIDAD DE ARBOLES	1
6	ZONAS LLANAS	1
7	ACCIDENTES GEOLOGICOS	1

MODELO DE ATRACTIVIDAD # 3

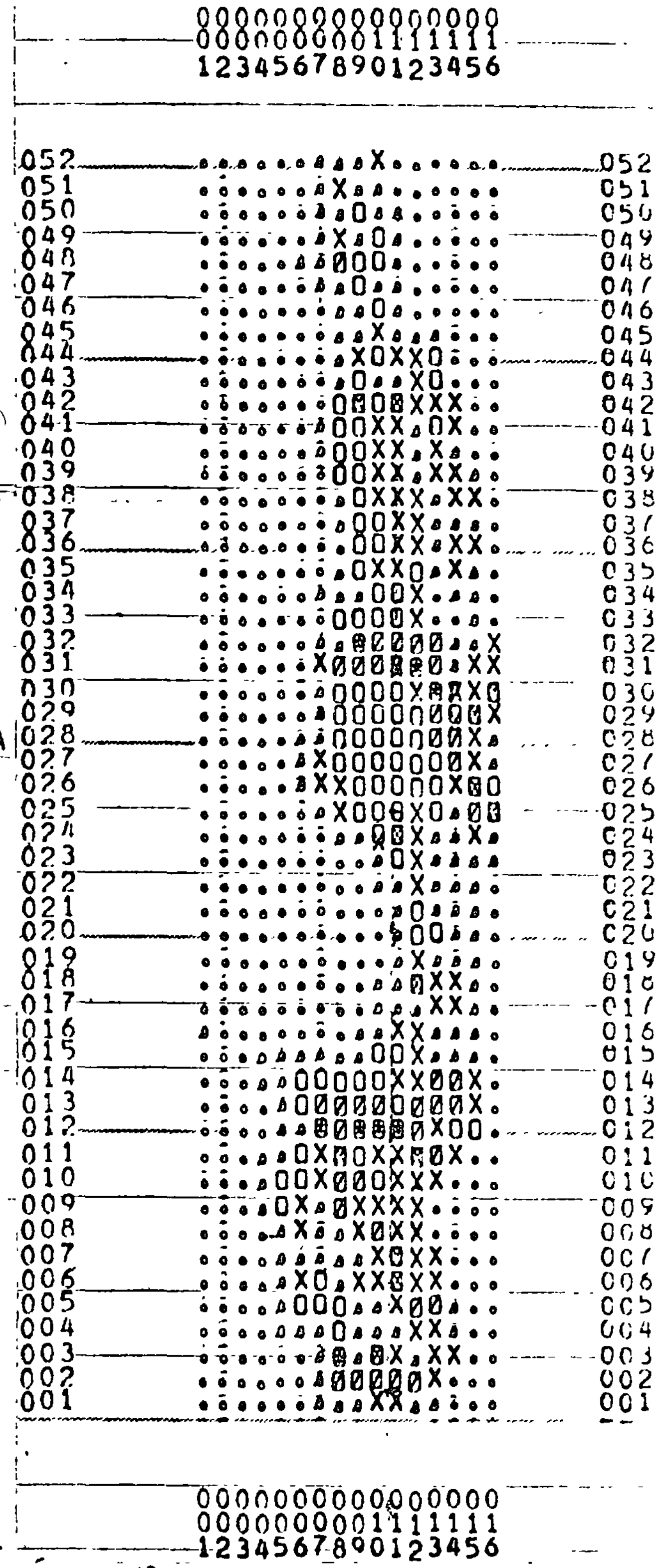
ESTRUCTURAS

AREA DE ESTUDIO VALLE DEL PACIFICO

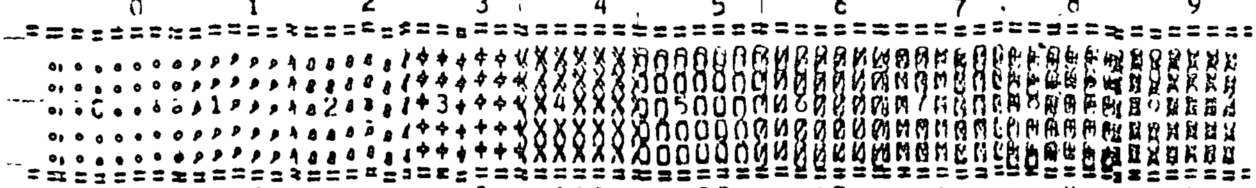
LABORATORIO DE PLANEACION URBANA

VARIABLES DATO USADAS EN EL MODELO

NO	NOMBRE	PESO
2	PORCIENTO DE PENDIENTE	1
5	PORCIENTO DENSIDAD DE ARBOLES	1
6	ZONAS LLANAS	1
7	ACCIDENTES GEOLOGICOS	1



NIVELES	0	1	2	3	4	5	6	7	8	9
SIMBOLOS										
FRECUENCIA	421	0	150	0	103	88	37	15	8	2



*** MODELO DE ATRACTIVIDAD NUMERO 4***

LOS INDICES DE ATRACTIVIDAD SON PUESTOS EN EL ARCHIVO EN DISCOLAEP/ATTR 4

EL NUMERO DE VARIABLES EN EL MODELO ES 2

ESTAS SON LAS TARJETAS DE LA HOJA DE CODIFICACION USADA EN EL MODELO :

2	0 0 0 9 0 5 0 0 0 0	1
9	0 0 0 2 5 9 0 0 0 0	1

TITULO DEL MAPA

MODELO DE ATRACTIVIDAD # 4: CABALLERIZAS

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

OPCIONES USADAS PARA ESTE MAPA

10 EL TEXTO DEL MAPA ES

VARIABLES DATO USADAS EN EL MODELO

NO	NOMBRE	PESO
2	PORCIENTO DE PENDIENTE	2
9	PROXIMIDAD AL AGUA	1

*** MODELO DE ATRACTIVIDAD NUMERO 5***

LOS INDICES DE ATRACTIVIDAD SON PUESTOS EN EL ARCHIVO EN DISCOLAEP/ATTR 5

EL NUMERO DE VARIABLES EN EL MODELO ES 6

ESTAS SON LAS TARJETAS DE LA HOJA DE CODIFICACION USADA EN EL MODELO :

2	0 0 0 9 0 3 0 0 0 0	2
5	0 0 0 6 8 9 0 8 0 0	2
6	0 0 0 0 0 9 0 9 0 0	2
4	0 0 0 0 6 9 0 0 0 0	1
9	0 0 9 9 7 3 0 0 0 0	1
1	0 9 5 0 0 0 5 0 9 0	1

TITULO DEL MAPA

MODELO DE ATRACTIVIDAD # 5: ESTACIONAMIENTO PARA TRAILERS

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

OPCIONES USADAS PARA ESTE MAPA

10 EL TEXTO DEL MAPA ES

VARIABLES DATO USADAS EN EL MODELO

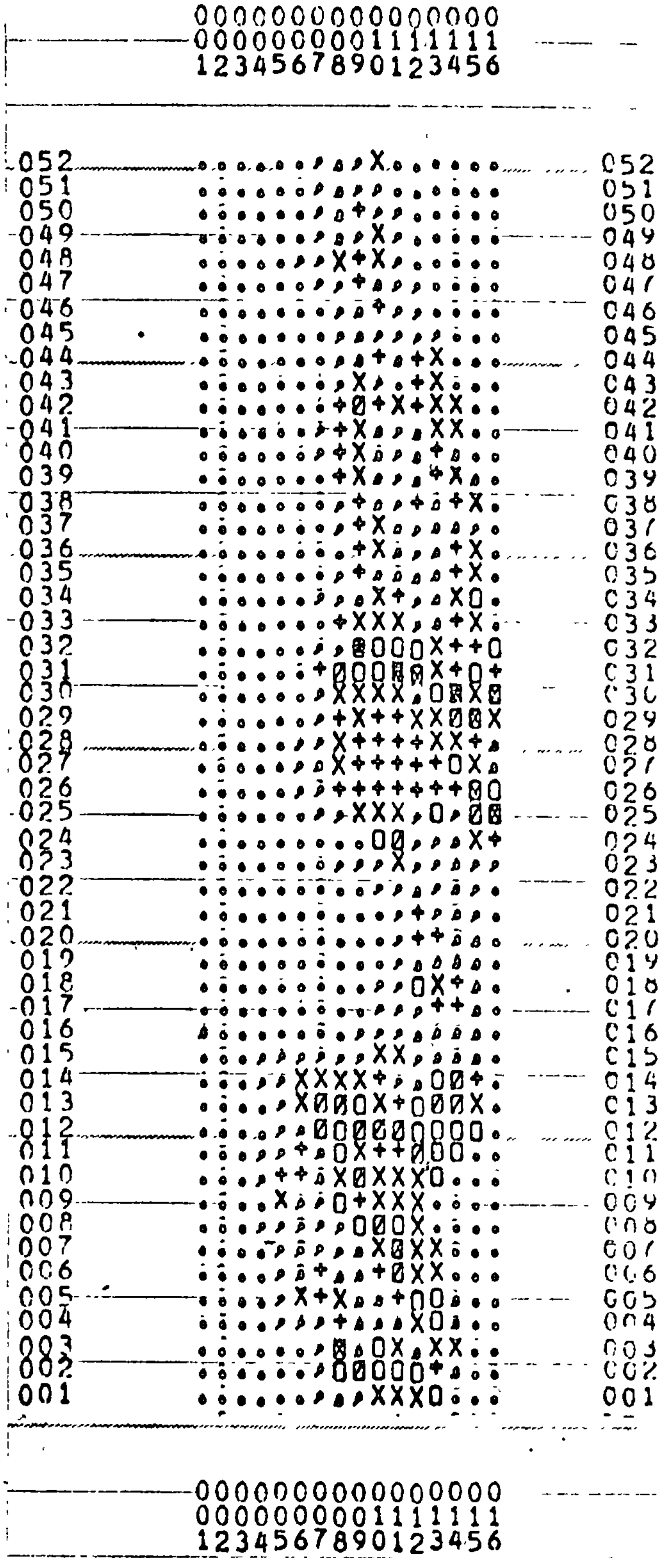
NO	NOMBRE	PESO
2	PORCIENTO DE PENDIENTE	2
5	PORCIENTO DENSIDAD DE ARBOLES	2
6	ZONAS LLANAS	2
4	VEGETACION POR ZONA	1
9	PROXIMIDAD AL AGUA	1
1	USOS DEL SUELO ACTUALES	1

MODELO DE ATRACTIVIDAD # 5

ESTACIONAMIENTO PARA TRAILERS

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA



VARIABLES DATO USADAS EN EL MODELO

NO	NOMBRE	PESO
2	PORCIENTO DE PENDIENTE	2
5	PORCIENTO DENSIDAD DE ARBOLES	2
6	ZONAS LLANAS	2
4	VEGETACION POR ZONA	1
9	PROXIMIDAD AL AGUA	1
1	USOS DEL SUELO ACTUALES	1

NIVELES	0	1	2	3	4	5	6	7	8	9
SIMBOLOS										
FRECUENCIA	416	116	70	73	82	37	19	8	1	2

*** MODELO DE ATRACTIVIDAD NUMERO 6***

LOS INDICES DE ATRACTIVIDAD SON PUESTOS EN EL ARCHIVO EN DISCOLAEP/ATTR 6

EL NUMERO DE VARIABLES EN EL MODELO ES 5

ESTAS SON LAS TARJETAS DE LA HOJA DE CODIFICACION USADA EN EL MODELO :

2	0 0 0 9 0 8 0 4 0 0	2
6	0 0 0 0 0 9 0 8 0 0	2
5	0 0 0 5 8 9 0 8 0 0	1
9	0 0 9 9 5 0 0 0 0 0	1
4	0 0 0 0 5 9 0 0 0 0	1

TITULO DEL MAPA

MODELO DE ATRACTIVIDAD # 6: CAMINATA

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

OPCIONES USADAS PARA ESTE MAPA

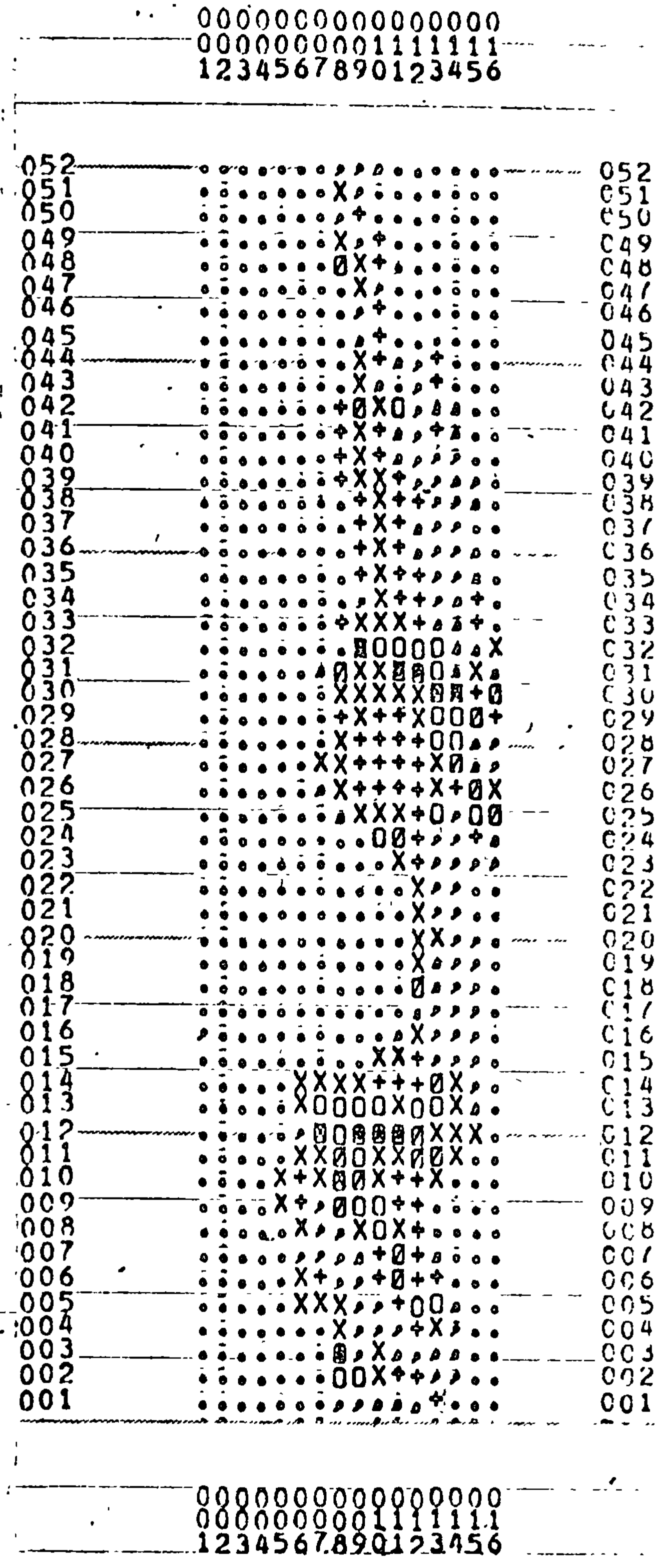
10 EL TEXTO DEL MAPA ES

VARIABLES DATO USADAS EN EL MODELO

NO	NOMBRE	PESO
2	PORCIENTO DE PENDIENTE	2
6	ZONAS LLANAS	2
5	PORCIENTO DENSIDAD DE ARBOLES	1
9	PROXIMIDAD AL AGUA	1
4	VEGETACION POR ZONA	1

MODELO DE ATRACTIVIDAD # 6

CAMINATA



AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

VARIABLES DATO USADAS EN EL MODELO

NO	NOMBRE	PESO
2	PORCIENTO DE PENDIENTE	2
3	ZONAS LLANAS	2
5	PORCIENTO DENSIDAD DE ARBOLES	1
9	PROXIMIDAD AL AGUA	1
4	VEGETACION POR ZONA	1

NIVELES	0	1	2	3	4	5	6	7	8	9
SIMBOLES	0	1	2	3	4	5	6	7	8	9
FRECUENCIA	497	78	42	76	81	28	19	3	5	3

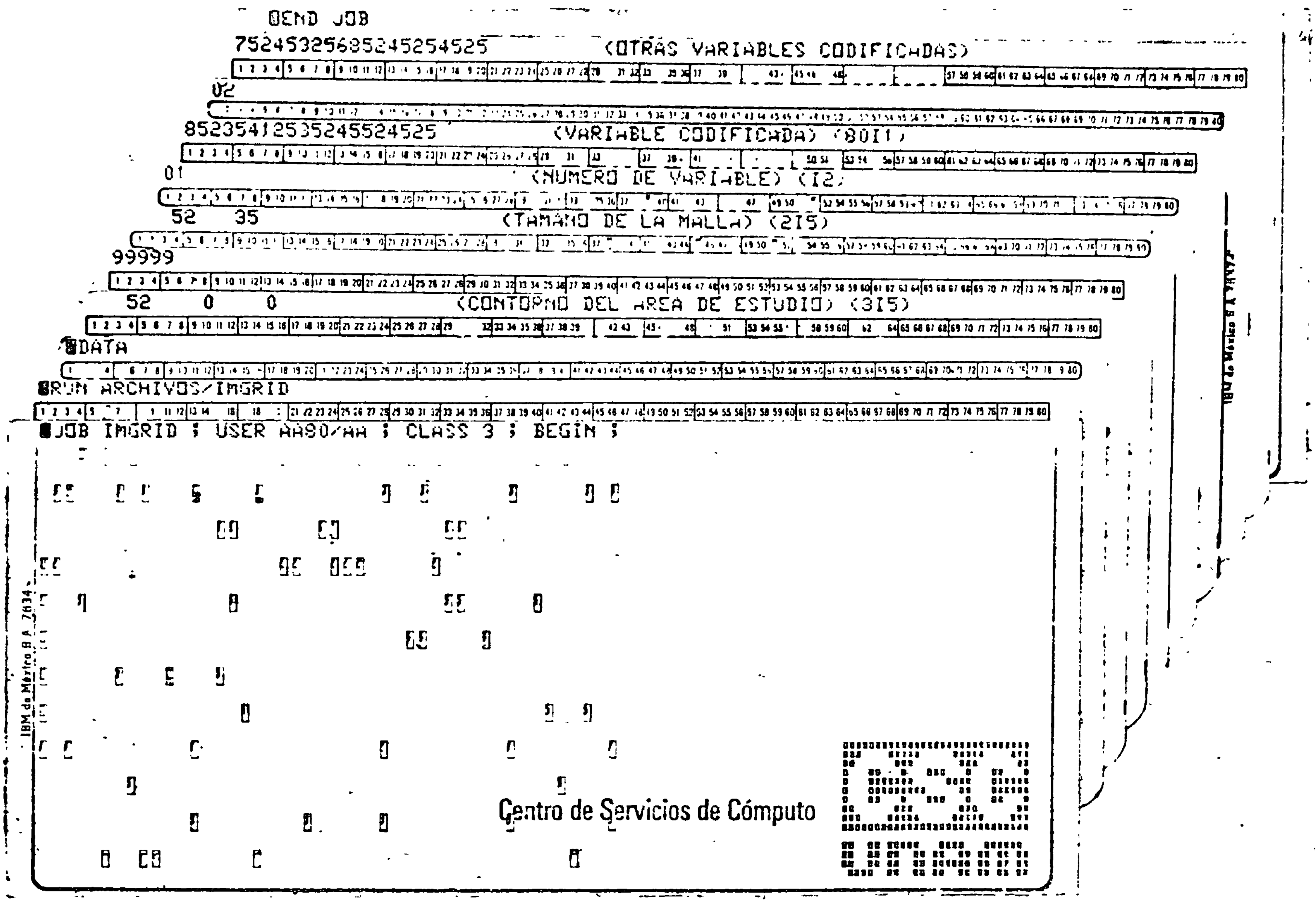


Figura 2. Forma de cargar el BANCO DE DATOS en el sistema B-6700 del CSC, UNAM.

MODELOS DE ATRACTIVIDAD

En esta fase del proceso con el paquete *IMGRID* debe usted establecer una definición detallada de los *criterios de atractividad, para cada uso del suelo (actividad) que se desea localizar dentro del area de estudio.*

El programa *ATRACTIVO/IMGRID*, facilita la evaluación de un máximo de hasta 20 modelos de atractividad en cada corrida.

Se recomienda al usuario seguir secuencialmente la lista de pasos descrita a continuación:

Paso 1 *DEFINIR EL PROPOSITO DE LOS MODELOS.*

En primer lugar se debe obtener una lista especificando cuales son las actividades o usos del suelo que se desean obtener localizadas en los lugares con mayor *atractividad* dentro del area de estudio.

Paso 2 *DETERMINAR LAS VARIABLES RELEVANTES EN LOS MODELOS.*

Para cada una de las actividades definidas en el paso anterior, se deben identificar, de las variables contenidas en el banco de datos, cuales son las que representan los *factores para una localización adecuada , a partir de los requerimientos específicos de las actividades analizadas en los modelos.*

El programa ATRACTIVO/IMGRID permite analizar hasta un máximo de 10 variables o factores de localización, en cada modelo.

Paso 3 *PONDERAR LAS VARIABLES RELEVANTES.*

En este paso se debe distinguir entre las variables consideradas aquellas que tengan mayor importancia en la localización de la actividad para la cual fue creado el modelo. El programa *ATRACTIVO/IMGRID* reconoce una jerarquia de 3 niveles para asignar prioridad a ciertos factores de localización sobre otros, de acuerdo a la siguiente escala:

1 *menos importante*
2 *importante*
3 *muy importante*

Paso 4 *DISTINGUIR LAS SUBVARIABLES IMPORTANTES.*

En este momento debe usted asignar un número del 0 al 9 a las subvariables que componen los factores de localización que intervienen en el modelo designando con esto *mayor preferencia a ciertas subvariables* en función de los requerimientos específicos considerados para cada una de las actividades que se pretenden localizar dentro del area de estudio.

Paso 5 *LLENAR LAS HOJAS DE CODIFICACION Y PERFORAR SUS TARJETAS.*

En las figuras 4 y 5 se explican tanto el orden, como los formatos en los que debe usted perforar sus tarjetas, *asegurese minuciosamente de cumplir con lo especificado en dichas figuras para evitar una sorpresa desagradable por una corrida frustrada, y recuerde que las pruebas de escritorio le ahorran un valioso tiempo de máquina.*

Paso 6 *ALIMENTAR SUS TARJETAS PERFORADAS EN LA COMPUTADORA.*

Una vez que ha verificado el orden y los campos en los que perforó su paquete de tarjetas deberá proceder a que este sea leido y procesado por la computadora mediante el programa *ATRACTIVO/IMGRID* .Este se encarga de analizar cada una de las células del area en estudio y determina de acuerdo a los criterios de *atractividad* proporcionados, si esa célula cuenta con los atributos necesarios para una buena localización del uso del suelo analizado, *construyendo un índice de atractividad para cada célula en función de la importancia o peso asignado a las variables y de las preferencias de cada una de las subvariables que las componen.*

El programa genera 20 archivos en disco magnético donde deposita los resultados de cada modelo de atractividad, estos archivos son llamados; *LAEP/ATTR01*, *LAEP/ATTR02*,..., *LAEP/ATTR20*. Los archivos *LAEP/ATTRn*, serán posteriormente utilizados por el programa *MAPAS/IMGRID* (vease la figura 6) para obtener el mapeo con los resultados de los modelos de atractividad, y por el programa para la evaluación de planes en etapas posteriores del paquete *IMGRID*.

Paso 7 *MAPEAR LOS RESULTADOS.*

Después de que todas las células han sido evaluadas se procede a obtener los mapas para cada modelo de atractividad con la utilización del programa *MAPAS/IMGRID* (descripción en la figura 6). En estos mapas la atractividad de cada célula se representa con una gama de caracteres sobreimpresos de tal modo de obtener una escala de 10 intensidades de grises, en donde el mas oscuro muestra la mayor atractividad para la localización de las actividades que fueron definidas como propósito del modelo.

Paso 8 *REVISAR LOS RESULTADOS.*

PAQUETE IMGRID

Programa ATRACTIVO/IMGRID

HOJA DE CODIFICACION PARA LOS MODELOS DE ATRACTIVIDAD

Nombre del(os) usuario(s):

Nombre de la actividad (usos del suelo) para la que se construyó el modelo:

Número del modelo (formato I2); --

Número de variables que intervienen en el modelo (formato I2); __

Columnas 1 y 2 Cuales son las variables relevantes nO.	Cual es la preferencia de las subvariables (del 0 al 9) Col-10	Col-12	Col-14	Col-16	Col-18	Col-20	Col-22	Col-24	Col-26	Col-28	Cual es la importancia de la variable (1al3 Col-34

Figura 4.- Hoja de codificación para los *MODELOS DE ATRACTIVIDAD*.

```
@END JOB
03      0 0 2 5 4 2 3 5 2 2     1     (OTRAS VARIABLES)
06      0 0 2 5 4 2 3 5 2 2     1     (OTRAS VARIABLES)
05      0 0 2 5 4 2 3 5 2 2     1     (OTRAS VARIABLES)
02      0 0 2 5 4 2 3 5 2 2     1     (OTRAS VARIABLES)
05                    (NUMERO DE VARIABLES QUE INTERVIENEN EN OTRO MODELO)
02                                  (NUMERO DEL MODELO)(I 2)
07      0 2 5 3 0 4 2 2 1 0     2     (I2,6X,10I2,5X,I1)
10      3 2 5 4 2 3 5 4 2 2     2     (PESO DE LA VARIABLE)(I 1)
09      2 4 6 5 8 7 5 4 2 0     1     (PREFERENCIAS DE SUBVARIABLES)(10 I 2)
04      2 5 4 3 5 8 7 9 8 5     3     (NO. DE VARIABLE(I 2)
04                    (NUMERO DE VARIABLES QUE INTERVIENEN EN EL MODELO) (I 2)
01                                  (NUMERO DEL MODELO)(I 2)
   52   20                          (TAMANO DE LA MALLA)(2 I 5)
99999
   52    0    0                     (CONTORNO IRREGULAR)(3 I 5)
@DATA
@EXECUTE ATRACTIVO/IMGRID
@JOB IMGRID ; USER AA80/AA ; CLASS 3 ; BEGIN ;
```

Centro de Servicios de Cómputo

Figura 5.- forma de ordenar su paquete de tarjetas para utilizar el programa ATRACTIVO/IMGRID en el sistema B-6700 del CSC de la UNAM.

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

MANEJO DE SISTEMAS DE INFORMACION GEOGRAFICA EN PLANEACION.

TEMA: BANCO DE DATOS CETENAL.

PROF. DR. ADOLFO GUZMAN ARENAS.

Noviembre, 1977.

INTRODUCCION

El Banco de Datos CETENAL (BDC) es una herramienta para manejo masivo de la información (producida por CETENAL y contenida en las diversas cartas), que describe el medio geográfico y socioeconómico.[4]

La información que contiene el BDC es la que un usuario obtendría de las siguientes cartas:

Geológica

Topográfica

Uso del Suelo

Edafológica

Uso Potencial

Se podrían vaciar al BDC otras cartas tales como la Carta Urbana, la de Climas y la Turística, ya que la estructura del mismo acepta gran variedad de información.

El BDC está diseñado para proporcionar varios niveles de consulta. Esto significa que se puede consultar a diferentes tamaños de

[4] Véase referencias 6, 8 pags. 180, 181

área, lo cual activa el acceso a la información.

La información de las cartas se clasifica en tres tipos:

1) Propiedades superficiales.- Comprenden todas las propiedades que están definidas por áreas, por ejemplo: el área que comprende el suelo aluvial, el área urbana, el área de nopalera, etc, están en el BDC como propiedades de este tipo.

2) Propiedades lineales.- Comprenden todas las propiedades que implican uniones entre determinados puntos o comunicación entre ellos, ejemplo: una carretera, vías de energía eléctrica, telegráfica, telefónica, etc.

3) Propiedades puntuales.- Estas propiedades están definidas por su localización, ejemplo: minas, pozos, puntos de verificación, etc.

Las propiedades superficiales están definidas por su clave y por su cantidad, dada como un porcentaje del área del nivel al que pertenecen. Las propiedades lineales y puntuales, están definidas por el número de ellas dentro del nivel al que pertenecen y también por su clave. Hay otras propiedades puntuales, como los servicios de la población y los servicios propuestos para la población en un pueblo, que también están en el BDC.

Como se puede observar, toda la información, ya sea geográfica o socioeconómica, puede ser clasificada dentro de los tres tipos de propiedades anteriores.

El vaciado de la información al BDC se realiza en la Oficina de Bases de Datos del Departamento de Informática de CETENAL con una metodología bien definida.

Las formas de salida son de dos tipos: mensajes escritos por impresora de la computadora, y mensajes por medio de pantallas de CRT, o sea por una terminal remota; esto permitirá tener acceso al BDC, aún estando distante de él. En este manual se describen ambas formas de salida.

Después de definir las funciones del lenguaje de usuario, se describe la forma de combinarlas.

En la parte última del manual se describen los programas de "aprendizaje" del BDC, con una serie de ejemplos para su análisis y conclusiones.

Un aspecto importante del método de "aprendizaje", es el de poder mejorar las experiencias anteriores y agregar nuevos parámetros y factores al BDC.

En los ejemplos de "aprendizaje", también se desarrolló un programa que verifica la confiabilidad del "aprendizaje".

CAPITULO I

DEFINICION DE LOS NIVELES DE CONSULTA

Para establecer los distintos niveles de consulta, se comienza por definir el nivel de regiones. Una región comprende el área dada por un rectángulo de 6° de longitud x 4° de latitud. El territorio nacional está cubierto por 20 regiones aproximadamente. Los siguientes niveles se establecen por subdivisiones sucesivas del rectángulo de menor tamaño que se tenga definido. Así, el nivel de subregiones se forma al subdividir cada región en 12 subregiones que son los elementos de un arreglo matricial de 3 x 4.

Las microregiones constituyen un arreglo matricial de 6 x 4 sobre una subregión. Cada microregión constituye una carta 1:50 000 de las producidas actualmente por CETENAL. La República Mexicana está cubierta por 2336 de estas cartas, aproximadamente. La información vaciada al BDC es extraída actualmente de estas cartas, aunque se vació también una carta urbana escala 1:5 000 y se ha proyectado vaciar las cartas de climas escala 1:500 000. Las microregiones se subdividen en 5 x 5 localidades o grupos y éstos en 4 x 3 sublocalidades o celdas. La figura 1.1 aclara los niveles establecidos y la tabla 1.1 resume estos niveles y sus características.

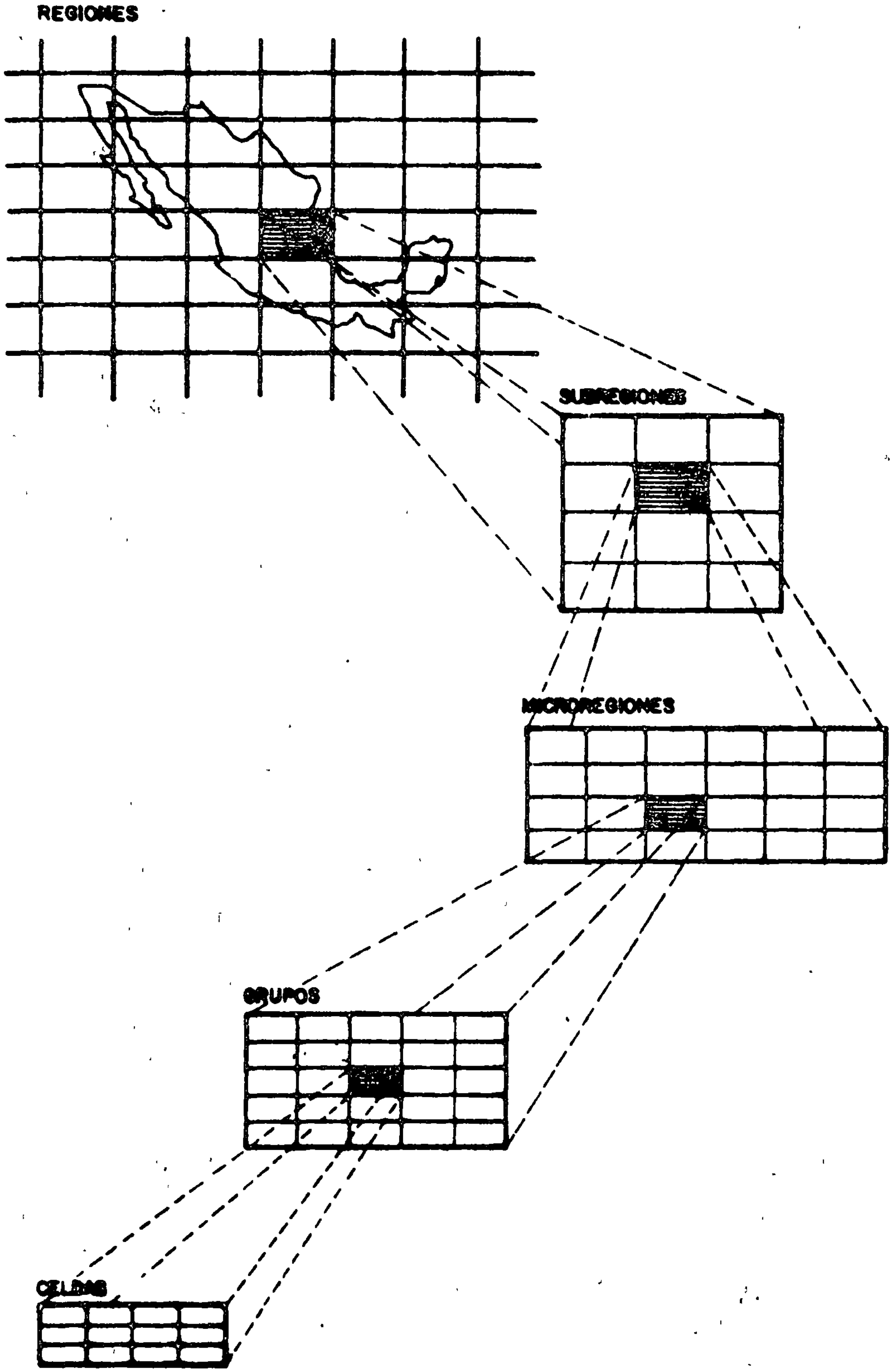


figura 1.1 NIVELES DE CONSULTA (2ª versión)

tabla 1.1 NIVELES DE CONSULTA (2a versión)

NIVEL	TAMAÑO		AREA APROX KM²	CONTIENE	EQUIVALE A
	LONG	LAT			
REGIONES	6°	4°	288 000 0	3 X 4 SUBREGIONES	12 CARTAS 1:250000
SUBREGIONES	2°	1°	24 000 0	6 X 4 MICROREGIONES	1 CARTA 1:250 000
MICROREGIONES	20'	15'	1000 0	5 X 5 GRUPOS	1 CARTA 1 50 000
GRUPOS O LOCALIDADES	4'	3'	34.7	4 X 3 CELDAS	
CELDAS O SULOCALIDADES	1'	1'	2.9		

Esta estructura de varios niveles agiliza las consultas y además cada subdivisión aumenta la precisión[5].

Una versión más antigua del BDC y que aún se aplica a las zonas de Celaya, Querétaro, Cortázar, Apaseo el Alto y Ojo Caliente, utiliza los niveles dados en la tabla 1.2 e ilustrados en la fig. 1.2[6].

tabla 1.2 NIVELES DE CONSULTA (1a versión)

NIVEL	NOMBRE	AREA APROXIMADA	CORRESPONDE A
4	SUBCUADROS	6 25 km²	
3	CUADROS	25 km²	4 SUBCUADROS
2	ZONA	1000 km²	48 CUADROS O, 192 SUBCUADROS
1	REGIONES	72 000 km²	72 ZONAS O,3456 CUADROS O, 13824 SUBCUADROS
0	REPUBLICA MEXICANA	2 500 000 km²	32 REGIONES O,2336 ZONAS O, 112128 CUADROS O, 448512 SUBCUADROS

5 Véase referencia 7
6 Véase referencias 9, 11

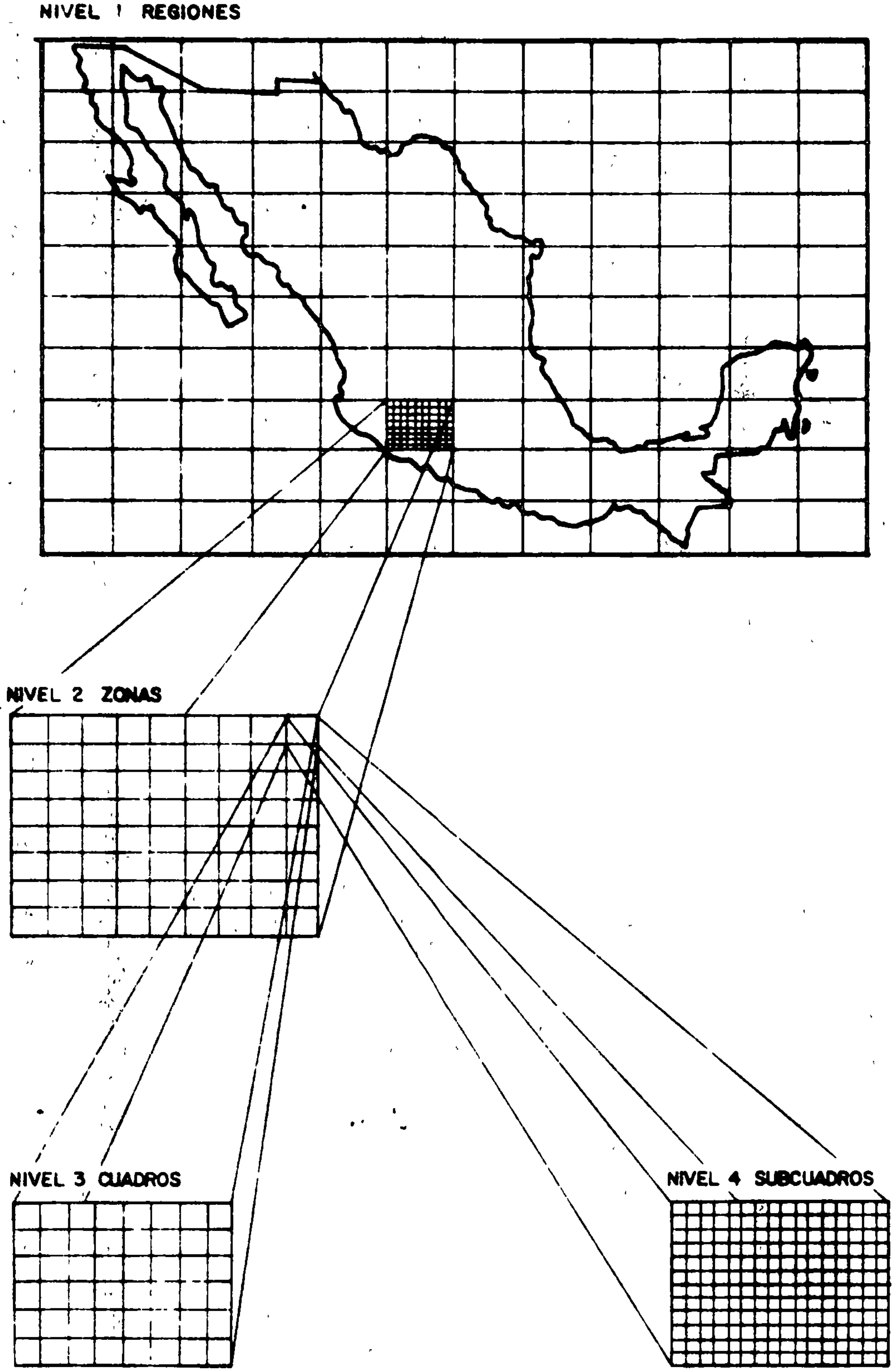


figura I.2 NIVELES DE CONSULTA (1ª versión)

Dentro del BDC a una carta se le llamó zona. Una zona tiene una área aproximada de 1 000 km^2; un nivel superior a éste, en área, es lo que en el BDC se llamó región y es el área comprendida por 72 zonas en un arreglo matricial de 9 x 8 zonas; esta división también aparece en las cartas de avance de CETENAL; cada región tiene una área aproximada de 72 000 km^2 y corresponde a 2° de latitud y a 3° de longitud. La República Mexicana queda contenida en un arreglo matricial de 12 x 10 regiones. Una zona se dividió en 48 cuadros, y un cuadro tiene una área aproximada de 25 km^2; cada cuadro se dividió en 4 subcuadros y cada subcuadro tiene una área aproximada de 6.25 km^2.

CAPITULO II

FORMAS DE SALIDA DE LOS RESULTADOS

Para explicar la forma de salida de los resultados por impresora, se hace referencia a resultados obtenidos con la primera versión del BDC que utiliza los niveles de zonas, cuadros y subcuadros. El significado de estas explicaciones puede hacerse extensivo a las impresiones de resultados obtenidos con la segunda versión del BDC aplicable al Estado de Aguascalientes, y que utiliza los niveles de microregión, localidad y sublocalidad. La única diferencia está en que los distintos arreglos matriciales de cuadros tienen distintas dimensiones.

Como respuesta a la búsqueda a un nivel determinado, la máquina produce en la impresora arreglos matriciales de cuadros de dos tipos: unos de cuadros claros y oscuros, y otros de diferentes tonos de gris (16 en total). Con el primer tipo de arreglos se puede distinguir en qué cuadros las preguntas son verdaderas y en qué cuadros son falsas[5].

[5] Véase figuras 2.1, 2.2, 2.3 pags. 18, 20

Los cuadros oscuros indican en ese lugar la pregunta se hizo verdadera, y los cuadros claros indican que en ese lugar la pregunta se hizo falsa.

Por ejemplo, con un predicado llamado "FRIJOL" se pregunta al BDC en qué zonas hay FRIJOL. Para esto se define la siguiente instrucción:

CALL, BUSCA (FRIJOL, 2, 79)

Y como respuesta se obtiene el resultado que muestra la figura 2.1 en la cual se ve un arreglo de matriz de 2 x 2 zonas. El área de una zona es igual al área de una carta CETENAL escala 1:50 000; las zonas corresponden a la región # 79.

Las zonas oscuras marcan los lugares en donde hay FRIJOL; en estos casos el predicado FRIJOL se cumple. Las zonas claras indican que en esas zonas no hubo FRIJOL, por lo tanto, el predicado buscado no se cumplió en esos casos; o lo que es lo mismo, que el conjunto de propiedades definido por el predicado FRIJOL no se encuentra en lugares donde aparecen zonas claras, (después se verá cómo definir al predicado FRIJOL). Resumiendo: zona oscura significa que sí hay lo que se busca y zona clara significa que no hay lo buscado.

Al buscar el predicado FRIJOL a nivel 3 (cuadros), se obtienen los resultados que muestra la figura 2.2. Esto permite tener una información más precisa que la del nivel 2, ya que dentro de este nivel, que en este caso es la carta de Celaya, se ve con más exactitud cuáles son los cuadros de esta zona que cumplen con el predicado FRIJOL; así se puede seguir haciendo a un nivel más pequeño, como se muestra en la figura 2.3. que corresponde al nivel 4 (subcuadros).

En síntesis, se resume que el predicado FRIJOL se hace verdadero en la zona 1, correspondiente a Celaya, de la región 79. A nivel de cuadro la función FRIJOL se hace verdadera en los cuadros: 5, 6, 13, 14, 17, 21, 22, 25, 26, 27, 28, 29, 30, 33, 36, 37, 40, 41, 42 y 45; esto permite tener más información que en el nivel anterior. Si se desea ampliar esta información se busca a nivel de subcuadro y se ve en qué subcuadro de cada cuadro se hace verdadero el predicado FRIJOL. Los subcuadros que cumplen la función FRIJOL son: los subcuadros **c** y **d** del cuadro 5, los subcuadros **b**, **c** y **d** del cuadro 6, el subcuadro **a** del cuadro 13, los subcuadros **b** y **d** del cuadro 14, los subcuadros **c** y **d** del cuadro 17, los subcuadros **c** y **d** del cuadro 21, los subcuadros **a**, **b** y **c** del cuadro 22, los subcuadros **a**, **b** y **c** del cuadro 25, los subcuadros **a**, **c** y **d** del cuadro 26, los subcuadros **c** y **d** del cuadro 27, los subcuadros **c** y **d** del cuadro 28, los subcuadros **a**, **b**, **c** y **d** del cuadro 29, el subcuadro **a** del cuadro 30, los subcuadros **a** y **c** del cuadro 33, los subcuadros **a** y **b** del cuadro 36, los subcuadros **a**, **c** y **d** del cuadro 37, los subcuadros **a** y **b** del cuadro 40, los subcuadros a y c del cuadro 41, el subcuadro c del cuadro 42 y el subcuadro a del cuadro 45. La tabla 2.1 explica detalladamente la nomenclatura usada[6].

Es preciso hacer notar la importancia que tienen los niveles de consulta en la estructura del Banco de Datos CETENAL.

Esto es, en el ejemplo de la función FRIJOL se buscó a nivel 4, qué subcuadros cumplieron con esta función.

Si no existieran los diferentes niveles de consulta, se tendría que buscar la función FRIJOL en aproximadamente 448 512 subcuadros, lo cual no sería eficiente ni práctico, aún desde el punto de vista del procesamiento por computadora. En cambio, usando los niveles de consulta, se pueden detectar rápidamente los subcuadros que cumplen la función FRIJOL en la forma siguiente:

1. Se busca a nivel 1 regiones[7] de toda la República Mexicana y los resultados se presentarán en un arreglo de matriz de 12 x 10; supóngase que solamente en la región 79 se cumplió la función.
2. Se busca a nivel 2 (zonas en la región 79) y se obtiene la figura 2.1, que muestra la zona 1 oscura; por lo tanto, la función

[6] Véase tabla 2.1 pag. 24

[7] Este nivel actualmente no se ha establecido debido a que el BDC todavía no cuenta con más de 72 cartas vaciadas para formar una región.

es verdadera en ese lugar.

3. Se busca en la zona 1 a nivel de subcuadro y se obtiene la figura 2.3.

Como se puede observar, con los pasos anteriores se obtuvieron en forma rápida los subcuadros que poseen la función FRIJOL, ya que cada nivel está constituido por su descriptor,[8] y no es necesario buscar en lugares donde no existe FRIJOL.

La figura 2.4 muestra una impresión de resultados para el nivel de sublocalidades de una microrregión de Aguascalientes.

El segundo tipo de arreglos, formados por cuadros de diferentes tonos de gris, se describe en el capítulo VI.

[8] Descriptor es el conjunto de información que existe en el BDC para cada nivel, véase referencia 1.

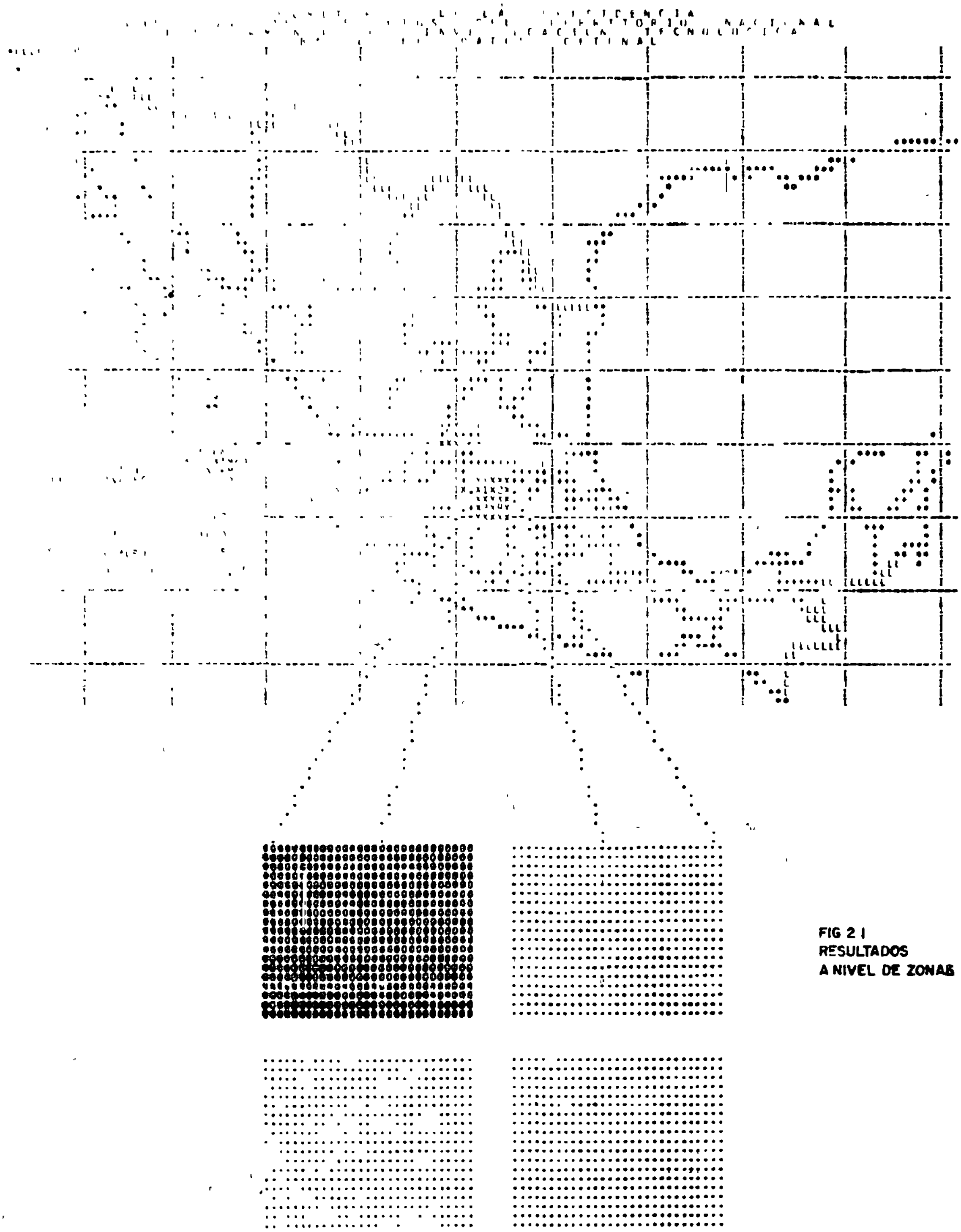


FIG 2.1
RESULTADOS
A NIVEL DE ZONAS

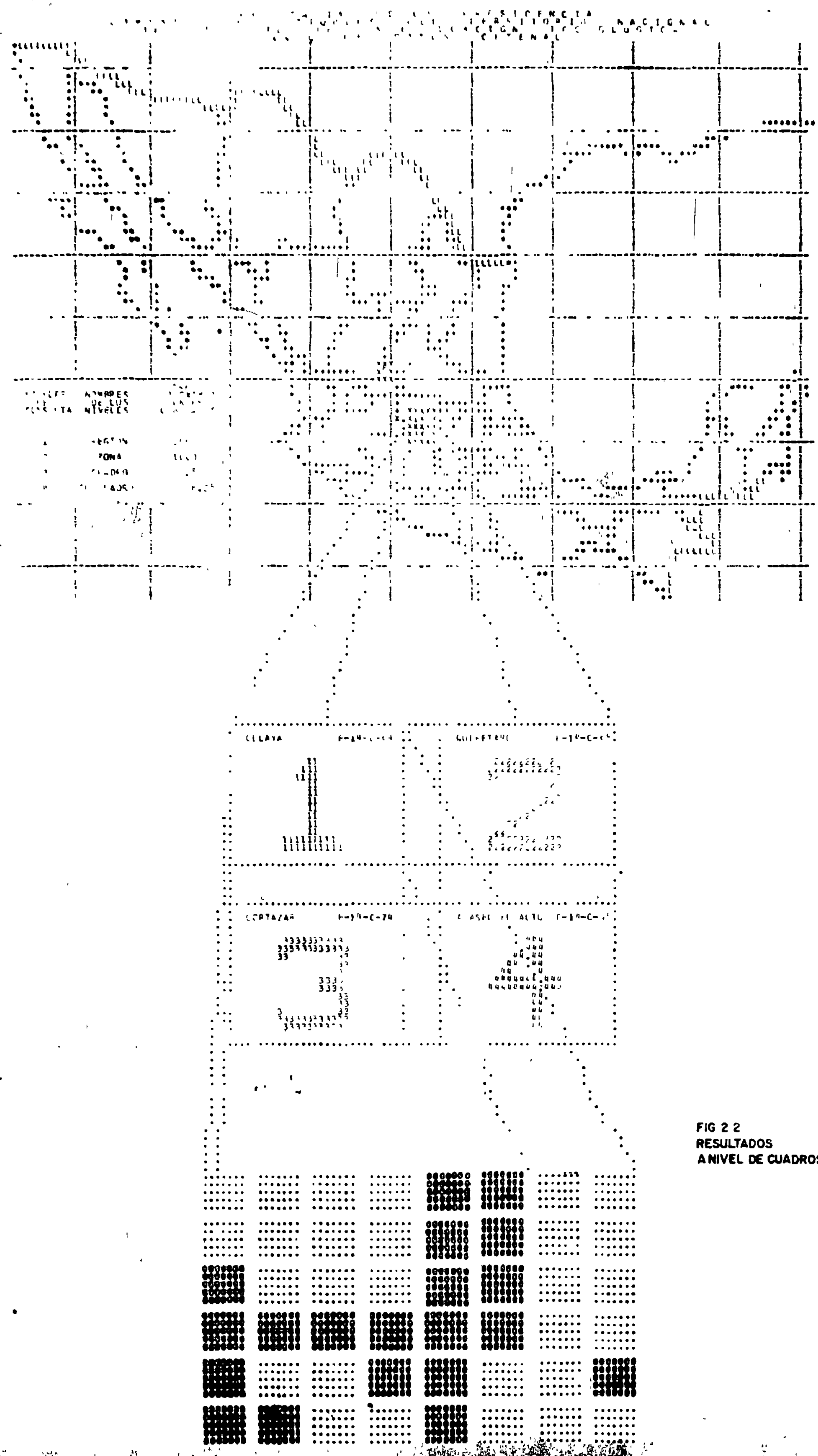


FIG 2 2
RESULTADOS
A NIVEL DE CUADROS

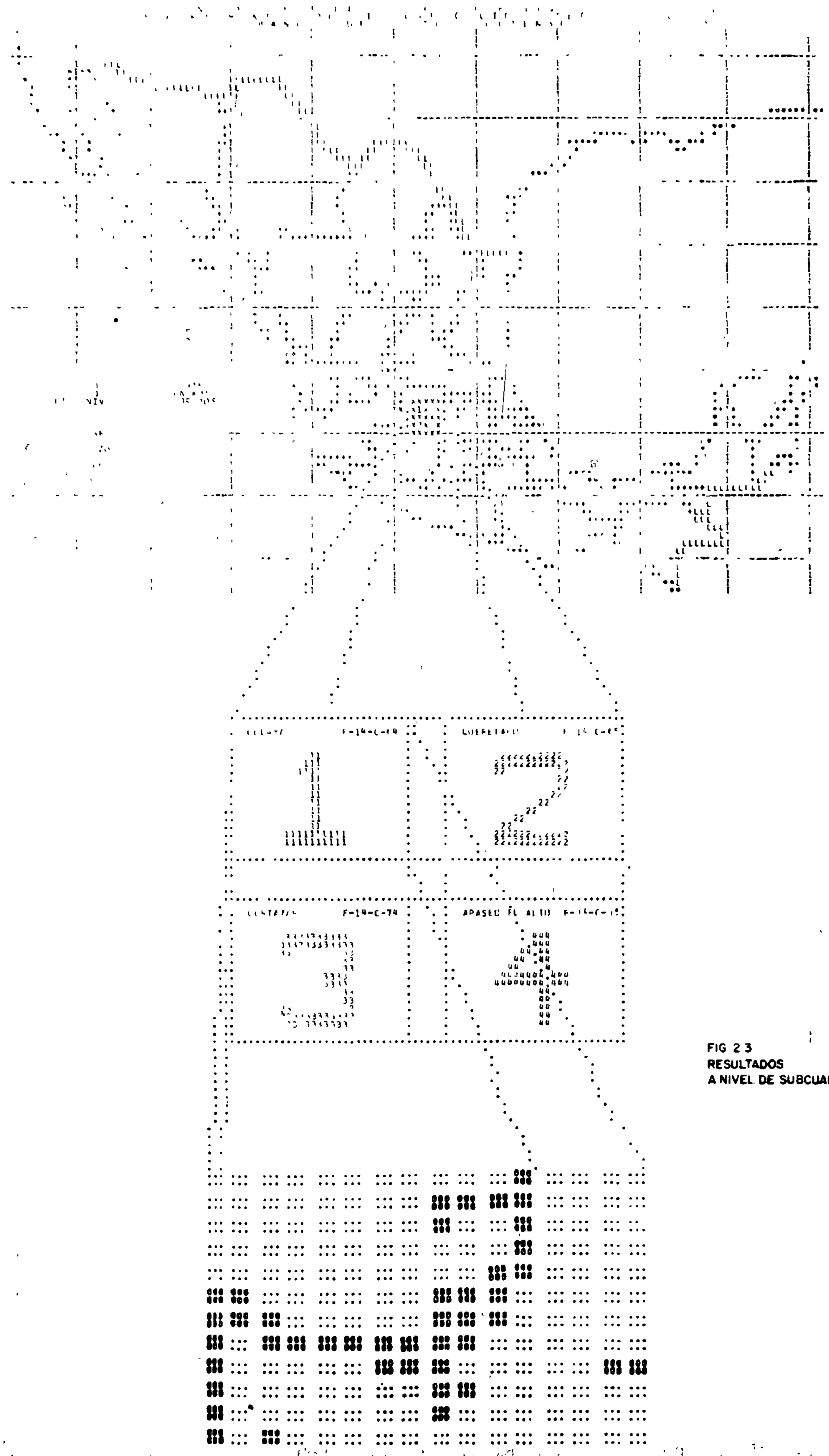


FIG 2 3
RESULTADOS
A NIVEL DE SUBCUADROS

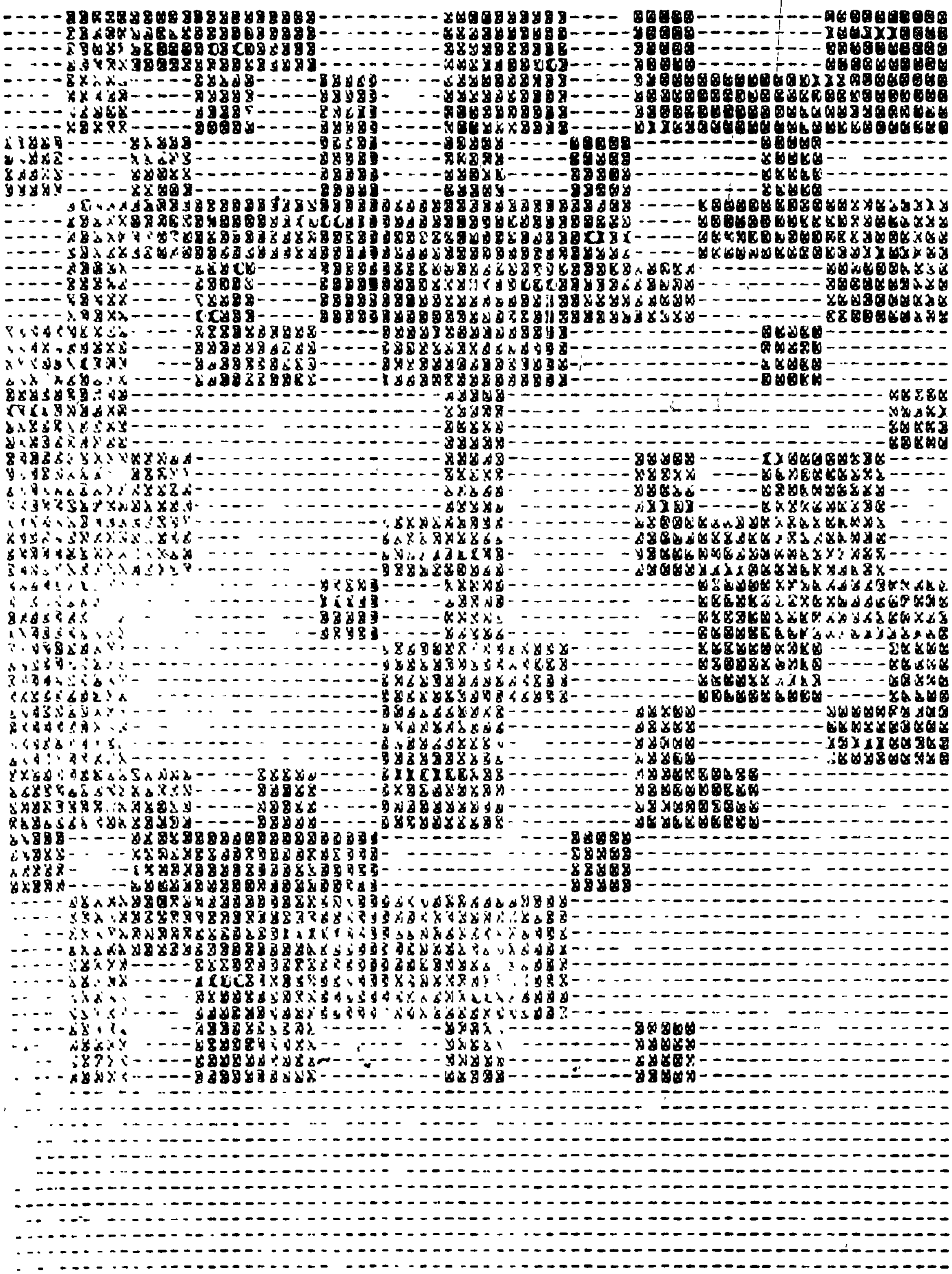


FIG. 2.4 RESULTADOS A NIVEL DE SUBLOCALIDAD

TABLA 2.1

NOMENCLATURA DE ZONAS, CUADROS Y SUBCUADROS

ZONAS

1	2	1. Celaya	2. Queré[illegible]
3	4	3. Cortázar	4. Apaseo el A[illegible]

CUADROS

1	2	3	4	5	6	7	[illegible]
9	10	11	12	13	14	15	[illegible]
17	18	19	20	21	22	23	[illegible]
25	26	27	28	29	30	31	[illegible]
33	34	35	36	37	38	39	[illegible]
41	42	43	44	45	46	47	[illegible]

SUBCUADROS

ab	ab	ab	ab	ab	ab	ab	[illegible]
cd	cd	cd	cd	cd	cd	cd	[illegible]
ab	ab	ab	ab	ab	ab	ab	[illegible]
cd	cd	cd	cd	cd	cd	[illegible]	[illegible]
ab	ab	ab	ab	ab	ab	ab	[illegible]
cd	cd	cd	cd	cd	cd	cd	[illegible]
ab	ab	ab	ab	ab	ab	ab	[illegible]
cd	cd	cd	cd	cd	cd	cd	[illegible]
ab	ab	ab	ab	ab	ab	ab	[illegible]
cd	cd	cd	cd	cd	cd	cd	[illegible]
ab	ab	ab	ab	ab	ab	ab	[illegible]
cd	cd	cd	cd	cd	cd	cd	[illegible]

En las siguientes secciones se describe el modo de consulta al BDC por medio de una terminal con pantalla de CRT.

Se cuenta con las siguientes funciones adaptadas a teleproceso.

PRO (CCCC, (MN, MY, IG), N)

PROP (CCCC, EN, N, M)

UNADE (CCCC, CCCC, (MN, MY, IG), N)

HAYVIA (CCCC, CCCC)

SERPRO (CCCC, N)

PUEBLO (CCCC)

SERPOB (N, (Y, O), S)

HABIT (N, N)

donde:

CCCC indica la clave utilizada según la tabla apropiada.

(MN, MY, IG) significa alguna de las funciones relacionales.

MN=MENORQ

MY=MAYORQ

IG=IGUALQ

Se utiliza sólo una de ellas.

N es un número que varía de 0 a 99999, según la pregunta.

EN=ENTRE

(Y, O) significa alguno de los operadores lógicos.

Y=AND

O=OR

Se utiliza sólo uno de ellos.

S Indica las subdivisiones de los servicios de la población y será un número que incluya a los dígitos 1, 2, 3, 4, 5, 6, 7 sin repetir ninguno según la equivalencia:

a = 1

b = 2

c = 3

d = 4

e = 5

f = 6

g = 7

Creación de una pregunta con base en las funciones anteriores.

- Cada función se construye con base en las especificaciones anteriores, sin incluir espacios intermedios.

PRO (6001, MY, 0) correcto

PRO (6001,   MY, 0) incorrecto

- La posición de cada función no tiene restricción.

- Los espacios en blanco entre funciones no tienen ningún efecto.

- Se utilizan los operadores lógicos Y, O, NO que son equivalentes a:

Y = AND

O = OR

NO = NOT

- Pueden usarse libremente los paréntesis.

- El número máximo de funciones es de 60, inclusive.

- El tamaño máximo de la pregunta es de 10 renglones en la pantalla.

Por medio de las teclas adecuadas se transmite la instrucción al nivel deseado, y se obtiene como resultado una matriz de caracteres donde * = cuadro oscuro y - = cuadro claro con los significados explicados anteriormente.

CAPITULO III

EJEMPLOS INTRODUCTORIOS

En el resto de este manual se usa la 1a. versión del BDC para describir y ejemplificar el lenguaje de usuario, y los programas de aprendizaje. En general, la 2a. versión funciona de igual manera, salvo pequeñas diferencias que se señalan cuando es necesario. Se utilizaron los cinco tipos de cartas (Edafológica, Topográfica, Geológica, Uso del Suelo y Uso Potencial), de cada una de las siguientes zonas; Celaya, Querétaro, Cortázar y Apaseo el Alto. La información que de estas cartas fue vaciada al BDC está descrita en las tablas del apéndice,[9] y corresponde aproximadamente a la información que un usuario extraería guiándose por las leyendas y explicaciones que aparecen en los márgenes de ellas.

Las funciones que dan acceso al BDC permiten describir los sitios (zonas, cuadros o subcuadros), que poseen determinada propiedad o cumplen con cierto predicado.

[9] Véase apéndice pag. 144

Se procede a dar algunos ejemplos sencillos, con el fin de mostrar la simplicidad de su uso. Cualquiera de las dudas que surjan en estos ejemplos, serán aclaradas en secciones posteriores. Por ejemplo, si se desea encontrar, a nivel 3, los cuadros en Celaya que poseen una superficie cultivada mayor del 40%, se define el predicado CULTI[10] mediante el postulado que define a CULTI como todo aquel lugar donde la propiedad 6104 sea mayor de 40, o sea, mayor del 40%. La tabla 1 del apéndice indica que la propiedad 6104 equivale a cultivos.

Una vez hecha esta definición, se procede a buscar CULTI a nivel de cuadro en Celaya, como sigue:

```
CULTI = PRO (6104, MAYORQ, 40)
CALL BUSCA (CULTI, 3, 1)
```

En la segunda versión, la segunda instrucción es como sigue:

```
CALL BUSCAS (CULTI, 1, MICROREG, 2, 3)
```

Las dos instrucciones anteriores, necesitan de ciertas declaraciones e instrucciones adicionales. Los programas completos son: programa que define el predicado CULTI.

```
LOGICAL FUNCTION CULTI (N)
LOGICAL PRO, MAYORQ
```

[10] El nombre dado no tiene importancia pero no debe pasar de 6 letras. Es conveniente usar nombres mnemónicos.

```
EXTERNAL MAYORQ
CULTI = PRO (6104, MAYORQ, 40)
RETURN
END
```

y programa principal que busca CULTI a nivel 3 en la zona 1, esto es, cuadros en Celaya, como se muestra en el listado 3.1.

```
LOGICAL CULTI
EXTERNAL CULTI
CALL BUSCA (CULTI, 3, 1)
STOP
END
```

El resultado se muestra en la figura 3.1 en la cual, se observan cuadros oscuros que son los que tienen más del 40% de su superficie cubierta por cultivos.

<u>Observaciones</u>.- Las declaraciones LOGICAL en los programas definen a CULTI, MAYORQ y PRO, como funciones que adquieren valores lógicos que pueden ser verdaderos o falsos. La (N) de CULTI (N) es una variable muda; las declaraciones EXTERNAL definen a CULTI y MAYORQ como nombres de funciones; las RETURN y END en el predicado, y STOP y END en el programa principal, terminan correctamente los programas. El usuario no necesita entender todo esto; sin embargo, es conveniente apuntarlo.

```
[illegible]CAL FUNCTION CULTI[illegible]
L[illegible]ICAL P[illegible]C,MAYOR[illegible]
[illegible]RNAL MAYOR[illegible]
[illegible]SEA SABER QUE CUADROS DE LA ZONA DE CELAYA TIEN N MAS [illegible]
[illegible] CIENTO DE SU SUPERFICIE CULTIVADA
[illegible]LT [illegible]C-61C4,MAYORC,40[illegible]
[illegible]URN
[illegible]
```

```
[illegible]CAL CULTI
[illegible]NAL CULTI
[illegible]LL [illegible]USCA(CULTI,3,1[illegible]
[illegible]
```

LISTADO 3.I ¿QUE CUADROS DE LA ZONA DE CELAYA TIENEN MAS DEL 40% DE LA SUPERFICIE CULTIVADA?

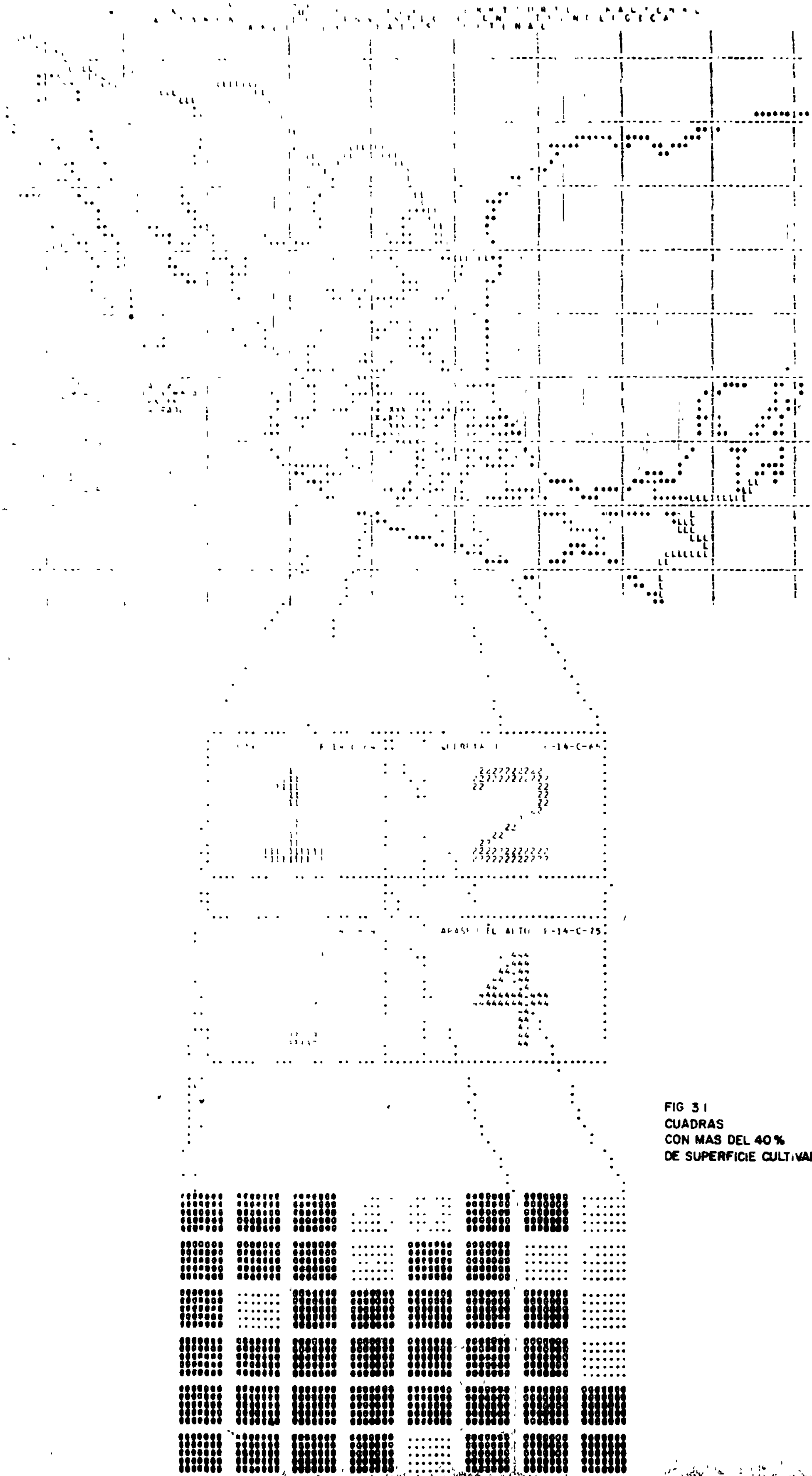


FIG 3 1
CUADRAS
CON MAS DEL 40%
DE SUPERFICIE CULTIVADA

Para buscar el mismo predicado CULTI a nivel de subcuadro, se dice la siguiente instrucción en el programa principal:

CALL BUSCA (CULTI, 4, 1)

Si hay interés por zonas en las que más del 60% de su superficie está cultivada o sean pastizales, propiedades 6104 y 6201, respectivamente, se dice:

ZONA = PRO (6104, MAYORQ, 60) .OR.PRO (6201, MAYORQ, 60)

Si se quiere hallar zonas que no tengan chaparrales, propiedad 6509, se dice:

NOCHAP = .NOT.PRO (6509, MAYORQ, 0)

Es decir, PRO (6509, MAYORQ, 0) define lugares donde sí existen chaparrales, y con un .NOT. anterior a la expresión, se niega esta propiedad.

El BDC, puede contestar también preguntas como las siguientes:

. ¿Qué lugares tienen pocos álamos y muchos pirules? Si se conviene en que poco signifique menos del 15%, y muchos signifique más del 70%, la función es:

ALAPIR = PRO (6312, MENORQ, 15) .AND. PRO (6319, MAYORQ, 70)

. ¿Qué lugares están comunicados por cualquier tipo de carretera? Si se usa la tabla 1 del apéndice,[11] se observa que las pro

[11] Véase apéndice pag. 144

piedades de la 801 a la 810 se refieren a diferentes tipos de carreteras. Con una de ellas es suficiente para comunicar el lugar por donde pasan. La función es:

COMUN = UNADE (801, 810, MAYORQ, 0)

Esta función se interpreta de la siguiente forma: al menos una de las propiedades 801 a la 810 debe ser mayor que cero, es decir, debe existir. Si se quieren conocer todos los lugares que están comunicados y que tienen agua almacenada, en presa o bordo, se dice:

COMAGU = UNADE (801, 810, MAYORQ, 0) .AND.
UNADE (401, 408, MAYORQ, 0)

Otras de las preguntas son:

- ¿Qué zonas se encuentran entre 1600 y 1800 metros sobre el nivel del mar?

RESUL = PROP (101, ENTRE, 1600, 1800)

- ¿En qué lugares hay palmares que ocupan el 25% de la superficie, y están a más de mil metros sobre el nivel del mar?

PALMAR = PRO (6501, MAYORQ, 25) .AND. PRO (106, MAYORQ, 1000)

- ¿Crecerán los nopales sobre suelos de tipo Gléyico?

Se define en que: "crecen los nopales" significa que más del 30% del suelo está cubierto de nopalera, propiedad 6507, y que un suelo es de tipo "Gléyico", cuando la superficie ocupada por la propiedad 3319, suelo Gléyico, es mayor de 70%. Entonces:

NOPG = PRO (6507, MAYORQ, 30) .AND. PRO (3319, MAYORQ, 70)

Se pueden hacer preguntas arbitrariamente largas. En el ejemplo que está a continuación, se buscan lugares de atractivo turístico; éstos pueden ser lugares donde haya cataratas, propiedad 601, o bosques, propiedades 6301 y 6302, o ríos, propiedad 503 y que estén en lugares no muy altos, esto es, a menos de 2 500 m snm; o bien aquellos lugares altos y con volcanes propiedad 2553 o lugares en donde haya más de 4 manantiales termales propiedad 2604, que estén bien comunicados por algún tipo de camino o carretera, propiedades 801 a 810 y que se hallen en terrenos no montañosos, propiedades 5802 y 5803.

Se puede proceder por partes y definir:

UNO = PRO (601, MAYORQ, 0) .OR. UNADE (6301, 6302, MAYORQ, 0)

.OR. PRO (503, MAYORQ, 0) .AND. PRO (105, MENORQ, 2500)

DOS = PRO (106, MAYORQ, 2500) .AND. PRO (2553, MAYORQ, 0)

TRES = PRO (2604, MAYORQ, 4) .AND. UNADE (801, 810, MAYORQ, 0)

.AND. .NOT. UNADE (5802, 5803, MAYORQ, 0)

Finalmente se dice:

TURIS = UNO. OR. DOS. OR. TRES

Otro ejemplo es el siguiente:

- ¿En qué lugares de Celaya se pueden explotar minas ?

En este ejemplo, se usan las funciones y rutinas de "aprendizaje".

Para proponer las minas que se pueden explotar, se necesitan experiencias anteriores. Los programas pueden adquirir estas experiencias de minas ya existentes en cualquiera de las cuatro zonas.

Se construye la función que busca minas, propiedad 2605, de la siguiente forma:

MINA - PRO (2605, MAYORQ, 0)

También se construye la función de "aprendizaje", en la cual incluye una variable de tolerancia; en este caso del 20%. Puede suceder que los programas de "aprendizaje" no encuentren nada debido a que la variable de tolerancia es muy severa, entonces se puede hacer mayor la tolerancia y así encontrar algunos lugares de los que se buscan.

Para buscar minas en CORTAZAR, se da la siguiente instrucción:

CALL BUSCA (MINAS, 3, 3)

A continuación, el aprendizaje de minas se lleva a cabo con la siguiente instrucción:

CALL LEARN (3, 2605)

Con las muestras (de minas) obtenidas en Cortázar se pueden proponer minas en Celaya como se describió anteriormente.

Según lo que se ha visto, también se puede "aprender" de otras zonas, y así tener mayor confiabilidad en los resultados.

También se puede determinar la confiabilidad con que se está tra

bajando, es decir, si se comparan las minas propuestas con las minas existentes, y se ve que se proponen 4 minas en lugares en donde existen 5, se está trabajando con una confiabilidad del 80%, esto se hace en la forma siguiente:

CALL CONFIA (MINAS, 3, 1)

La rutina llamada por la instrucción anterior, dice la confiabilidad con que se está trabajando.

NOMENCLATURA USADA EN LOS EJEMPLOS ANTERIORES.

PREDICADO: Pregunta construida por el usuario.

FUNCIONES RELACIONALES: MENORQ, MAYORQ, DIFERE, IGUALQ, ENTRE, MAYORI, MENORI, NOIGUA.

FUNCIONES LOGICAS: PRO, PROP, UNADE, CERCA, SERPOB, SEPRO, PUEBLO HAYVIA Y APRE11.

CAPITULO IV

DESCRIPCION DE LAS FUNCIONES RELACIONALES

Las funciones relacionales (MAYORQ, MAYORI, MENORQ, MENORI, ENTRE, IGUALQ, NOIGUA, DIFERE) se usan dentro de las funciones lógicas PRO, PROP y UNADE, para delimitar los valores de una propiedad.

Son útiles para construir el predicado que finalmente se utiliza es decir, el que contiene todas las propiedades que interesan. PRO se usa cuando se necesitan tres argumentos y PROP cuando se requieren cuatro.

Regla 1.- Cuando en la definición de estos predicados se usa alguna de las funciones PRO, PROP, UNADE, o alguna de las funciones relacionales, tal función antes de ser usada debe declararse que es una función lógica, ejemplo:

Supóngase que se quiere hallar los lugares donde hay agricultura de temporal nómada, o temporal permanente. Esto es, cualquiera de las propiedades 6002 ó 6003.

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

MANEJO DE SISTEMAS DE INFORMACION GEOGRAFICA EN PLANEACION

TEMA: BASE PARA PLANEACION URBANA DISTRITO FEDERAL.

PROF. DR. JORGE CAIRE LOMELI.

Noviembre ,1977.

INSTITUTO PANAMERICANO DE GEOGRAFIA E HISTORIA

ORGANISMO ESPECIALIZADO DE LA O E A

REVISTA CARTOGRAFICA

NUMERO 2[illegible] MEXICO JUNIO 1976

INSTITUTO PANAMERICANO DE GEOGRAFIA E HISTORIA

REVISTA CARTOGRAFICA

NUMERO 29 MEXICO JUNIO 1976

SUMARIO

BASE PARA LA PLANEACION URBANA.
EL CASO DEL DISTRITO FEDERAL

Jorge Caire Lomelí. *

RESUMEN

El crecimiento natural y social de la población que cada vez registran mayores tasas, generan nuevos problemas, así como la incrementación de los existentes. Estos resultan ser agudizados en los centros urbanos del país, siendo el Distrito Federal donde predomina este hecho geográfico

Las medidas que deben tomarse para la resolucion de este problema tendrían que estar basadas en una planeación adecuada e integrada, que tenga origen en un diagnóstico y para que éste, pueda cumplir con su funcionalidad, debe acudirse a una cartografía integrada en el marco urbano–rural El aprovechamiento de la información cartográfica urbana y rural que se desarrolla en las dependencias oficiales sería de gran valor si se tomara en cuenta que además de sus objetivos particulares se integraran a un patrón común nacional.

La presente investigación trata la forma de llevar a cabo la integración cartográfica, utilizando el proyecto adoptado a nivel nacional, teniendo en cuenta los requerimientos cartográficos urbanos en el Distrito Federal, Capital de la República Mexicana

1.– INTRODUCCION

El continuo y acelerado proceso de urbanización que tiene su origen a partir de la Revolución Industrial, hace que la ciudad sea un interesante tópico de estudio, en el que es necesario el análisis e interpretación de los fenómenos que en ella ocurren situándola en un marco especial determinado

* El Sr Jorge Caire Lomelí es Cartógrafo de Petróleos Mexicanos y profesor de Cartografía de la Universidad Nacional Autónoma de México y del Instituto Politécnico Nacional Es ingeniero Topógrafo e Hidrógrafo con Maestría y Doctorado en Geografía en la Universidad Nacional Autónoma de México.

El analisis y representación cartográfica de estos fenómenos hace posible visualizar la problemática actual del medio urbano, para con esto se puedan presentar soluciones al problema de desarrollo desarmónico de la ciudad y obtener así una planeación integral de ella en beneficio de sus habitantes

Al realizar los estudios urbanos dentro del contexto regional es imprescindible tener una serie de mapas que tengan un común denominador para ser interpretados por los diferentes hechos geográficos y plantear una serie de políticas a través del diagnóstico que presente; para el logro de tales objetivos se ha propuesto utilizar una proyección cartográfica que además de cumplimentar los requisitos del área urbana y del área rural puedan ser interconectados en cualquier circunstancia y sea posible aprovechar los estudios que con un objetivo sectorial se hicieron al ámbito regional o nacional

II.– NECESIDAD DE UNA INTEGRACION CARTOGRAFICA

La vida urbana rebasa ampliamente el marco de la ciudad y se proyecta sobre las comunidades vecinas de modo que muchas de las actividades urbanas encuentran su explicación en el medio rural El hombre de la ciudad depende en gran medida del hombre rural, es por esto que cualquier estudio que se realice de la ciudad debe ser efectuado no como un hecho aislado sino que interrelacionado al resto de los fenómenos del paisaje geográfico

Entendido el mapa urbano como la base de todas las expresiones de la vida urbana se hace necesario la integración de su representación con el resto de los mapas regionales

Muchos son los aspectos que fundamentan esta situación ya que varios son los nexos urbanos–regionales

1.– Las comunicaciones

Los vínculos de la ciudad con su hinterland dependen en primera instancia de los medios de comunicación existente, su existencia facilita las comunicaciones del hombre rural En aquellos lugares en que el sistema vial es inoperante o deficiente se hace necesario su incrementación o trazado debiendose por lo tanto diseñar una carta para este fin Como los caminos y carreteras conducen al hombre hacia áreas externas de la ciudad, éstas deben estar trazadas de modo que empalmen con las carreteras ya trazadas, por lo tanto el empalme de un plano a otro debe ser exacto evitanto de este modo superposiciones o errores de trazado

Estas vías de comunicación deben tener una integración exacta con las circundantes dado que son muchos los nexos de comunicación que por ellas se realizan, tales como abastecimiento, difusión de periódicos y revistas, esparcimiento, etc.

2.– Funciones administrativas:

La ciudad es un centro administrativo por excelencia, en ella se concentran los poderes y centros administrativos, parten las normas legislativas que son aplicadas en toda su área de influencia. El control de su aplicación puede efectuarse a través del señalamiento cartográfico, vale decir, con la localización en forma homogénea de centrales administrativas que efectúen dichas labores necesitandose por lo tanto mapas a una misma escala y proyección para tener una expedita realización de estas actividades

3.— Funciones comerciales:

La ciudad es la oficina de negocios de la región, de ella parten los comerciantes con sólo sus muestrarios y diseños para ser vendidos en otra área. La representación comercial debe ser diseñada en un mapa que permita fácilmente realizar la integración de las transacciones.

4.— Expansión del área urbana:

En caso de seguir extendiéndose el radio urbano de un poblado éste debe seguir con los mismos lineamientos de la ciudad origen, es decir su mismo plano, calles, banquetas e infraestructura. El trazado de estos aspectos requiere de una continua cartografía para no incurrir en errores. Al presentarse el fenómeno de la conurbanización, la anexión de nuevas áreas antes rural deben realizarse en forma ordenada de manera que se vayan ocupando los terrenos de menor valor agrícola, estas características solo pueden ser cualificadas y cuantificadas en base a un mapa con las mismas características que el núcleo en expansión, es decir la ciudad primitiva.

5.— Uso del suelo:

El conocimiento del uso del suelo del área cercana al medio urbano con fines de expansión debe ser previamente evaluado sobre una carta o mapa, la que como punto de referencia debe contener las mismas características de las que señale las del uso del suelo urbano.

6.— Jerarquización de ciudades e interconección de ellas.

Un país que desee efectuar una jerarquización de sus ciudades señalándole diversas funciones a cada una de ellas, debe disponer antes que nada de un mapa que las localice, señale su extensión y sus caracteres principales en una carta o varias cartas con las mismas características de modo que se pueda fácilmente efectuar su comparación y evaluación, esta política es de gran significado para la planeación de un país que da vital significado a las ciudades.

7.— Planeación integral.

Desde el punto de vista de la planeación, se requiere contar con una cartografía básica general y no es posible concebir la planeación sectorial sin tomar en cuenta la nacional y general.

Todo plan ya sea a corto, mediano o largo plazo requiere estar enlazado a las políticas generales del país, así la planeación industrial no puede fijar sus metas y objetivos en forma aislada sino que se debe enmarcar a las políticas generales. De este modo la planeación urbana debe estar también de acuerdo a este principio y no figurar como ente aislado.

Toda la planeación debe apoyarse en una base cartográfica por lo que analizadas las características anteriores no se podría efectuar una planeación general sino se contara con cartas integradas a un tipo de proyección que presentara caracteres homogéneos.

La proyección cartográfica óptima para ser utilizada en los mapas urbanos debe reunir básicamente dos condicionantes:

1.— Que sea posible llevar una integración cartográfica a nivel nacional a diversas escalas mediante las reducciones respectivas, es decir con mapas urbanos catastrales, reguladores, directores de servicios, equipamiento y uso del suelo, etc. y mapas rurales (topográfico, geológico uso del suelo, etc.).

2.– Que los mapas r u or .les o también llamados rurales sean integrados a nivel nacional, así como para las áreas urbanas a escalas grandes.

Estos dos aspectos se han venido elaborando en forma aislada, en la primera utilizando el plano local y en la segunda la proyección Universal Transversa de Mercator.

Ante estas circunstancias el autor ha investigado una solución para estar en condiciones de establecer la integración cartográfica urbano–rural (escalas grandes con las chicas).

III.– ADAPTACION A LA PROYECCION UNIVERSAL TRANSVERSA DE MERCATOR.

La proyección cartográfica adoptada a nivel nacional para escalas chicas es la Universal Transversa de Mercator, condición secante tipo Gauss – Kruger con un factor de escala de 0.9996 para los meridianos centrales (M C) teniendo bandas de contacto a 180,000 metros de cada uno de ellos (factor de escala igual a 1), los meridianos centrales son los únicos que se confunden con las ordenadas de la cuadrícula Universal Transversa de Mercator, siendo el resto paralelas y por lo cual no convergen, de tal forma que cada meridiano central cubre una zona de seis grados y la cuadrícula dos de divergencia como máximo Además esta proyección trabaja para estos usos con las distancias reducidas al nivel del mar Ante estas características no es posible que se puedan elaborar los mapas urbanos que se caracterizan por sus escalas grandes ya que tanto el factor de escala como la reducción de distancias al nivel del mar acusarán diferencias de consideración y serán más críticas cuando se conjuguen estas, como acontece en el Distrito Federal en que se tiene el meridiano central de 99° de longitud oeste y una altitud de 2,300 metros.

Para estar en condiciones de que la citada proyección sea óptima en la elaboración cartográfica urbana integrada, se tendría que tener en cuenta los aspectos:

1 – Apoyar todos los levantamientos urbanos en cuanto al azimut y a la representación cartográfica al meridiano central correspondiente.

2 – Establecer como origen de coordenadas las mismas que rigen en la proyección cartográfica nacional, para lo cual deben ser transformados los apoyos geodésicos que intervengan en la proyección Universal Transversa de Mercator.

3.– Obtener el parámetro común por el cual todo el mapa urbano mediante procesos de reducción o amplificación cumple con todas las especificaciones de la proyección cartográfica Universal Transversa de Mercator.

Los dos primeros incisos involucran todo el desarrollo interno del mapa urbano, el cual se levantara con las precisiones que demanden sus objetivos y se tratará como una simple proyección ortogonal, es decir por los métodos tradicionales en la cartografía urbana. Al cumplirse estos dos aspectos, el tercer inciso es un valor registrado en la parte marginal del mapa urbano para establecer su integración a cualquier otro mapa de tipo rural o regional.

AZIMUT EN LA PROYECCION UNIVERSAL TRANSVERSA DE MERCATOR.

Los azimutes se determinan a través de las direcciones (ángulo entre una línea o plano y una l...

de referencia o plano arbitrariamente seleccionado) observadas en los levantamientos de campo. El proceso de promediar los puntos observados, de corregir las excentricidades de estaciones y de ajustar los ángulos, se hacen por procedimientos convencionales de acuerdo con los levantamientos. La transformación de las direcciones observadas en direcciones que se puedan usar en la Proyección Universal Transversa de Mercator, es una operación de cálculo, así como la transformación inversa de valores en la proyección para pasar a los del campo.

El azimut de una dirección es el ángulo diedro formado por el plano meridiano que pasa por el lugar y el plano vertical que contiene a la dirección dada, se mide en sentido retrógrado de 0° a 360° y de acuerdo con el meridiano a que esté referido será el tipo de azimut. Los azimutes de cuadrícula están dados con origen en el Norte de ella.

Los símbolos que se emplean para las diferentes clases de azimut son:

(t) = Azimut plano.
(T) = Azimut geodésico proyectado.
(α) = Azimut geodésico.
(α') = Azimut geodésico inverso.

Los azimutes (t) y (T) se refieren frecuentemente sin distinguirlos a los de cuadrícula en los levantamientos de poca precisión y con distancias cortas. Para los levantamientos de mayor precisión con líneas largas, es necesario distinguir el azimut (t) del (T) porque llevan valores numéricos diferentes de proporciones significadas. El azimut geodésico aparece como una línea curva en la Proyección Universal Transversa de Mercator (excepto cuando coincide con el Meridano Central) y determina un ángulo con la meridiana geodésica que también aparece en la proyección como línea curva.

Las correcciones que se aplican a los valores angulares son:

$\Delta\alpha$ = Simboliza la convergencia de meridianos y se aplica en la ecuación del azimut inverso:

$$\alpha' = \alpha + 180^\circ \pm \Delta\alpha$$

C = Declinación de cuadrícula, causada por la convergencia de meridianos y consiste en la separación de la línea Norte–Sur de cuadrícula con la línea meridiana; se aplica en la ecuación:

$$(T) = \alpha \pm C + 180^\circ$$

t–T= Torsión que es la diferencia angular entre el azimut plano (t) y el azimut proyectado (T) que salen del mismo punto y su valor es llamado corrección por torsión.

Cálculo del azimut plano (t). Por definición el azimut plano de una línea AB es el ángulo medido en sentido retrógrado a partir del Norte de cuadrícula a una línea dada AB, que es recta en la proyección y curva en la superficie terrestre. En función de sus coordenadas el azimut plano (t) de A hacia B se determina por la ecuación:

$$\tan(t) = \frac{X_2 - X_1}{Y_2 - Y_1} = \frac{\Delta X}{\Delta Y}$$

X_1 y Y_1 ' coordenadas del punto inicial A

X_2 y Y_2 ' coordenadas del punto final B

El resultado de esta ecuación es un ángulo simbolizado por la que estará determinado en cualquier cuadrante, y para referirlo al Norte de cuadrícula y en sentido retrógrado se hace por medio de los signos que se obtengan para ΔX y ΔY.

El azimut plano (t) también puede ser determinado conociendo el azimut geodésico y una referencia de cuadrícula aproximada, utilizando la ecuación $(T) = \alpha \pm C + 180^\circ$ en donde C es la convergencia calculada por medio de la ecuación:

$$C = (XII)\, p + (XIII)\, p^3 + C_5 = (XV)\, q - (XVI)\, q^3 + F_5$$

en donde las funciones XII, XIII, XV, XVI y F_5 están calculadas en los manuales de la proyección Universal Transversa de Mercator.

La forma de obtener el azimut geodésico proyectado, es por medio del azimut plano al que se le aplica la correción llamada de torsión (t–T), este valor se encuentra a partir de la ecuación:

$$(t-T) = (-\Delta Y)(2X'_1 + X'_2)\ 6.875\,5\ (XVIII)\ 10^8$$

El valor del azimut plano (t) se determina en función de las coordenadas de cuadrícula, la función XVIII establecida en los manuales y X'_1 ' X'_2 son las abscisas a partir del meridiano central.

Conociendo los valores del azimut plano (t) y la corrección de torsión (t–T) el azimut geodésico proyectado se obtiene por la ecuación:

$$T = t - (t - T)$$

Origen de coordenadas en la Proyección Universal Transversa de Mercator.

El origen de este sistema, está con el valor de X = 500 000 metros para el meridiano central, que para el Distrito Federal le pertenece el de 99° al oeste de Greenwich eY = 0 metros en el Ecuador.

Para referirse a este origen es necesario conocer las coordenadas geodésicas de un vértice (φ, λ) y llevar a cabo su transformación por medio de las ecuaciones:

$$Y = (I) + (II)\, p^2 + (III)\, p^4 + A_6$$

$$X = 500\,000 \pm X'$$

$$X' = (IV)\, p + (V)\, p^3 + B_5$$

Las funciones I a V así como los términos A_6 y B_5 se encuentran en los manuales correspondientes.

Parámetro común Se obtiene mediante el factor de escala por la reducción al nivel del mar y al estar consignado como información marginal, es posible transformar toda la cartografía urbana a la regional, puesto que la primera tiene que efectuar sus levantamientos al nivel medio de la altitud del area considerada y con un factor de escala igual a la unidad.

El factor de escala se obtiene mediante las ecuaciones:

$$\frac{1}{K_t} = \frac{1}{6}\left(\frac{1}{K_1} + \frac{4}{K_3} + \frac{1}{K_2}\right)$$

$$K = K_0 [1 + XVIII\, q^2 + 0.00003\, q^4]$$

K_0 = 0.9996 factor de escala en el meridiano central.

q = 0 000001 X"

X" = Abcisa al meridiano central

K = Factor de escala de un punto determinado

XVIII = función establecida en los manuales

K_1 = factor de escala del punto inicial de la línea

K_2 = factor de escala del punto extremo final de la línea

K_3 = factor de escala del punto medio de la línea.

K_t = factor de escala de la línea considerada.

La distancia reducida al nivel del mar. Todos los levantamientos geodésicos establecen esta singuridad con el objeto de poder apreciar las tolerancias establecidas a lo largo de las diferentes altitudes que se atraviesen, de tal forma que para referir una línea base medida y corregida se reduce al nivel del mar por medio de la ecuación:

$$d = D\left(1 - \frac{h}{R} + \frac{h^2}{R^2}\right)$$

D = distancia medida y corregida

h = altitud media.

R = radio medio.

d = distancia reducida al nivel del mar.

El Distrito Federal está localizado entre los paralelos de latitud norte 19° 04′ a 19° 36′ y entre los meridianos 98°58° a 99°22′ de longitud oeste del meridiano de Greenwich con una superficie de 1,482 kilómetros cuadrados, altitud de 2,278 metros sobre el nivel del mar, distancia norte-sur 55 kilómetros y distancia este-oeste 52 kilómetros

De acuerdo con su situación geográfica al aplicarle el factor de escala arroja un valor de 0,999639 y por reducción al nivel del mar 0.999615. La reducción total es de 0,999254, es decir su parámetro común, por el que tiene que multiplicarse para estar en la proyección Universal Transversa de Mercator.

Por otra parte, no es posible que los estudios urbanos se hagan a partir de una cartográfica a base de la proyección citada, puesto que para el Distrito Federal si se considera la línea norte—sur se encuentra un error relativo de 1:1380, que no justifica las tolerancias especificadas, puesto que en 55 kilómetros se cometerían 41 metros de error

Al basar el error gráfico de un mapa, (la apreciación del ojo humano 0,25 milímetros) las tolerancias van siendo más bondadosas a partir de que las escalas van siendo más chicas, puesto que los errores pequeños se van absorbiendo, así se detecta: a la escala 1:500 se aprecian 0.125 metros; en escala 1:5000, 1.25 metros y en escala 1:10 000, 2.5 metros.

Es de considerarse la labor efectuada hasta la fecha dentro de las actividades cartográficas urbanas en el país y que para llevar a cabo esta transformación se requiere de grandes esfuerzos, pero que al realizarla se tendrá como consecuencia la base para la planeación integrada urbano—rural.

IV.– CONCLUSIONES

En forma de conclusión se deduce que si una área urbana (escala grande) se trabaja en proyección ortogonal, con las condicionantes siguientes, puede ser integrada al área rural (escala chica)

a.– Disciplinada al azimut de origen (meridiano central).

b.– Que las distancias se reducen al horizonte exclusivamente, o se proyectan al nivel medio y se tomen los orígenes de coordenadas de la proyección cartográfica adoptada.

c.– Se utilice el parámetro común, como dato de aplicación para integrarla a la cartografía regional.

Habiéndose demostrado la necesidad de efectuar la integración cartográfica urbano—regional, se concluye la necesidad de realizar estudios que conduzcan a evitar el adolecimiento de esta situación en la República Mexicana por lo que el desarrollo de esta metodología significa un avance técnico en esta disciplina que conduce a obviar los problemas de regionalización existentes para así poder efectuar la planeación del desarrollo sobre una misma base cartográfica evitando así los trabajos aislados de detalle.

Se fundamenta esta proposición en el hecho que la ciudad no puede ser considerada como un fenómeno espacial aislado sino integrado e interrelacionado con los demás fenómenos del espacio geográfico.

De la misma manera la realización de esta integración cartográfica conducirá a disminuir y simplificar las inversiones económicas que para el desarrollo de este aspecto se realizan en forma aislada, las cuales muchas veces inclusive llegan a estar duplicadas con la ruinosa pérdida de esfuerzos humanos y económicos lo que incide directamente en el tiempo.

REFERENCIAS

1. **BLACHUT, T.J.** Technical and organizational problems in urban surveying and maping. 1969.

2. **BLACHUT, T.J.** Confección integral de mapas a escala grande. 1971.

3. **GONZALEZ, G.E.** Los estudios de planeación y uso del catastro urbano. El caso del Distrito Federal. 1975.

4. **BLACHUT, T.J.** Cadastre, various functions characteristic tecniques and the planning of a modern land record system. 1974.

5. **GEORGE, PIERRE.** Geografía urbana. 1974.

6. **UNIKEL, LUIS.** La dinámica del crecimiento de la Ciudad de México. 1974.

7. **CAIRE, L.J.** La proyección cartográfica para Petróleos Mexicanos. 1974.

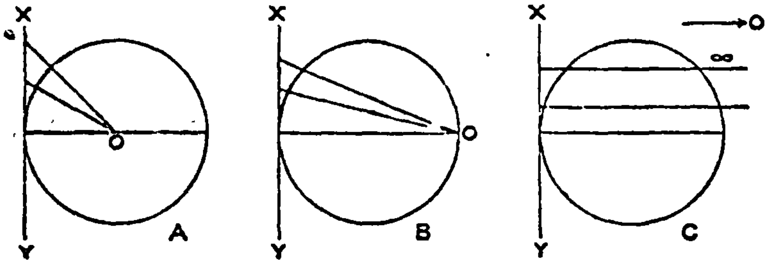
centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

MANEJO DE SISTEMAS DE INFORMACION GEOGRAFICA EN PLANEACION.

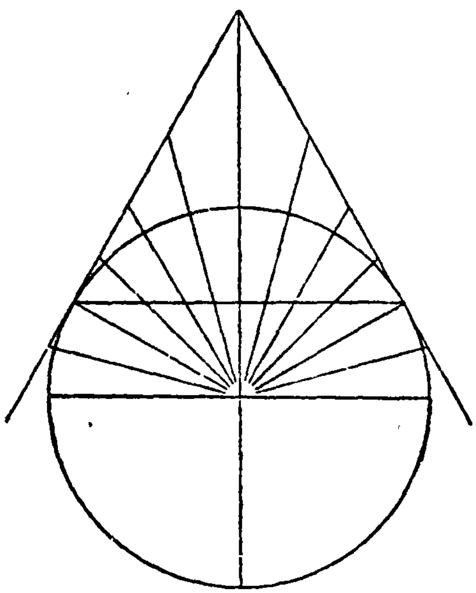
TEMA: P R O Y E C C I O N E S.

PROF. DR. JORGE CAIRE LOMELI.

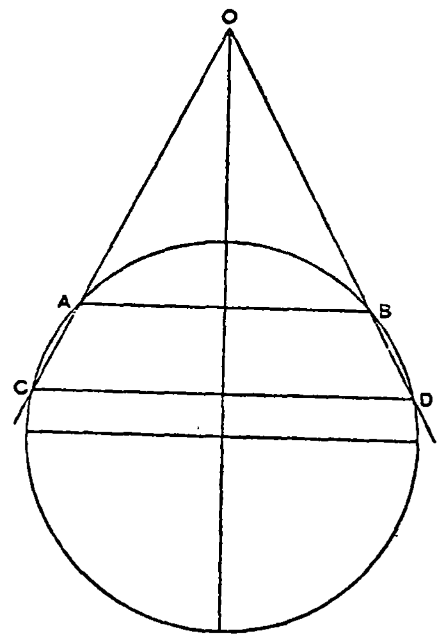
Noviembre ,1977



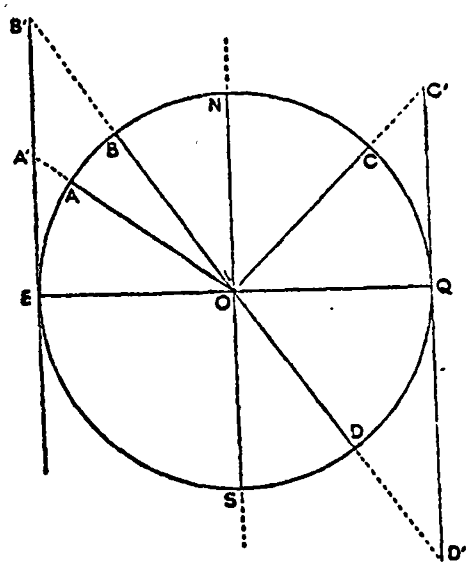
Simple Perspective Zenithal Projections



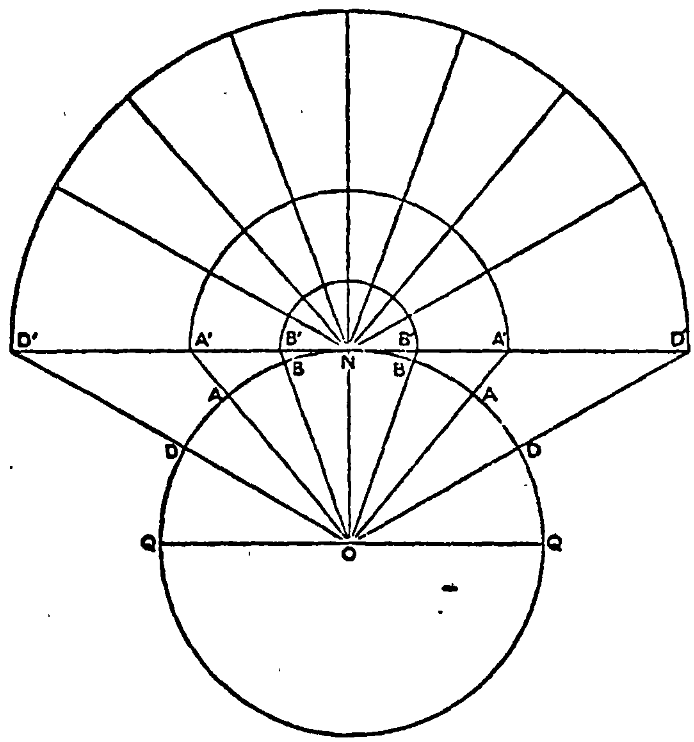
Natural Conical Projection



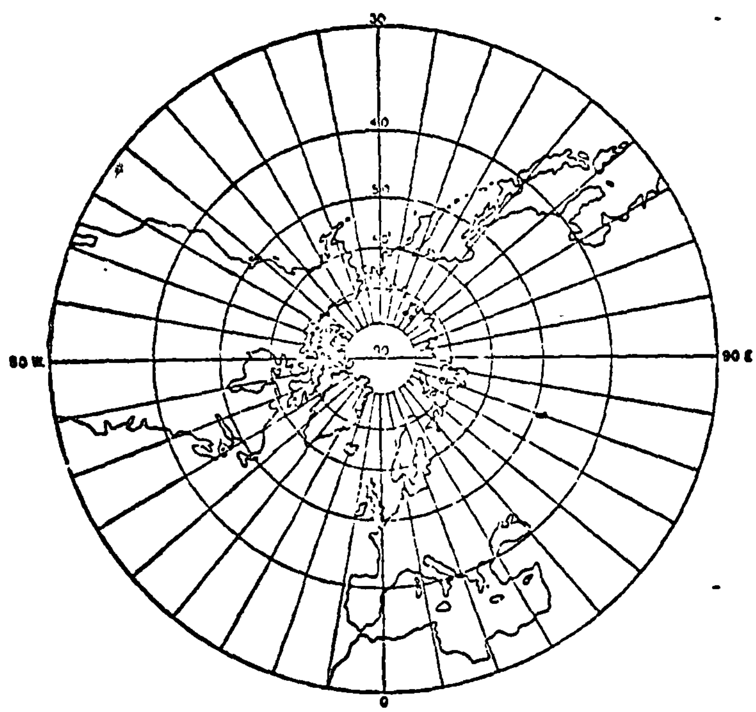
A True Secant Conic Projection



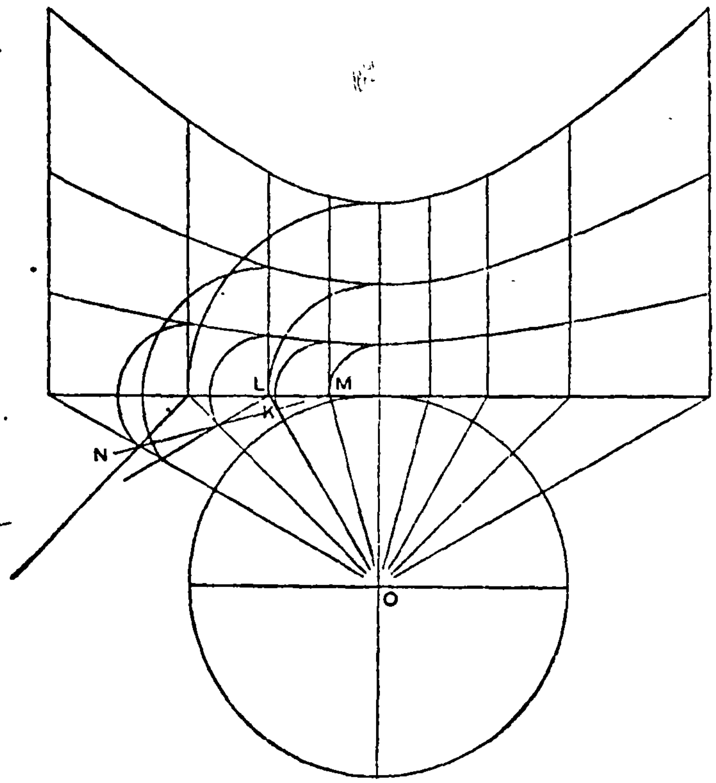
Cylindrical Projection



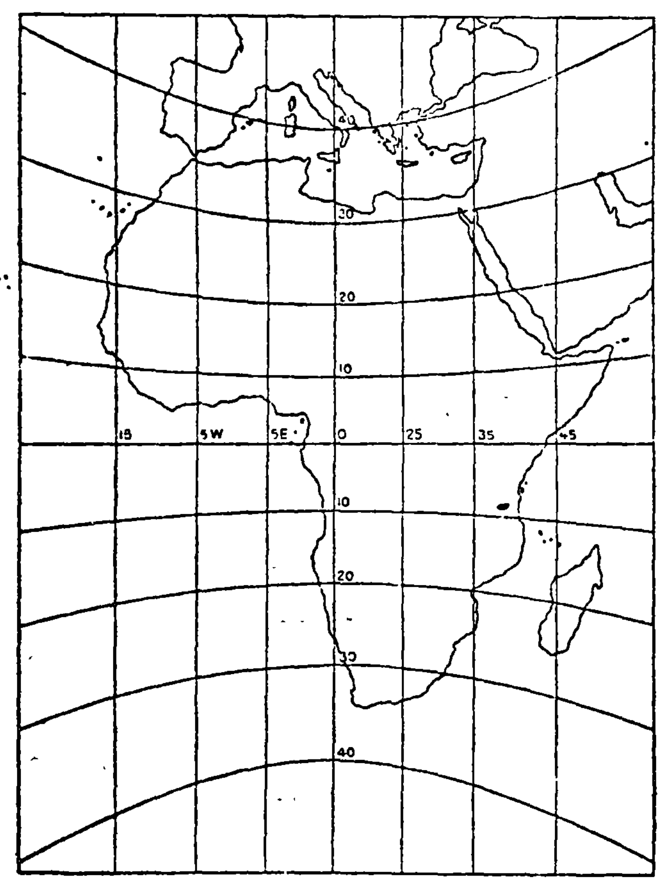
Graphical Construction of Polar Gnomonic



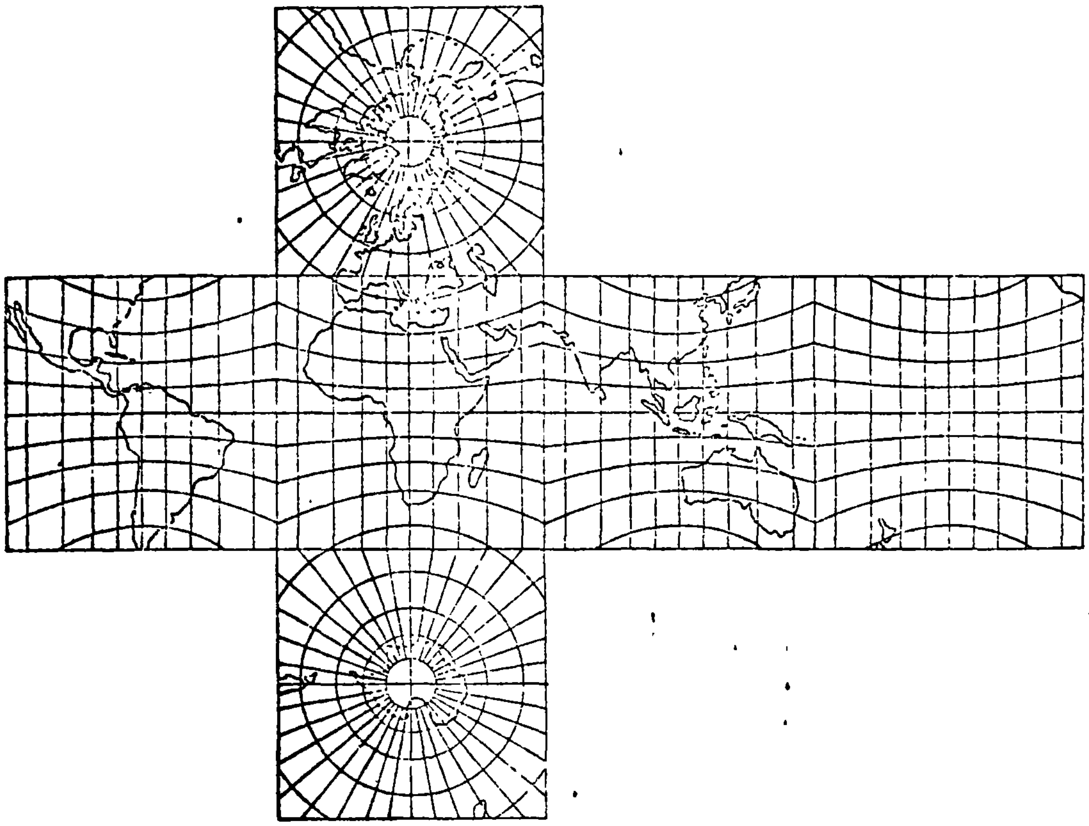
Gnomonic Map of North Polar Areas

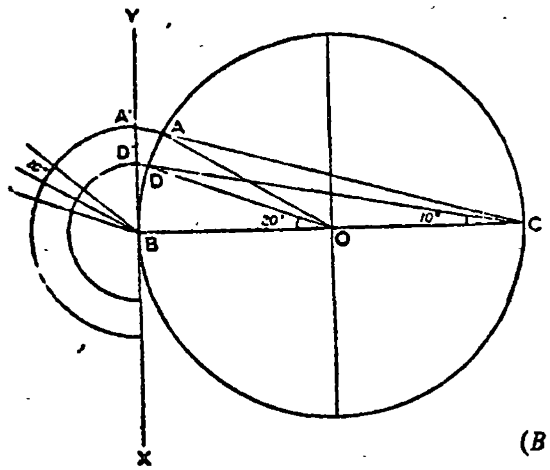


Graphical Construction of Equatorial Gnomonic

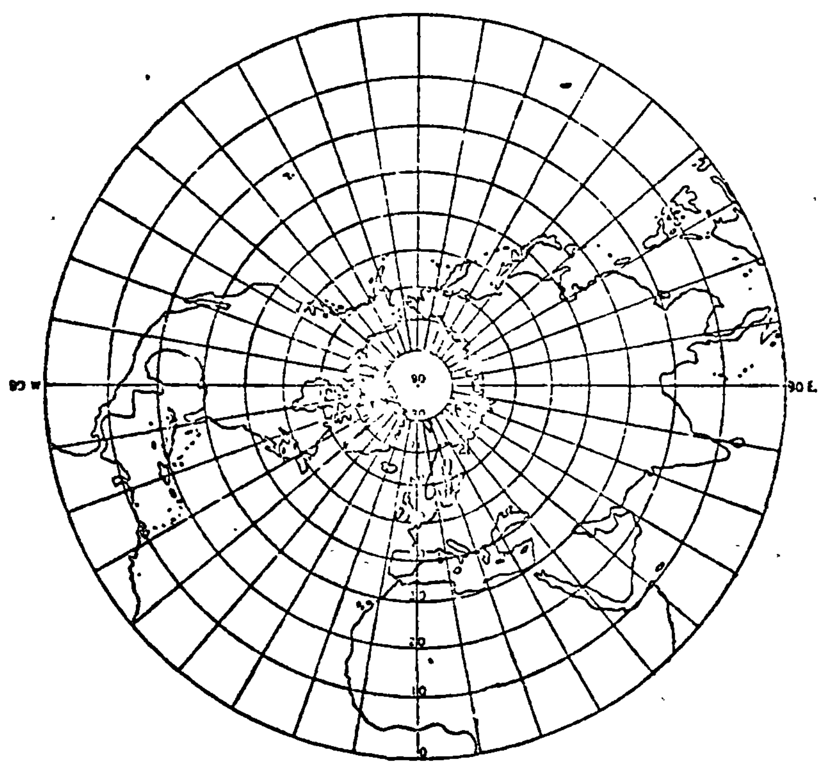


Africa on a Gnomonic Projection

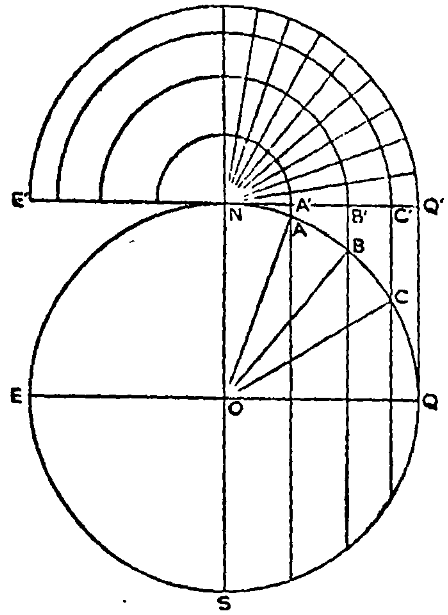




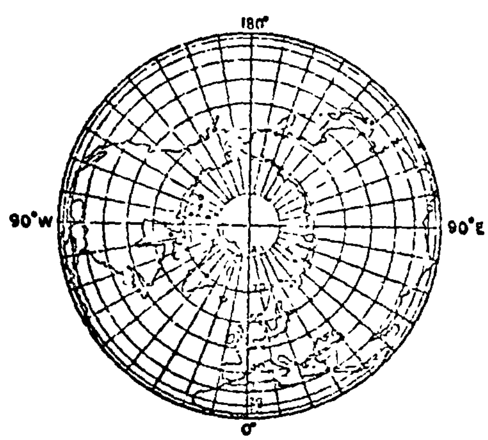
Construction of the Polar Stereographic (B represents the north pole)



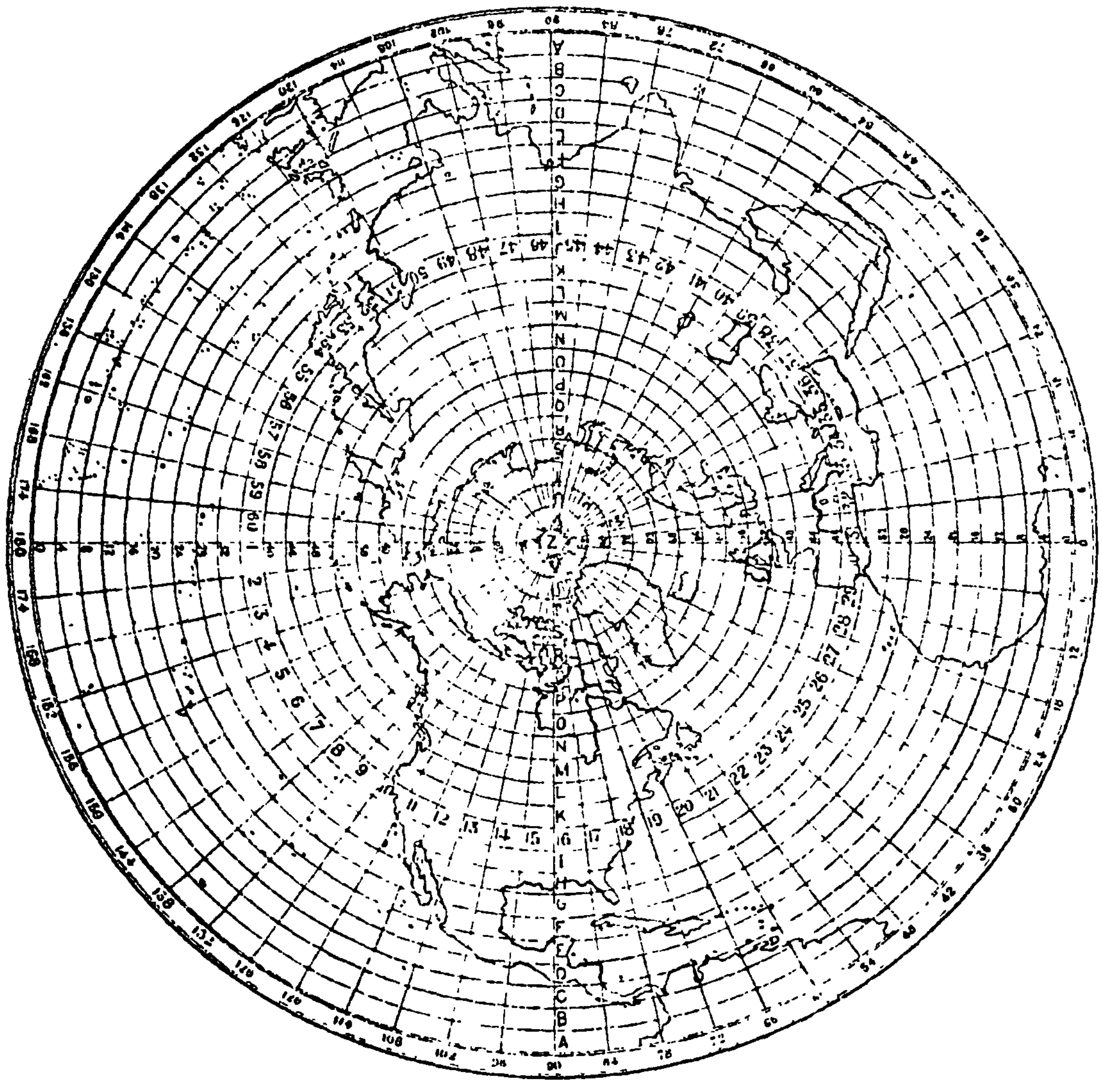
Stereographic Map of the Northern Hemisphere



Graphical Means of Constructing the Polar Orthographic



Orthographic Map of the Northern Hemisphere



DIRECTORIO DE ASISTENTES AL CURSO SOBRE :

MANEJO DE SISTEMAS DE INFORMACION GEOGRAFICA EN PLANEACION

ABEL AGUILAR ARELLANO
S.A.H.O.P.
Insurgentes Sur 1476
México, D.F.

PASCUAL LUNA 21
Tezoyuca, México

ANTONIO ALVARADO DOMINGUEZ
D . E. S . F . I., UNAM
Tel.: 550.52.15 E. 4492

Amado Nervo 31-C
México 4, D.F.
Tel.: 541.25.38

FRANCISCO J. ALVAREZ CASO
Dir. de Monum. Históricos I.N.A.H.
Jardín de Churubusco y Gral. Anaya
México, D.F.
Tel.: 544.42.41

Ricardo Palmerina 11
México 20, D.F.
Tel.: 65.14396

ALBERTO ASPE Y ROSAS
AEROFOTOGRAMETRIA, S.A.
V. M. Alemán 81
México 18, D.F.
Tel.: 277.35.99

8 DE SEPTIEMBRE 43
México 10, D.F.
Tel.: 589.86.70

CARLOS BELMONT GUZMAN
Cía. de Luz y Fuerza del Centro, S.A.
Av.M. Ocampo 171 Oficina 411
México, D.F.
Tel.: 592.02.35

Calle 1 Manz. 5 #18
México 21, D.F.
Tel.: 544.45.68

RAMIRO BERMUDEZ MALLOC
DETENAL
Secc. de Prog. y Pres.
S. A. Abad 124
México 8, D.F.
Tel.: 588.45.35

Unicornio 42
México 13, D.F.
Tel.: 582.14.58

ENRIQUE BUENDIA PINEDA
S. A.H.O.P.
Insurgentes Sur 1476
México, D.F.

B.Juárez 3
S.P. Tecalco, Edo. de Méx.

JORGE CAIRE LOMELI
PEMEX
Marina Nal. 329
México 17, D.F.
Tel.: 545.74.60 E. 2830

Cerrada de Bartolache 9
México 12, D.F.
Tel.: 524.08.93

DANIEL G. CASTILLA P.
ICATEC, S.A.
González de Cosio 24
México 12, D.F.
Tel.: 536.85.60

GUILLERMO CORANTE MONROY
Cía. Mex. de Exploraciones, S.A.
Río Balsas 101 - 8°
México 5, D.F.
Tel.: 533.62.46

Albatroces 46
Izcalli Jardines, Ecatepec, Edo. de

ROBERTO CORANTE MONROY
Cía. Mex. de Exploraciones, S.A.
Río Balsas 101-2°
México 5, D.F.
Tel.: 533.62.46

Retorno 204 # 34
México 13, D.F.

JONATHAN DAVIS
Lab. Sección Planeación
DESFI, UNAM
Tel.:

P. del Pedregal 262
México 20, D.F.
Tel.: 568.19.00

EFREN DE GORTARI RABIELA
D.G. Caminos Rurales
S.A.H.O.P.
Dr. Barragán 779-8°
México 13, D.F.
Tel.: 579.05.35

Viena 271-4
México 21, D.F.
Tel.: 534.13.53

NESTOR DUCH GARY
Dir. Gral. de Estudios del Territorio Nal.
S. A.Abad 124-3°
México 8, D.F.
Tel.: 588.45.35

Magisterio "Nacional 112"
México 22, D.F.
Tel.: 573.46.53

JIMENEZ CRUZ CASIANO
Instituto de Geofísica
Circ. Exterior
Cdad. Universitaria
Tel.: 548.99.38

Manz. 78 Lote 533
México 22, D.F.
Tel.: 548.99.38

YOLANDA LAZO TISCAREÑO
S.A.H.O.P .
Dir. Gral. de Plant. Terr. Asent. H.
México , D.F.

Zempoala 295-204
México 13, D.F.
Tel.: 563.80.76

VICTOR E. LOPEZ GONZALEZ
S.A.H.O.P.
Dr. Barragán 779
México 12, D.F.
Tel.: 590.62.80

Siempreviva 126
México 21, D.F.
Tel.: 549.54.81

ALVARO LOPEZ MARTINEZ
S.A.R.H.
Reforma 107-8°
México 1, D.F.
Tel.: 566.06.88 E.112

Callejón Gral. Anaya 118
México 21, D.F.
Tel.: 544.55.43

ROBERTO MARI RODRIGUEZ
Fundación Javier Barrios Sierra
Tacuba 5 Entrepiso
México 1, D.F.

Coruña 27-9
México 13, D.F.
Tel.: 530.68.99

JAVIER MARTINEZ NORIEGA
Fac. de Filosofía y Letras
Colegio de Geografía
UNAM

Pallares y Portillo 156 B-301
Coyoacán
México 21, D.F.
Tel.: 544.93.35

JORGE V. MIRANDA LOZADA
INFONAVIT
Bca. del Mto. 280
México 20, D.F.
Tel.: 524.97.36

Donizetti 91
México 14, D.F.
Tel.: 537.75.94

JUAN MANUEL NUCHE CABRERA
U. A. M.
Av. Sn. Pablo S/N
Azcapotzalco
México 16, D.F.
Tel.: 561.37.33 E. 281

Mar Mediterráneo 120-33
México 17, D.F.
Tel.: 527.12.73

ISMAEL ORTIZ ALONSO
Cía. Mex. de Exploraciones, S.A.
Río Balsas 101-8°
México 5, D.F.
Tel.: 533.62.46

JOAQUIN PEON ESCALANTE
Teleinformática de México, S.A.
Arenal 40
Villa Alvaro Obregón, San Angel
México 20, D.F.
Tel.: 550.80.33

Puebla 218 Bis Depto. 301
México 7, D.F.
Tel.: 528.74.58

RAFAEL E. RANGEL CHAVEZ
Lab. de la Sección de Planeación
DESFI, UNAM

Antonio Caso 107-3
México 4, D.F.
Tel.: 535.64.92

FRANCISCO EDUARDO RIVERA PORTO
Fundación Javier Barros SIERRA
Tacuba 5 Entrepiso
México 1, D.F.
Tel.: 512.35.68

Xochicalco 703-4
México 13, D.F.
Tel.: 559.12.42

CARLOS RUIZ DE VELASCO G.
Inst. de Invs. en Matemáticas
Aplicadas y Sistemas
UNAM
Tel.: 548.33.60

Fenix 43
México 14, D.F.
Tel.: 577.50.01

VICTORIA CLAUDIA SAEZ ROLDAN
Dir. Gral. de Bibliotecas
UNAM
Tel.: 548.97.80

Yacatas 120
México 12, D.F.
Tel.: 532.68.55

LUIS SALAZAR ZAZUETA
Subdirección de Promoción y Programas
S.A.R.H.
Reforma 35-11°
México 1, D.F.
Tel.: 591.03.83

Niño Pérdido 545-9
México 12, D.F.

CARMEN SAMANO PINEDA
Facultad de Filosofía y Letras
Colegio de Geografía
UNAM

Av. Universidad 1900 Edif. 44-204
México 20, D.F.
Tel.: 548.92.44

JOSE PAULINO SANCHEZ PEREZ
Cía. Mex. de Exploraciones, S.A.
Río Balsas 101-4°
México 5, D.F.
Tel.: 533.62.46

Inst. de Higiene 13
México 17, D.F.

HECTOR SERRANO BARQUIN
Dirección General de Equipamiento Urbano
y Vivienda
S.A.H.O.P
Liverpool 3-11°
México, D.F.
Tel.: 535.95.26

Fco. Murguia 203 Pte.
Toluca, México
Tel.: 5.05.95

EDWIM SIMPSON ESPINOSA
Ing. Miguel Montes de Oca y Asociados, S.A.
Av. Patriotismo440-301
México, D.F.
Tel.: 516.15.53

Av. Contreras 387
México 20, D.F.
Tel.: 595.44.46

ROBERTO SOTO ALVAREZ
S. A. R.H.
Reforma 35
México, D.F.
Tel.: 591.03.83

Jengibres 35
Villa de las Flores, Coacalco
Edo. de México

LUZ MARIA ORALIA TAMAYO P.
Inst. de Geografía
UNAM
Tel.: 550.52.15 E. 4294

Juventino Rosas 4
México 2, D.F.
Tel.: 583.09.84

MATILDE TORRES GARCIA
S. A. H.O.P
Insurgentes Sur 1443-4°
México, D.F.
Tel.: 563.52.21

Pensilvania, 265-1°
México, D.F.
Tel.: 543.22.39

JAIME R. TRASFI S.
ICATEC, S.A.
González de Cosio 24
México 12, D.F.
Tel.: 5.36.85.60

Magdalena 214-201
México 12, DF.
Tel.: 536.52.10

ESTRELLA VILLARREAL DE TOVAR
Fac. de Filosofía y Letras
Colegio de Geografía
UNAM

Luz y Fuerza 11-C 1
México 20, D.F.
Tel.: 593.12.41

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.

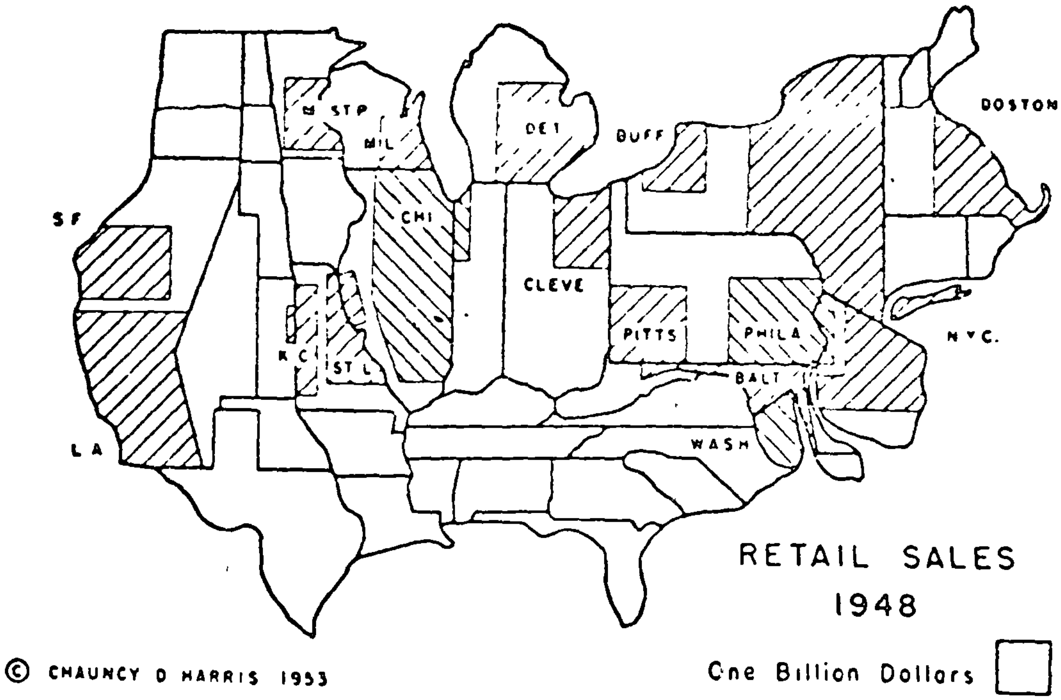


Figure 1: Cartogram Illustrating Retail Sales in the U.S. for 1948

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 to a new origin)
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

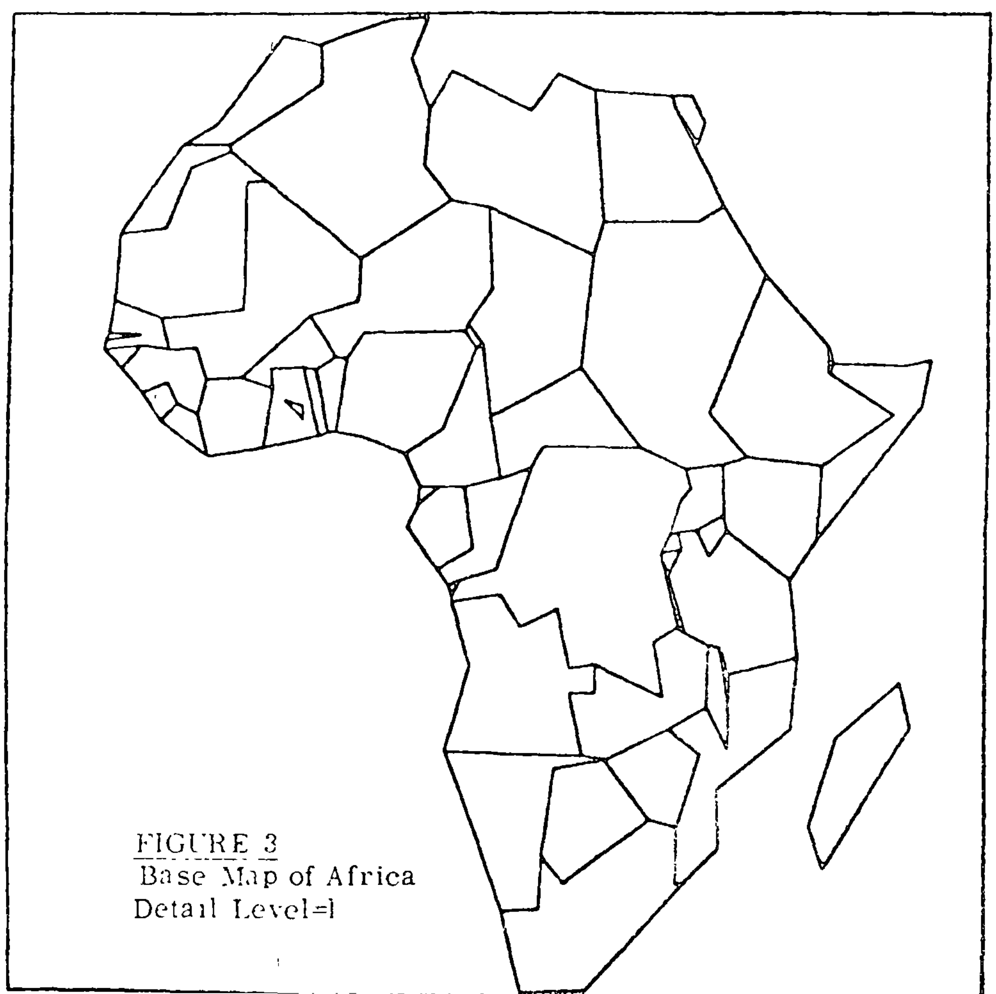


FIGURE 3
Base Map of Africa
Detail Level=1

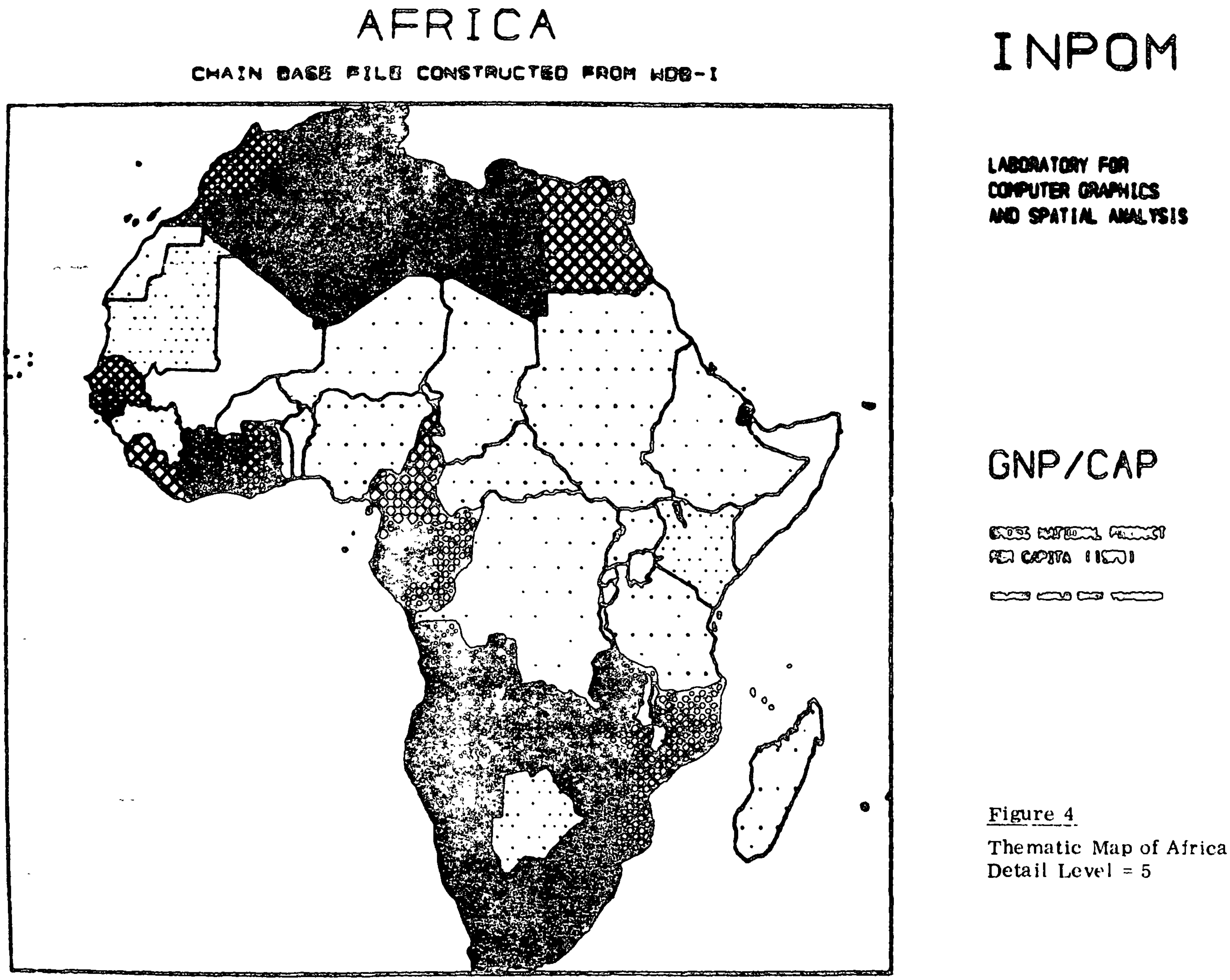


Figure 4

Thematic Map of Africa
Detail Level = 5

WEST AFRICA

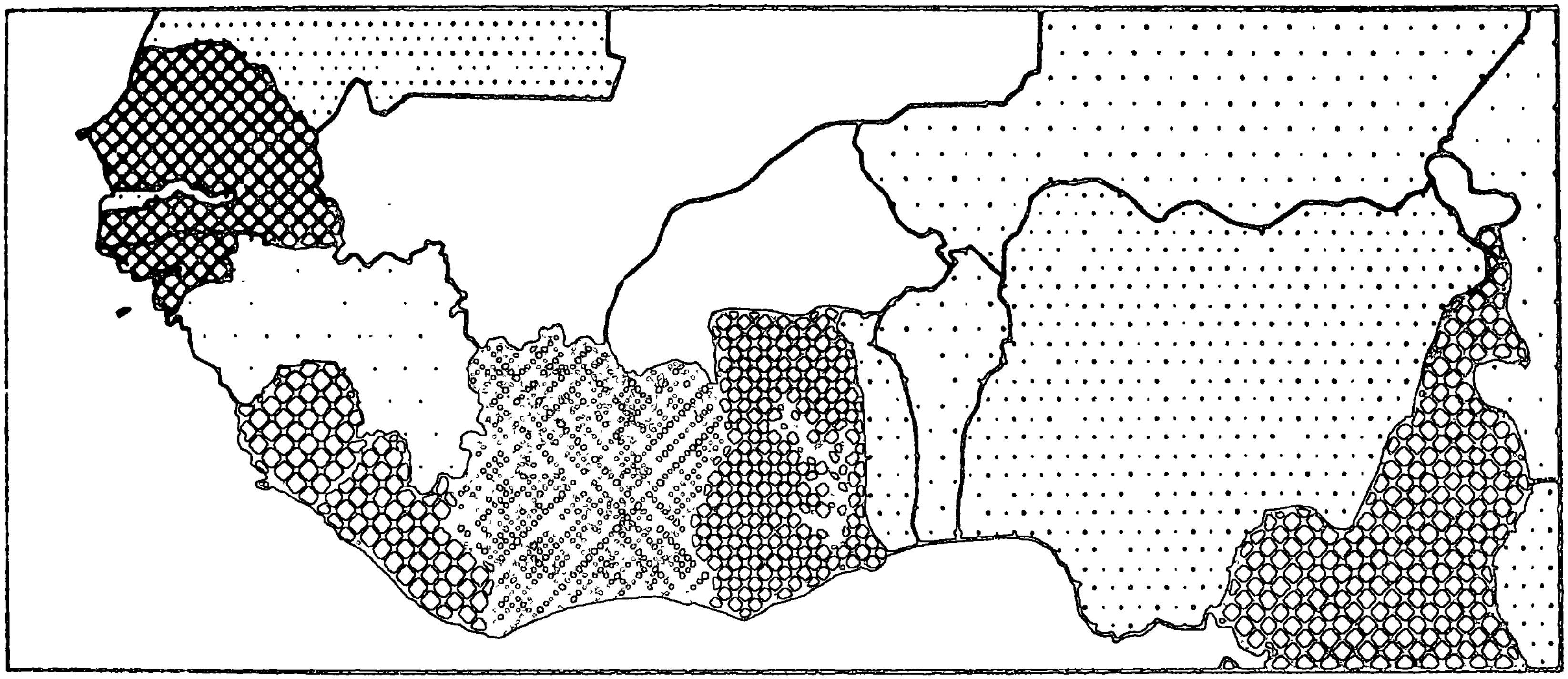


Figure 5

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.

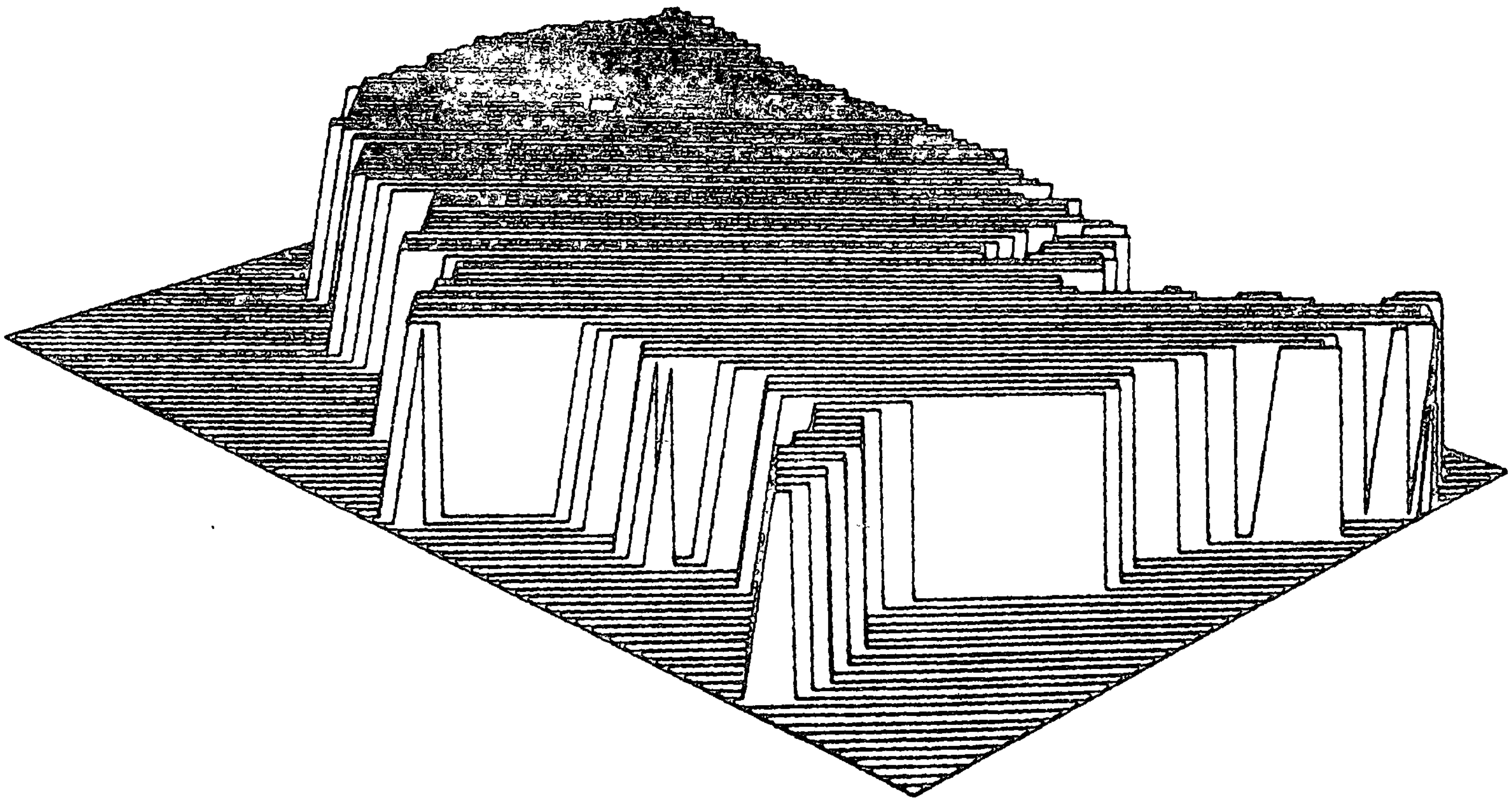


Figure 6

Truncated Height Values

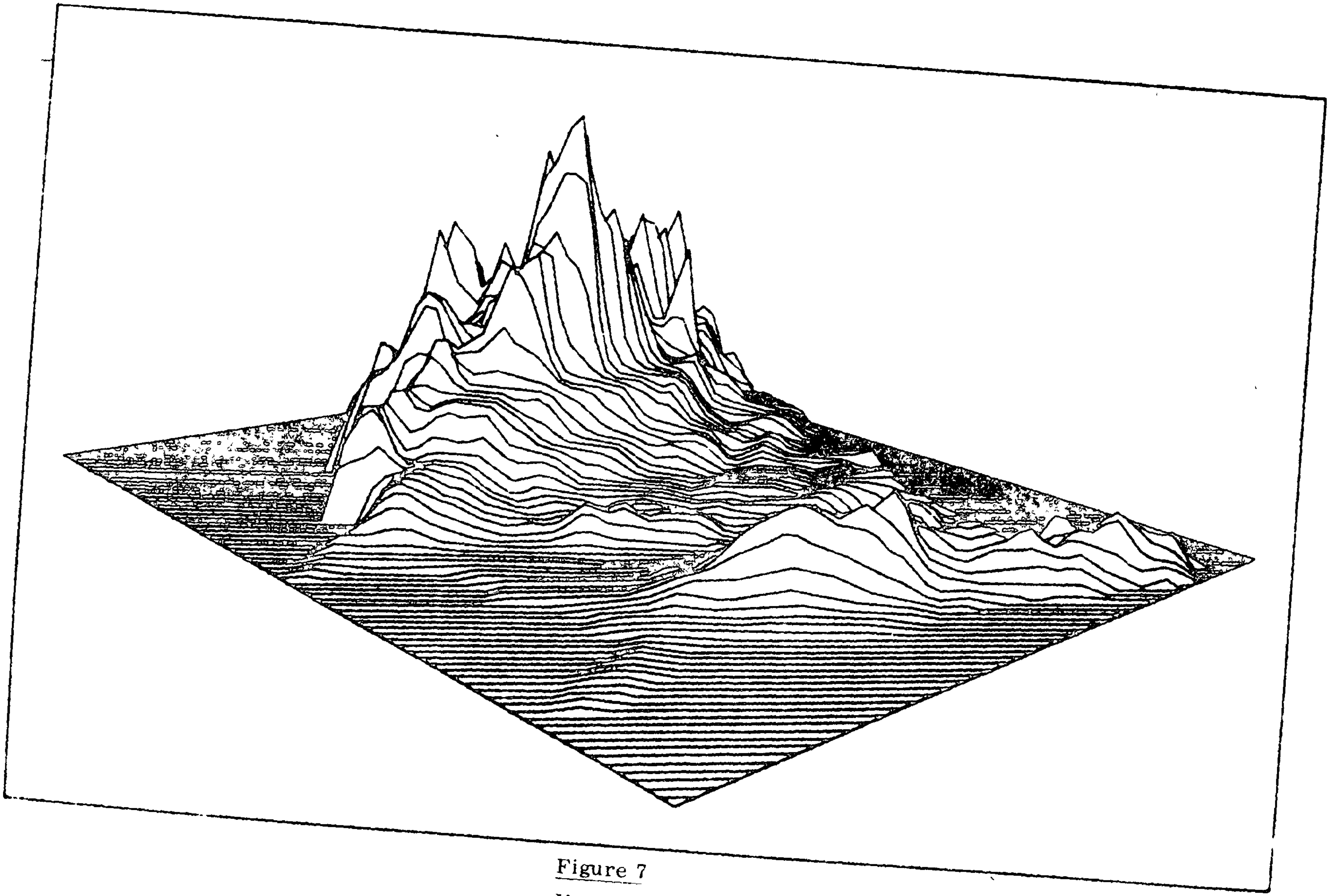


Figure 7

View from the South East

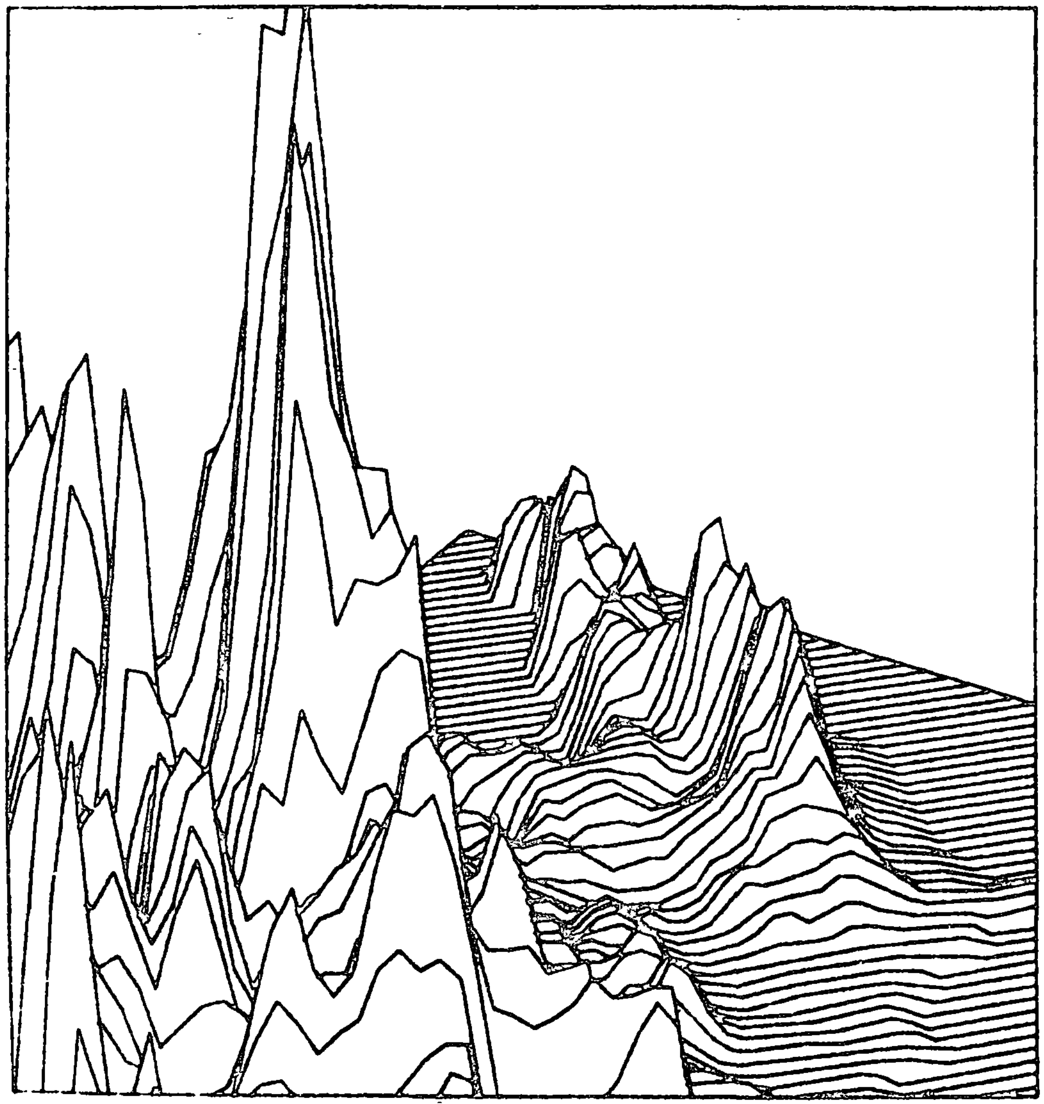


Figure 8

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.

'Urban and Regional Information Systems for Social Programs', Papers from the 5th Annual Conf. of the URISA, 1967, 207-218

Donald F. Cooke, staff
William H. Maxfield, staff
New Haven Census Use Study

THE DEVELOPMENT OF A GEOGRAPHIC BASE FILE AND ITS USES FOR MAPPING

ABSTRACT: The Address Coding Guide (ACG) of the Census Bureau was first used as the geographic base file. But, weaknesses in the ACG system lead to an experimental mapping data base, Dual Independent Map Encoding (DIME), which employs redundant encoding of map features which allows machine detection of errors.

Details of coding and diagrams make both the theory, file preparation, file editing and actual use in mapping very explicit.

INTRODUCTION

One of the objectives of the New Haven Census Use Study is to evaluate methods of computer mapping of small area data. The Study began experimenting with computer-produced maps in February of 1967. Several methods of linking geographic coordinates to data identified by Census block or by address were studied.

The Census Bureau Address Coding Guide (ACG) was first used for the geographic base file, with the ACG's "block-faces" as reference units for data aggregation and display. However, several weaknesses in the ACG-based system led to the proposal of an experimental mapping data base, (called "DIME" for Dual Independendant Map Encoding) employing redundant encoding of map features which allows machine detection of errors. The mapping file thus created has been successfully edited, has proven useful and accurate for mapping, and shows promise for use in many areas of city planning.

The following sections outline the reasons for the choice of redundant encoding, the system's theory, the techniques for developing and editing the file, discussions of the file's use, its cost, and recommendations for its use in other cities.

BACKGROUND

The Census Bureau's Address Coding Guide - part of the mechanism of the mail-out/mail-back Census proposed for 1970 appeared to be the most logical starting point for preparing a geographic base file for computer mapping. The address coding guide is a listing of block faces, which are, as the name suggests, sides or "faces" of blocks. A block face record in the ACG contains the following information:

State—16 009 150 06510 Andrews St. 21 103 2-48

County

Place

Zip Code

Street Name

Census Tract

Census Block

Address Range

In exploring the possibilities of a base file using the coding guide, block face terminals were digitized - that is, coordinates of the endpoints of each block face were measured and recorded, using a semi-automatic "Coordinate Locator" built by the Census Bureau. With the addition of the block face terminal coordinates, the format of the geographic base file became:

State—16 009 150 06510 Andrews St. 21 103 2-48 (x,y) (x',y')

County

Place

Zip Code

Street Name

Census Tract

Census Block

Address Range

Coordinate pair for one block face terminal

Coordinate pair for other block face terminal

However, several difficulties appeared in the use of this file:

1) Digitizing was inefficient, because each cross-street intersection was read eight times.

2) Non-street block boundaries such as shorelines, railroads, and rivers were not included (since there are no addresses on them).

3) Areas of blocks could not be calculated.

4) Winding roads and circular streets could not be described.

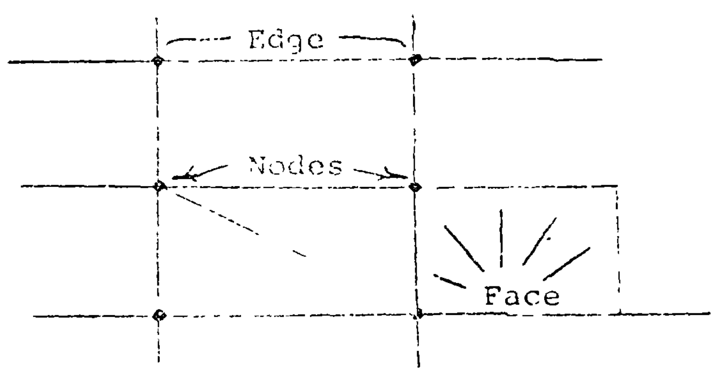
5) Clerical errors were present, and could not be detected by machine edits.

Editing the file proved very difficult and time consuming. As these problems compounded, it became apparent that the Address Coding Guide, though ideal for a mail-out/mail-back Census, could not easily be adapted to computerized mapping.

Therefore, an alternative Mapping Data Base was proposed by James Corbett of the Bureau's Operations Research Division and George Farnsworth, a Census Statistical Intern. The method was tested in New Haven and has yielded a geographic base file which is useful for mapping, and, in addition, may have far wider applicability in traffic analysis, routing of city services, creation of special urban districts, and in the placement of future city facilities. The system is based on graph theory - a branch of mathematics with direct applicability to mapping.

THEORY

The approach suggested by Corbett and Farnsworth is derived from the fact the Metropolitan Maps of the Census

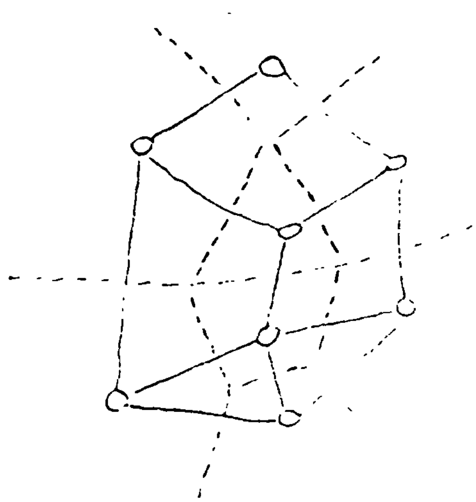


Node = Intersection
Edge = Street Segment
Face = Block

Bureau can be considered mathematically to be linear graphs.

A linear graph is a set of nodes, edges, and faces on a plane, which form a network - such as a street pattern.

Such a network can be described by any of three "incidence matrixes", which define the relationships between the edges and nodes, or nodes and faces, or faces and edges. Whereas only one of the matrixes is sufficient to describe the network, a redundant encoding of the incidence matrixes allows error detection routines to be run on computers until the street network is described perfectly. The network described by one matrix is compared to the supposedly identical network described by another matrix. Differences in the network descriptions can be detected by machine and corrected by hand. Not only the primal network, but also its dual can be edited in this fashion.



Primal Network (solid lines) and its duel (dotted lines)

PREPARATION OF THE FILE

The materials used to create a geographic mapping base file may vary. In New Haven, the materials used were a Metropolitan Map set (which has street names and block numbers labeled) and a city directory to supply address range information for each block. All elements not specifically named on the maps were labeled in a straight forward manner. Nodes were given numerical identifiers and names such as "shoreline" or "Connecticut Turnpike Ramp" or "Mill River Centerline" were assigned to their corresponding line segments.

Node numbering was done sequentially and systematically to facilitate digitizing. Nodes which were numbered included: street intersections, ends of dead-end streets, and places where railroads, shorelines or rivers constituted Census block boundaries and crossed streets or each other. Places where lines on the map curved appreciably were assigned node identifiers also. The criterion for "appreciable" bending of a street depends on the needs of the user, that is, whether he needs very accurate area calculations or wants to reproduce the street network extremely accurately.

Once all nodes were numbered, encoding of the incidence relationships were done in the following fixed length format, each record in the file describing a line segment - or section of a line between two nodes.

Street—Andrews St. 73 129 1 102 1 101 06510 1-25 1 0

Node (1)

Node (2)

Tract (L)

Block (L)

Tract (R)

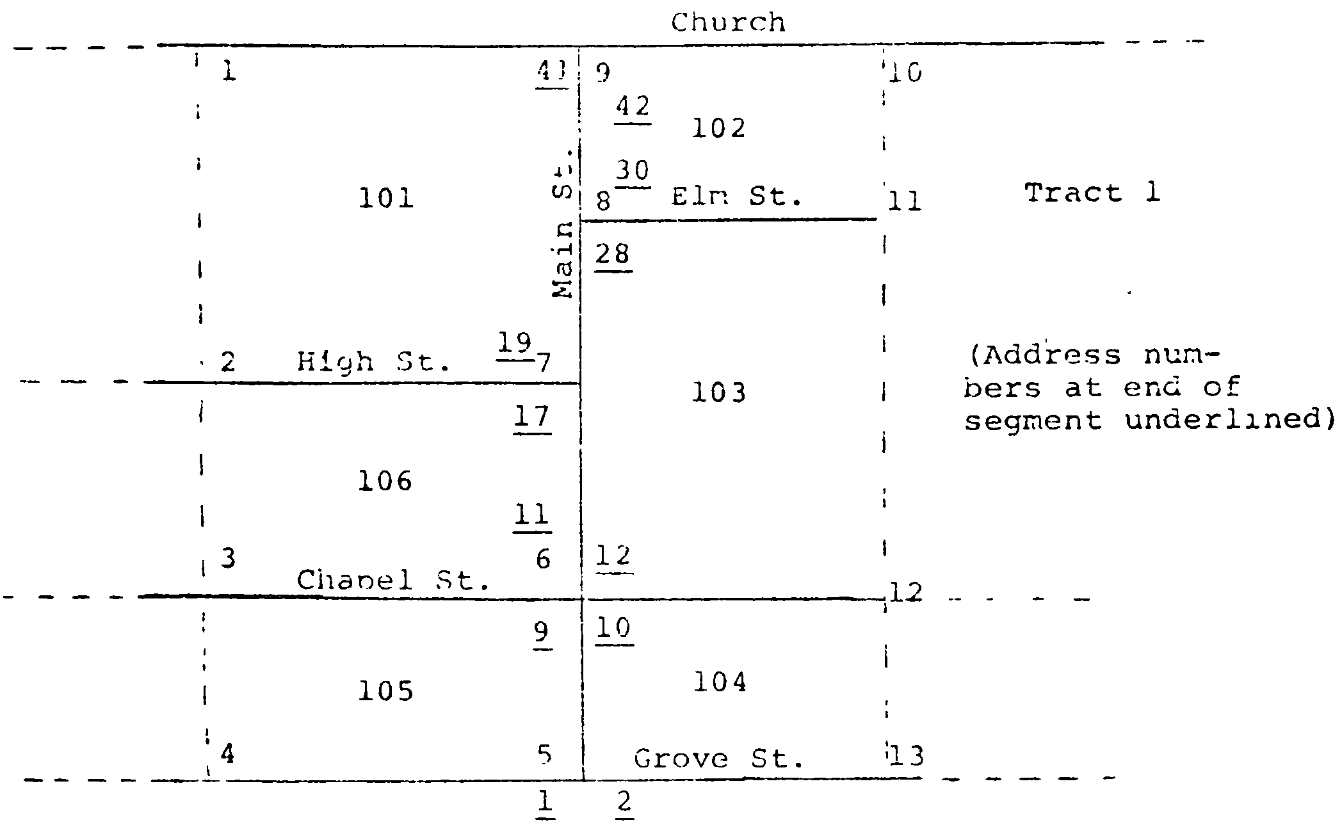
Block (R)

Zip Code

Address Range

Address Code (1 means odd left)

Segment Code (0 means street segment)



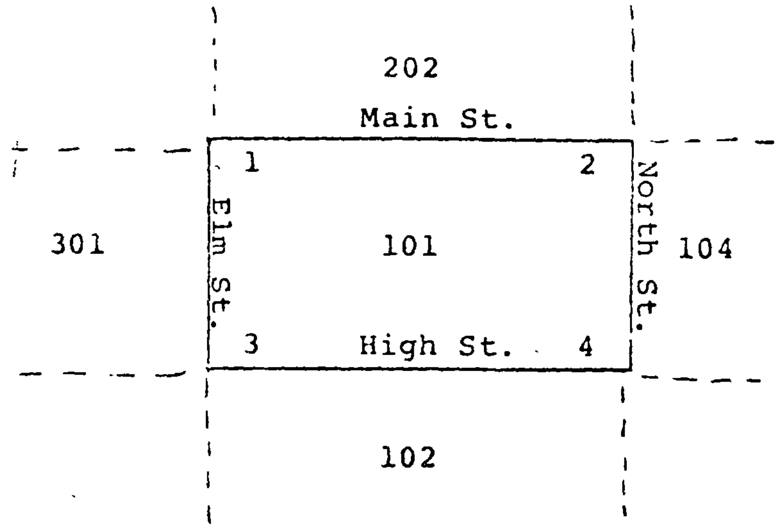
Census Address Coding Guide records for Main Street

Street	Tract	Block	Low Add	High Add
Main St.	1	102	30	42
Main St.	1	103	12	28
Main St.	1	104	2	10
Main St.	1	105	1	9
Main St.	1	106	11	17
Main St.	1	101	19	41

DIME Street Segment Records for Main Street

Street	Node	Node	Tract	Block	Tract	Block	Low Add	High Add	
Main St.	5	6	1	105	1	104	1	10	*
Main St.	6	7	1	106	1	103	11	18	*
Main St.	7	8	1	101	1	103	19	29	*
Main St.	8	9	1	101	1	102	30	42	*

* = Odd left



Primal edit sorts for boundary segments of block (101)

Street	Node[1]	Node[2]	Block[L]	Block[R]
Main St.	1	2	202	101
Elm St.	1	3	101	301
High St.	3	4	101	102
North St.	4	2	101	104

The edit attempts to traverse the boundary by placing the segments in order by matching on Node numbers.

Street	Node[1]	Node[2]	Block[L]	Block[R]
Main St.	1	2	202	101
North St.	2	4	104	101
High St.	4	3	102	101
Elm St.	3	1	301	101

Block L and Block R have been switched if the nodes had to be reordered to effect a traverse. (This occurs for all streets except Main)

In this example, the traverse is successful, and the block numbers have been encoded correctly. Note the relationship of the node numbers and block numbers.

Node 1 and Node 2 are the node numbers defining the line segment. Their order is such that Node 1 is at the low address end of the segment. Tract L and Block L are the Census

The implication of this is that the clerks who perform the encoding process need not even be able to tell their right hand from their left. However, in New Haven we had excellent clerical help, and very few errors in the Block L - Block R encoding were encountered.

A final machine edit sorts out all segments records associated with each intersection. The module two sum of the block numbers associated with any node should be zero. If this is not the case, an error message is written.

Another useful edit is a visual check of the street network as reconstructed from the segment file using digitized coordinates. An almost equivalent edit is to plot mechanically the node numbers on an overlay to the source map. However, the former technique exploits the pattern recognition abilities of the human eye, and erroneous digitizing is easily discovered.

Edits which we hope to run as soon as possible will check the address range information for internal consistency. Any gaps or overlaps of address ranges will be detected and corrected by hand. When the address range checks have been completed, the data base file will be complete for New Haven and West Haven, Connecticut. We hope sometime to expand the directory to include the rest of the S.M.S.A. The final mapping file will include calculated block centers and areas.

USE OF THE FILE IN NEW HAVEN MAPPING

The file without coordinates is usable for address coding, and for many functions which require knowledge of the contiguity relationships of blocks. Mr. Farnsworth mentioned some of these last night: studying the "neighborhood" concept, moving or mosaic averaging, aggregation of statistics of any area by specifying the boundary of the area, districting, and avoiding disclosure of confidential information. We have coordinate information in New Haven, which allows the further ability of graphic display, area and density calculation, calculation of distances along streets, and aggregation of statistics for districts specified by coordinates, for instance, a count of all people living within one mile of City Hall.

Our primary interest lies in graphic display, however, and experiments indicate that we have a very flexible tool for this purpose. In testing the applicability of the mapping file, we have produced maps of [illegible] from block statistics using SYMAP, a printer mapping program de-track and block numbers to the left of the segment as seen by an observer standing at Node 1 and looking towards Node 2. The address range is derived from the city directory and the Address Code indicates odd numbered left side of the street or odd numbered right side of the street. Note that this is an inclusive address range for both sides of the street while each record in the Census Address Coding Guide gives address range information for one side of the street. Segment Code is a number assigned to indicate whether the record is a street, railroad, shoreline, or other type of segment. It was also used to indicate problems encountered by the encoding clerks. Coordinates of the nodes defining the segment may be incorporated into the record, or may be carried separately in a list which can be indexed by node number.

EDITING THE FILE

Street segment records were keypunched, and the first of several computer edits was run. This edit sorts out for each block all segment records associated with the block, then further sorts out all boundary segments. Note that a boundary segment can be identified as having different block numbers in the Block L and Block R fields.

An interior street such as a dead end will have the same block number on each side. The edit program attempts to "traverse" the boundary of the block by linking the segments together in order by matching node numbers. Success or failure messages are written out, and are used to guide clerks in correcting whatever errors or omissions are present. The edit program is run and corrections and additions are made until no traverse failures are reported.

The second edit is a test of the accuracy of the encoding of the Block L - Block R information. Again a traverse of the boundary is made, this time with the Block L and Block R fields being switched if the nodes had to be reordered to effect a traverse. The program then checks to see if the block number of the block being checked appears entirely in the Block R or entirely in the Block L fields. If this condition fails, there is an error in encoding the block pair. It may be of interest to note that if the coordinates of the nodes are digitized and inserted into the segment records by matching on the node numbers, then the areas of the blocks can be calculated as soon as a block can be traversed. The sign of the calculated area will indicate the direction of the traverse. Knowing the direction of the traverse enables one to know which of the Block L or Block R fields should contain the block number of the block being edited.

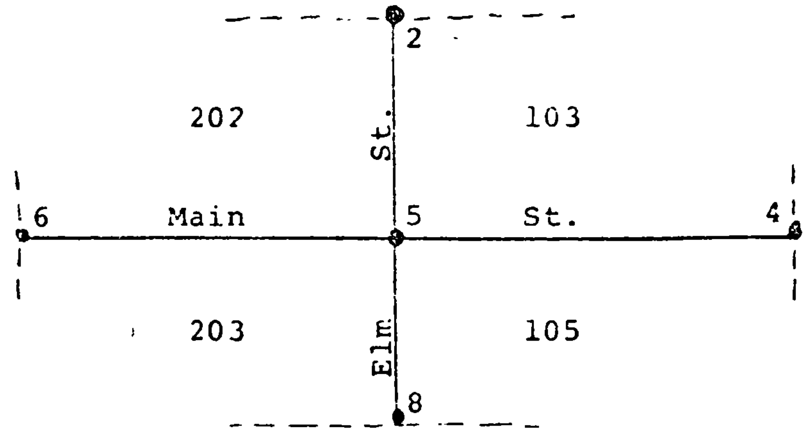
veloped by Howard Fisher of the Harvard Laboratory for Computer Graphics. The SYMAP options we have tested require that the boundary of the map section to be plotted be described by coordinate pairs of vertices in order around the boundary. Retrieving this information from the Geographic Base File was easily done and can be done by machine. Block center coordinates, and the area of the blocks (needed for the density calculation) were calculated directly from the segment file.

We are testing another printer mapping package - MAP 01, for use with the mapping file to map Census data. We are also writing a mapping program to plot data maps on CRT, pen plotter, or Geospace plotter output devices. The figure of merit of any maps produced by the Census Use Study will be the usefulness, if any, to local planners who use Census data. Hopefully some of the techniques we develop in New Haven will be significant enough and cheap enough to be adopted by the Census Bureau for their publication problams sometime in the future.

EXTENSION OF THE SYSTEM TO OTHER CITIES

In recommending adoption of this system in other cities, we would modify the foregoing procedure as follows: we would stress the importance of accurate source materials. Much time was lost trying to read detailed information off maps where the scale (1" = 800') made the features ambiguous. We would suggest using 1" = 400' or even larger scale maps, preferably ones with address numbers on the maps. We would further suggest that two address ranges be carried for each segment record - one for each side of the street. Working in this manner from maps such as the Sanborn series would enhance the address coding accuracy of the file (which we consider to be the weakest part of the New Haven file).

Once the source maps are encoded, machine plotting of the street segments, block numbers, and street lables allows rapid, accurate and inexpensive production of work maps at any desired scale. Aside from the size of the file, there is no technical reason why the system of network encoding cannot be extended to the parcel level, where parcel boundary segments, rather than street segments, are the units of the directory. An experiment to test this is being run on a four tract section of New Haven, primarily to produce a highly accurate address file to test the address coding capability of the New Haven mapping file. Extension of the system to the parcel level is straightforward, and not as time-consuming as might be expected. A file such as this, matched to tax



Dual edit sorts for all segment records associated with intersection (5)

Street	Node[1]	Node[2]	Block^L	Block^R
Main St.	6	5	202	203
Main St.	5	4	103	105
Elm St.	8	5	203	105
Elm St.	5	2	202	103

If the intersection being edited (5) appears in the Node field, multiply Block L by -1. Otherwise, multiply Block R by -1.

Street	Node[1]	Node[2]	Block^L	Block^R
Main St.	6	5	202	-203
Main St.	5	4	-103	105
Elm St.	8	5	203	-105
Elm St.	5	2	-202	103

Next, all the block numbers are added:

```
   202
  -103
   203
  -202
  -203
   105
  -105
 + 103
 -----
     0
```

If all segments are encoded correctly, the sum will be zero.

R. G. Loomis and J. J. Lorenzo
IBM Corporation
New York Scientific Center

EXPERIMENTS IN MAPPING WITH A GEO SPACE PLOTTER

ABSTRACT: The Geo Space Plotter is a computer mapping system using wet photographic process. It is versatile in that images are stored on disk, transferred to tape for reading and restored on disk for updating or overlay with additional image components to produce a photographic image of a special purpose map.

Using the New Haven Census Use Study data, it proved to be capable of plotting block numbers in irregular blocks and base maps depicting streets, tract boundaries and railroads. Experiments in block shading for density studies were less promising.

Future experiments in equipment and programming include use of a "Fast ALPACA" system allowing construction of entire image before plotting computer storage is used, and the development of a faster dry photographic process.

INTRODUCTION

Within the preceding year the IBM New York Scientific Center initiated a study to investigate advanced computer applications in the area of urban affairs. At a very early state of this study it became quite clear that one of the distinguishing characteristics of many urban planning and operational activities is their significant dependence on spatially organized data - a fact which requires no further substantiation for those present at this meeting. As a consequence, it was decided to undertake exploratory studies directed to what were judged to be three major aspects of computer processing of spatial data; namely

1) techniques for storing and manipulating spatial data within a computer,

2) systems which would permit planners and other users to directly interact with such data by means of on-line graphic terminals,

3) methods for providing hard copy graphical represent-

records, could yield extremely accurate, detailed, land use information. In addition, the file could yield the percentage of land that is residential, commercial, street and even the area of the sidewalks.

FILE DEVELOPMENT COSTS

One final note on the New Haven street segment base file - the cost. Even using very generous pay scales, and overestimating computer and digitization usage, we found it difficult to push the cost estimate for the New Haven file over $3,000.00, including writing of the edit programs. This figure works out to about 2 cents per person, or $3.00 per block. We cannot extrapolate the cost to other cities at the present time, but we do feel that this method yields an inexpensive and accurate geographic base file with broad application as a tool for urban analysis.

TOWARD CREATION OF A DIGITAL TERRAIN DATA BASE

Arthur A. Noma
DEFENSE MAPPING AGENCY
TOPOGRAPHIC CENTER
Washington, D. C. 20315

Presented to
The Annual Convention of the
AMERICAN CONGRESS ON SURVEYING AND MAPPING
and the
AMERICAN SOCIETY OF PHOTOGRAMMETRY

St. Louis, Missouri
10–15 March 1974

TOWARD CREATION OF A DIGITAL TERRAIN DATA BASE*

Arthur A. Noma
Defense Mapping Agency Topographic Center
Washington, D.C.

ABSTRACT

A number of digital terrain data collection systems have come into operation at the Defense Mapping Agency Topographic Center during the past decade; the oldest being the *D*igital *T*opographic *D*ata *C*ollection *S*ystem (DTDCS), most current being the *S*emi-*A*utomated *Cart*ographic *S*ystem (SACARTS). This paper describes the digital contents of terrain data collected by these systems and its availability for computer applications. It also discusses the use of the collected digital terrain data for establishing a digital data base for areas of Department of Defense interest. The paper concludes with some recent applications of the digital terrain data.

* Presented to the Annual Convention of the American Congress on Surveying and Mapping–American Society of Photogrammetry (ACSM-ASP), St Louis, Missouri, 10–15 March 1974.

TOWARD CREATION OF A DIGITAL TERRAIN DATA BASE

Arthur A. Noma

INTRODUCTION

The effort toward exploiting digitized terrain data for mapping and charting began in the late 1950's, only a few years after the advent of the first commercial electronic computers. At the Defense Mapping Agency Topographic Center (DMATC), then the Army Map Service (AMS), the effort in utilizing the digital technology for cartographic work began shortly after the installation of the first commercially available computer, the UNIVAC I,* in 1952. Three systems began in the era and their successors continue in operation today at DMATC. These systems are:

- The Integrated Mapping System.
- The Automatic Model Carving System.
- The Symbols and Names Placement System.

These systems in one form or another utilize digital media for collecting, processing, and controlling the graphic product output. The systems diagram, Figure 1, from the publication entitled "Let's Go Over the Hill – Potential Benefits of Profile Scanning the Stereo-Model" by Messrs. Spooner, Dossi, and Misulia in 1957,[1] shows the creation of a byproduct labeled "Topo Tape Library."

During the early and middle sixties, techniques and ideas pioneered by these systems were refined and incorporated into a number of cartographic systems. The activities of the late sixties showed the rapid swing in mapmaking toward exploitation of digital techniques. Today, perhaps more importantly, cartographic technology is on the threshold of making available map information in digital form for more comprehensive planning. The forecast by William C. Aumen in the midsixties of a "A New Map – The Numerical Map"[2] may have arrived. This paper will describe three systems now operational at DMATC for the collection, processing, and storage of digitized terrain and other topographic data. These systems are:

- The Digital Topographic Data Collection System (DTDCS).
- The Universal Automatic Map Compilation Equipment (UNAMACE) System.
- The Semi-Automated Cartographic System (SACARTS).

SACARTS encompasses a much larger scope than DTDCS and the UNAMACE System. In addition to collecting, processing, and storing digital data, SACARTS considers the total cartographic steps needed for the production of topographic maps. Hence, this

* Any mention herein of a commercial product does not constitute endorsement by the United States Government.

paper addresses only that portion of SACARTS currently operational for the creation of digital data. Subsequent papers are forthcoming on full exposition of SACARTS.

DIGITAL TOPOGRAPHIC DATA COLLECTION SYSTEM (DTDCS)

A decade ago, DMATC (then AMS) began collecting digital terrain data on a production basis. As originally conceived and elaborated in the publication entitled "Programming Topographic Maps for Automatic Terrain Model Construction" by Messrs. Noma and Misulia in 1959,[3] the system automated the carving of three-dimensional terrain or relief maps.

During 1961–62, the initial approaches were modified and a new digitizing technique and processing methods were implemented. The equipment constructed, and concept still in use, was the Digital Graphics Recorder (DGR), sometimes called the Floating Arm Graphics Recorder (FAGR), Figure 2. The April 1963 issue of the Communications of the ACM, Vol 6, No. 4, the journal of the Association for Computing Machinery, presented a description of a "prototype" system.[4] The system consisted of a digitizing table which collected the incremental motion of traced contour lines on a magnetic tape. This was followed by a sequence of computer programs to verify the data by producing plots, filling gaps between contours by interpolation, and outputting a tape for directing a tape-controlled milling or carving machine, Figure 3.

Before the fully automatic modeling capability was put into use, the system's ability to produce digital terrain data triggered production programs to provide data for a number of Department of Defense (DoD) agencies.

The requirement to provide elevation data temporarily shelved the completion of a milling device until 1967–68, and the system was successfully put into operation in 1969. The publication "Automated Terrain Modeling" by Messrs. Mays and Noma in 1970[5] described the system, Figure 4. Figure 5 illustrates the surface appearance depictable using the system. One of recent experiments carried out is the digital enlargement of the horizontal and vertical scale. That is, data collected and stored from 1:250,000 scale have been carved at 1:125,000 and even 1:62,500. The shrinkage of scale is also possible.

The increased production requirement by the DoD community for digital elevation data accelerated the acquisition of more rapid and efficient collection equipment, culminating in the Digital Topographic Data Collection System (DTDCS) developed in 1966–67, Figure 6. The DTDCS is designed to handle five DGR tables (maximum) via data flow traffic control by computer. The system collects "on-the-fly" traced data to a core buffer area and thence to assigned tracks on a 1.5 million word capability Disk

Pack. The system will, on the average, hold at least one shift's worth of traced data for each table. Before the data on the Disk Pack are transferred to tape, they are plotted for validation and necessary corrections are made on the disk.

The production rates are for the system to handle 200 to 300 digital terrain sheets per year, or from 4 to 6 sheets per week, or at least one a day. This level of production would require two full DTDCS systems, with 10 DGR's in operation.

The processing software has similarly evolved from original concepts and was modified as production rates increased. The procedure for processing was first published in the paper entitled "Digitizing Graphic Data at the Army Map Service" by Messrs. Mays, Noma, and Aumen. [6] Figure 7 illustrates the current process used.

Present digitizing practice calls for identification of two types of data:

- Contour lines and spot elevations.
- Stream and ridge lines.

On the initial sort/merge step, the digitized data are sorted into profiles and the two types of data are separated. The sorted contour and spot elevation file is used as a control file for the subsequent steps that follow.

The sorted stream and ridge line file is matched to the control file to obtain the elevation values. The data are reordered back into stream and ridge lines and all elevation values for these lines are obtained by ordinary linear interpolation from the intersection points. These data are again sorted back into profiles and merged with the control file.

Next the sheet neatline information is derived. This process uses coordinates provided as input and first generates X's and Y's defining the neatline. Using these points, a match is made with the control file to obtain height information for interpolation in defining all neatline elevations. Here, two types of interpolation have been used; a moving cubic spline fit and an ordinary point-to-point linear interpolation. Experience indicates that the spline produces more rational values if attention is paid to collecting sufficient control points at max-min and inflection points; however, this type of precision normally requires interactive processing which adds to time and effort needed to complete the job. Thus, the current practice has turned more and more toward the use of the simpler but sufficiently accurate linear method.

A repetitive interpolation scheme, called planar interpolation, is used to determine all undefined surface points within the sheet region. The derivation of the method has been described in a number of earlier works and a detailed exposition will not be given here. As shown in Figure 8, the approach is to obtain surface values point by point by passing planes through three known points. The algorithm operates within an ordered array and, hence, at completion of the calculation all points within the sheet are defined.

The process at DMATC is specifically tied to the recording density of the collection system which is 0.01 inch in X and Y for the DGR. Thus, the digital array produced reflects the unit of the digitizer and is independent of source material scale.

As stated earlier, the digitizing and collection of elevation data began over 10 years ago at DMATC. Within this timespan, the completion of digital elevation coverage of the continental United States, Figure 9, is a major accomplishment. This index shows the latest inventory check made of the agency holdings of available digital data. Gaps in the coverage are sheets being verified and inventory checks still taking place. The bulk of the data has been collected and processed from the standard topographic 1:250,000 scale series map and retained on magnetic tape reels by map sheet number. Copies of the data are available in DMATC standard compacted elevations on industry-standard magnetic tapes. The procedure for handling requisition and distribution are described in the current military map supply catalogs.

The current size of library holdings at DMATC is about 1,600 reels. The major coverage is of the United States and plans call for collection in other world areas. The primary source is the 1:250,000 scale topographic map; however, 1:50,000 scale topographic maps will be digitized and processed for areas of specific interest based upon current available coverage.

UNAMACE

Figure 10 illustrates the data flow and processing of the UNAMACE System. The use of digitized data from the UNAMACE System has been described in "Automatic Contouring at the Army Map Service" by Messrs. Viticllo, Biggin, and Middleton in 1968.[7] Additionally, a more recent publication by Biggin, entitled "Computer Generated Contours from Numerical Data," in 1971,[8] describes current work at DMATC. The system produces digitized elevations directly from rectified stereo imagery. As described in the cited papers, the primary effort is directed toward techniques for processing the collected digital data for the automatic plotting of contour lines.

Figure 11 shows the current UNAMACE data flow. First, the data from each model are collected by an update program and stacked onto a master. After all models have been collected for the mapping area, the data are processed through a contour program. The program is provided with orientation parameters and a smoothing criterion. Proper scaling and smoothing is performed on the data followed by a process to digitally "draw" lines for the contour elevations specified. The resulting output of the process are (1) contour plot tape for drawing labeled contour lines and (2) smoothed elevation data in profile to drive an orthophoto device, such as the Gigas-Zeiss, for producing orthophotography.

Two major operational problems in the processing of UNAMACE data to produce usable terrain data are the inability to detect areas of inconsistencies on the digital data and the inability to panel or mosaic digital data to construct consistent terrain data for an entire map region. Areas of inconsistencies occur because of equipment limitations where correlation is lost. Locating these areas and correcting the errors are necessary for effective use of the digital data. Techniques for digital mosaicking are now being developed and plans call for their incorporation as a subsidiary step prior to the contour program.

In addition to the two points discussed, the question comes up as to what smoothing does to the raw data. Is smoothing done to produce cartographically acceptable contour lines or to better represent the true ground configuration? Are the two the same? The point of immediate concern related to the UNAMACE data is whether the master or the smoothed be retained. Perhaps certain cartographic digital definition may be needed to make this determination.

There are approximately 400 UNAMACE master tapes at DMATC. However, their current utility is limited since the data are strictly discrete individual digital models. The effectiveness of using the UNAMACE terrain data will depend upon the successful implementation of a system to collect and combine data to cover entire map sheet areas.

SACARTS

The initial phases of the *S*emi-*A*utomated *Cart*ographic *S*ystem (SACARTS) have just been put into production use at DMATC. The system is the combination of both hardware and computer software for aiding the cartographer in the production of repro quality materials, Figure 12. Inputs to the system are both hypsography, from processed UNAMACE type elevation data, and planimetric data, from vector or line digitizers such as the CALMA and BENDIX units. The processing of the data will be done on a large central computer such as the UNIVAC 1108. Presently, the table digitizer is used for servicing, correcting, and adding features; however, plans call for interfacing the *D*igital *I*nput-*O*utput *D*isplay *E*quipment *S*ystem (DIODES) for rapid interactive editing. This subsystem consists of cathode ray tubes and digitizing table for displaying and correcting digitized data. For production of repro quality material, a high-precision digitally controlled drafting machine such as the Concord plotter is used.

For more rapid output of high-content data, a raster-type plotter will be used. The raster unit can also be used for rapid scanning of manuscript as input to the processing steps.

SACARTS consists of a number of different computer-controlled subsystems; however, central to the concept is data processing on a large-scale computer. The initial version of processing software has been developed and tested, called Graphic Improvement Software Transformation System I (GISTS I). Details of GISTS data bases have been published in the paper "A System for Automated Cartographic Analysis and Map Production" by Messrs. Burdette, Dario, and Spencer, presented at the October 1973 Regional Cartographic Conference for Asia and the Far East in Tokyo, Japan.[9] Figure 13 shows the general schematic data flow and logical processing of GISTS. Additional details on the system are given in two other papers to be presented at this 1974 Spring ACSM-ASP Convention, "Improved Cartographic Copy from Digitized Map Compilation" by Henry R. Cook[10] and "Color Separation Symbolization in Semiautomated Map Production" by William Burdette.[11] The former paper provides additional details on systems design and data structure while the latter discusses the software techniques used to produce control tapes for outputting repro quality, color-separated, symbolized plots.

This system, as presently constituted, is designed for efficient production of graphics; however, elements for the construction of a planimetric digital data file are there. Data fields are expandable to handle a wide variety of coded line and point features. In anticipation of a raster scanner/plotter, software for determining the inside and outside of regions, which can be highly relevant to an effective digital data base, can be incorporated into the GISTS file structure.

The present digital holding from this system is small and for actual use available only for graphic manipulation; however, as the system is put into production, auxiliary processes will be designed to manage the planimetric data files.

Current production plans call for the system to be phased in, beginning with rates up to 50 map sheets per year to a capacity of 500 map sheets per year in a 3- to 4-year span. With acquisition of additional and new equipment as well as improved software, the system can be pushed upwards toward a 1000-maps-per-year rate. These production rates will obviously create an explosion of digital tape data. Hence, in a few years this system should be a significant generator of digital data, especially planimetric information.

FILE ATTRIBUTES

The three major digital data collection systems can be interfaced for the creation of a common digital terrain, or more generally called topographic, data base. The base should consist of possibly two distinct, however, interrelated files; namely, elevation and planimetric data.

The elevation data should be in the form of elevation arrays organized into uniform logical regions, for instance, 1° by 1° blocks. The planimetric data can be broken into two categories – line and point information. Line data such as drainage, roads, railroads, and other communications networks, should be stored using some form of chain encoding. Point information such as buildings, mine locations, wells, etc., on the other hand, should be stored in position lists by information category for a specified region.

A point for consideration is how area data should be retained. The classification of area data covers such things as lakes, swamps, forests, rice paddies, urban outlines, etc. – in general, cartographic "open windows." These types of data can be retained in two forms: first, as a direction-coded line to indicate the inside and outside of the region outlined by the line, or, second, like elevation data, carried in array form. The choice of the structure may depend on response requirements needed to meet user needs. Array form should be much more effective in responding to inquiry of whether a point is inside or outside of a region, since, if data are in line form, either a search for line must be made or the line data expanded to array to respond.

In order to establish a responsive system, efficient file maintenance and distribution will be necessary. However, before such a system is put into operational use, the present file, as well as files to be generated, must be structured for universal coverage. The current digitized data are being retained in some arbitrary local form; in most cases, the instrument system. These files should be organized into a universal, contiguous system; for example, into latitude bands.

Two major problems must be solved before a universal digital terrain data base or file can be effectively instituted:

- Selection of a universal coordinate system for organizing the digital data.
- Developing a matching process to combine digital data such that information is contiguous.

Aumen has suggested using geographic coordinates as the universal reference. The use of the latitude-longitude system ties effectively to the idea of maintaining data in latitude bands. A fairly uniform system can be developed if the longitude interval increases as latitude increases. However, the selection of a coordinate system is far from complete and further studies should be made in the light of user needs and requirements.

As regards the matching of data to make adjoining data contiguous, different techniques have been studied by a number of groups. One method which has been tried is smoothing by spreading the difference over an arbitrary band adjoining a junction line. Other techniques used have involved mathematical surface fitting. The issue, however, is that a technique should be selected to operate effectively on a universal data base.

Another major point of consideration on data base creation is data compaction. Some studies have been made in this area. However, the current and near-future requirements indicate that use of current computer technology should more than keep up with present files and those being created. Such items as efficient data compaction by chain encoding and use of high-density storage media can comfortably keep pace with the collection rates. As an example, over 6,000,000 elevation points can be stored on a reel of tape using the present DTDCS procedures. By installation of available high-density tape units, the capacity can be doubled. In fact, there are available systems for a four-fold increase if needed. The compaction problem should certainly be studied to be ready to handle the long-range needs, although it is not paramount to the present establishment of a universal digital terrain data base.

SUMMARY

Once the system is established, a positive investment of resources will be needed to maintain the file and handle its distribution. As new data are collected, information should be processed to be incorporated into the data base. This will require a substantial allocation of computer resources, and manpower for reviewing and editing to retain data fidelity.

The distribution of the data will also require the development of software and commitment of computer resources. The copying of digital data for distribution will be one of the simpler functions; but, for effective response, the distribution system should have the capability to selectively retrieve data from only the specified regions as well as extract the feature requested. It should also have the ability for some amount of format changing of the data to fit the user's need. Additionally, future needs may dictate the design of logical synthesizing software to extract for the user implied topographic information for more effective decision making.

The foundation for the creation of a universal digital terrain or topographic data base is available today.

LITERATURE CITED

1. SPOONER, C.S., Jr.; DOSSI, S.W.; and MISULIA, M.G. "Let's Go Over the Hill–Benefits of Profile Scanning the Stereo-Model." *Photogrammetric Engineering.* December 1957.

2. AUMEN, W.C. "A New Map–The Numerical Map." Presented to the 32d Annual Meeting of the American Society of Photogrammetry, Washington, D.C. March 1966.

3. NOMA, A.A. and MISULIA, M.G. "Programming Topographic Maps for Automatic Terrain Model Construction." *Surveying and Mapping.* September 1959.

4. "Terrain Model System." *Communications of ACM.* Association for Computing Machinery. April 1963: Vol 6, No. 4.

5. MAYS, R.R. and NOMA, A.A. "Automated Terrain Modeling." Presented to the 30th Annual Meeting of the American Congress on Surveying and Mapping, Washington, D.C. March 1970.

6. MAYS, R.R.; NOMA, A.A.; and AUMEN, W.C. "Digitizing Graphic Data at the Army Map Service." Presented to the American Society of Photogrammetry–American Congress on Surveying and Mapping Convention, Washington, D.C. March 1965.

7. VITIELLO, D.; BIGGIN, M.J.; and MIDDLETON, G. "Automatic Contouring at the Army Map Service." *Photogrammetric Engineering.* October 1968.

8. BIGGIN, M.J. "Computer Generated Contours from Numerical Data." Presented to the American Congress on Surveying and Mapping American Society of Photogrammetry Convention, Washington, D.C. March 1971.

9. BURDETTE, W.; DARIO, E., and SPENCER, N. "A System for Automated Cartographic Analysis and Map Production." Presented to the Seventh United Nationals Regional Cartographic Conference for Asia and the Far East, Tokyo, Japan. October 1973.

10. COOK, H R "Improved Cartographic Copy from Digitized Map Compilation." Presented to the Annual Convention of the American Congress on Surveying and Mapping–American Society of Photogrammetry, St. Louis, Missouri. March 1974.

11. BURDETTE, W.H. "Color Separation Symbolization in Semiautomated Map Production." Presented to the Annual Convention of the American Congress on Surveying and Mapping–American Society of Photogrammetry, St. Louis, Missouri. March 1974.

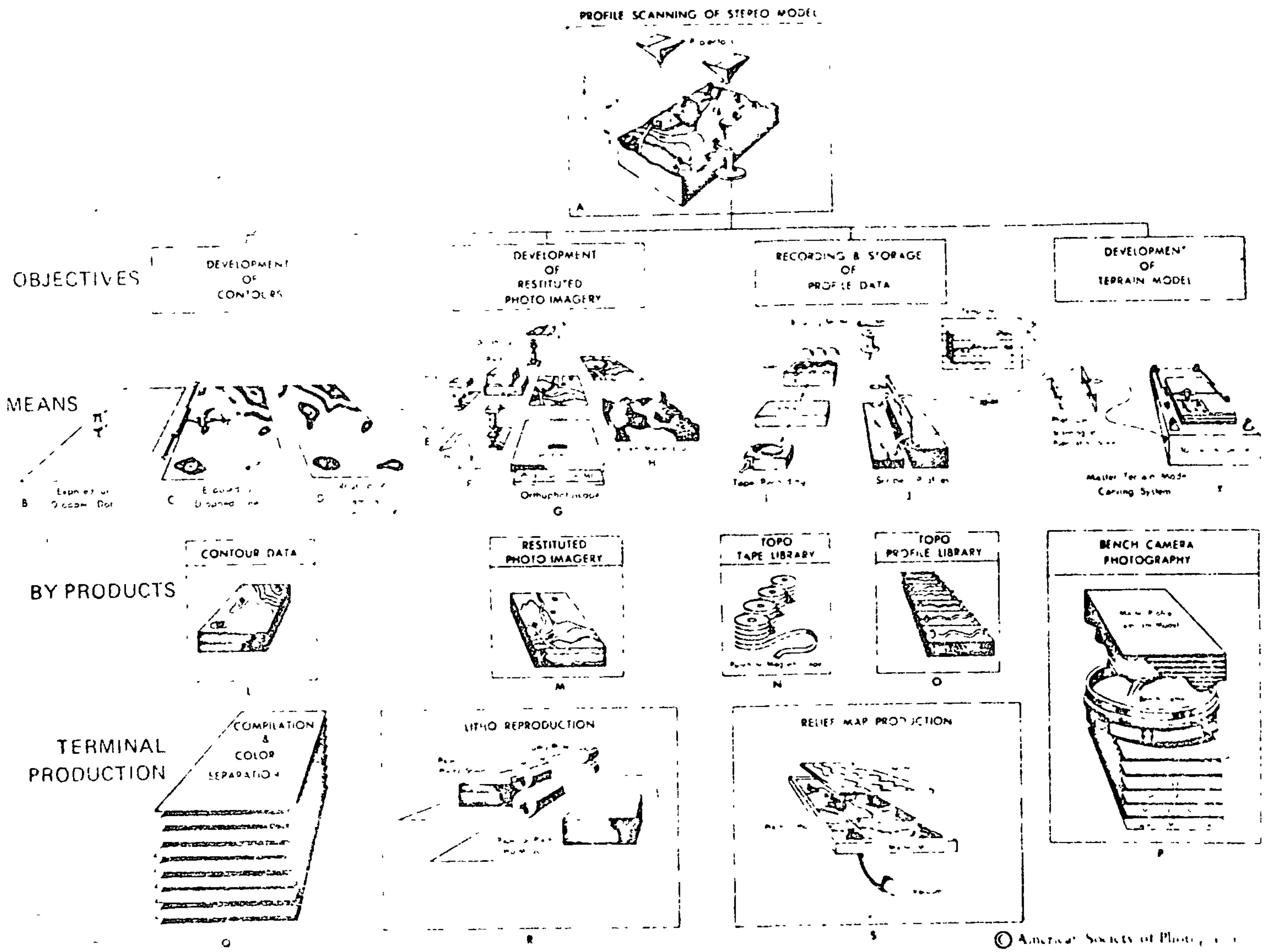


Figure 1. Integrated mapping system diagram.

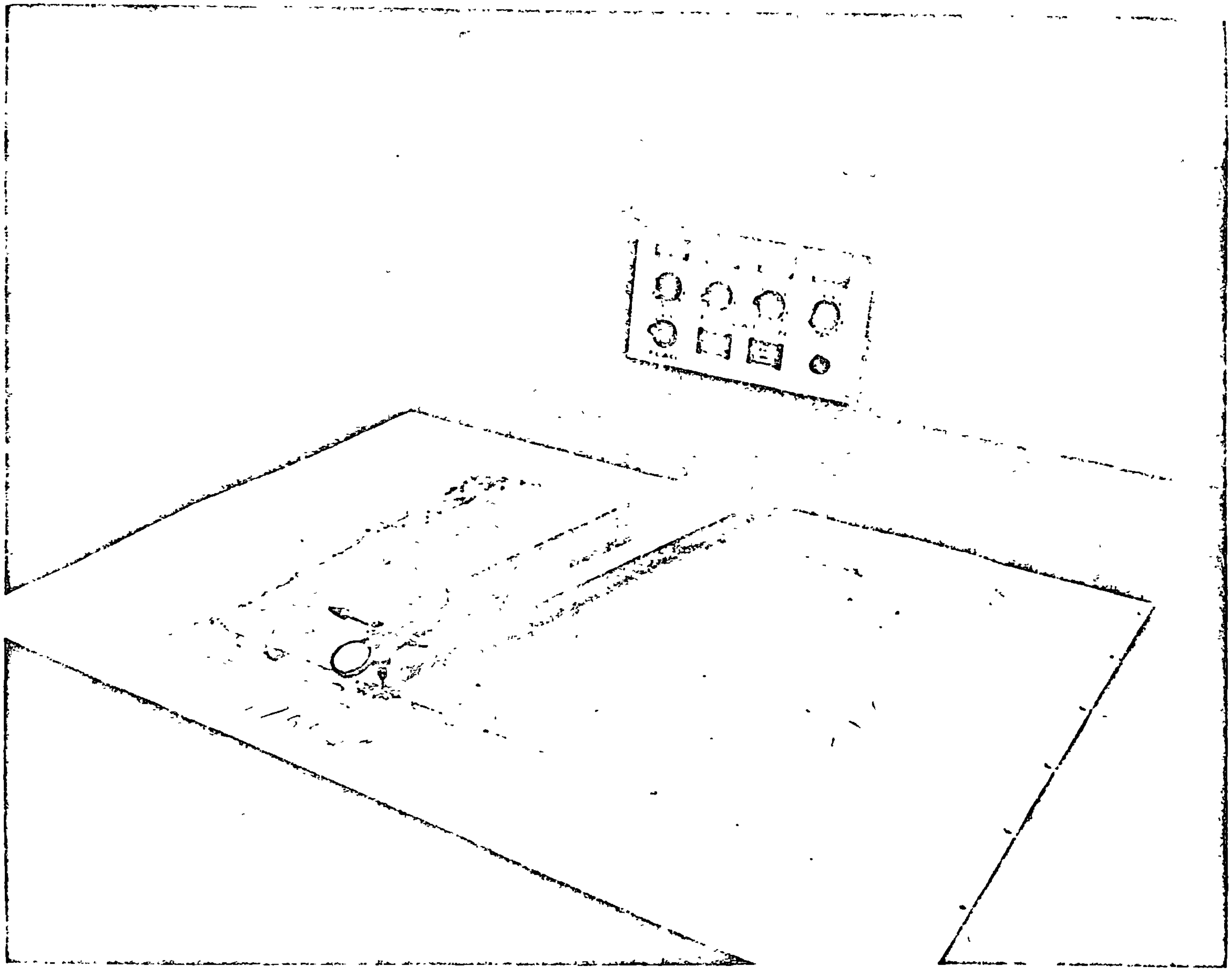


Figure 2. Floating arm graphics recorder.

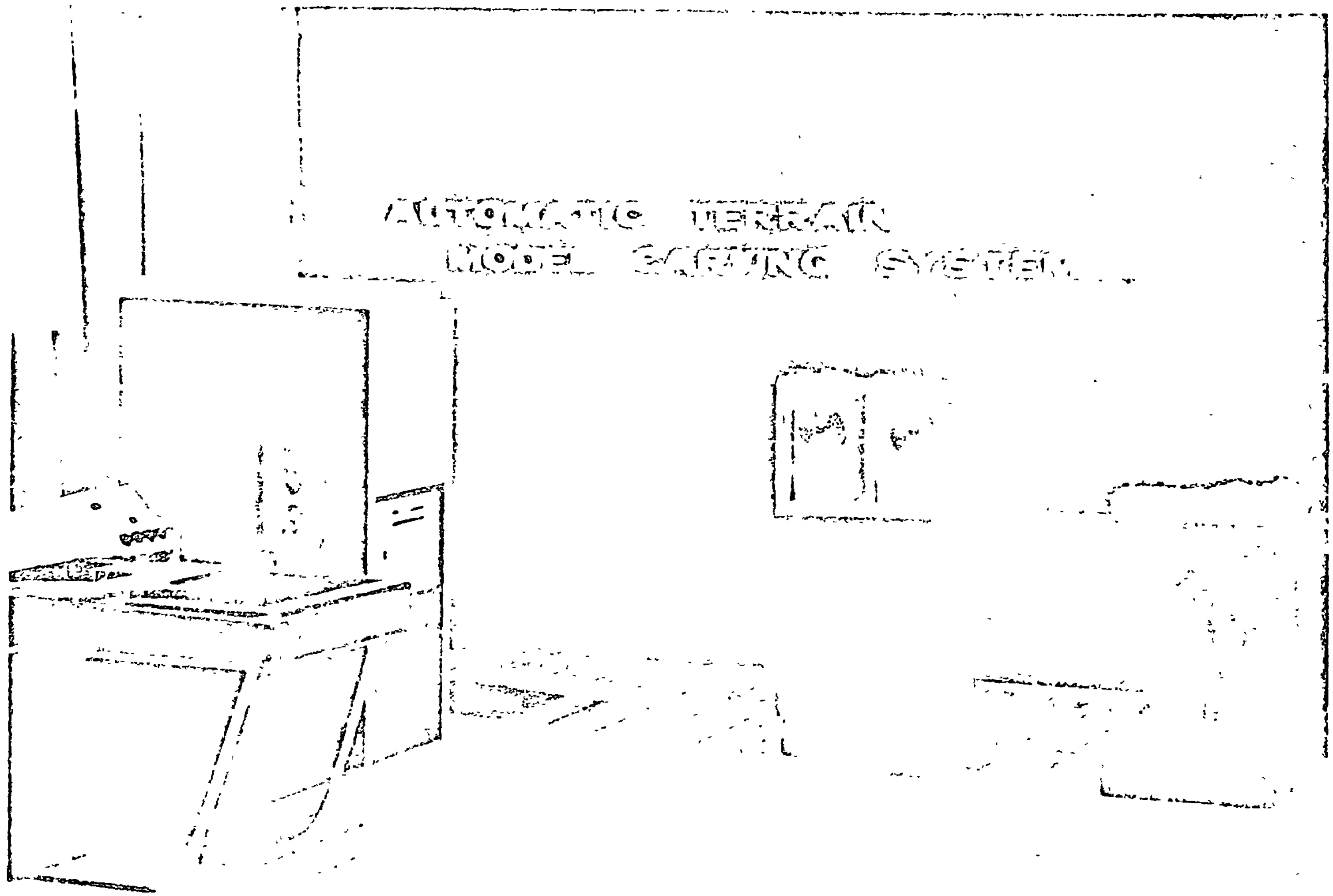


Figure 3. Automatic terrain model carving system–model of layout.

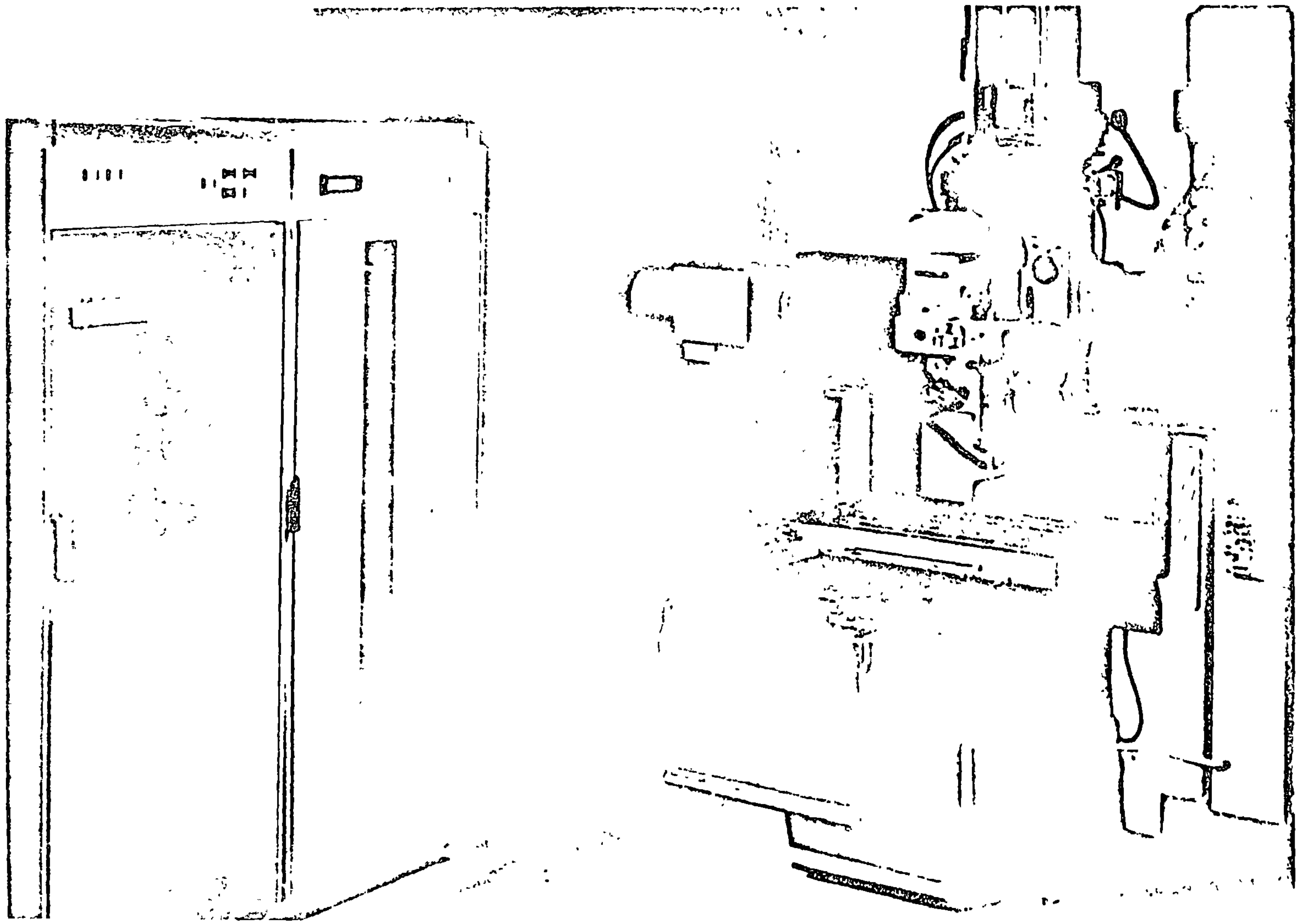


Figure 4. Automated terrain model carver.

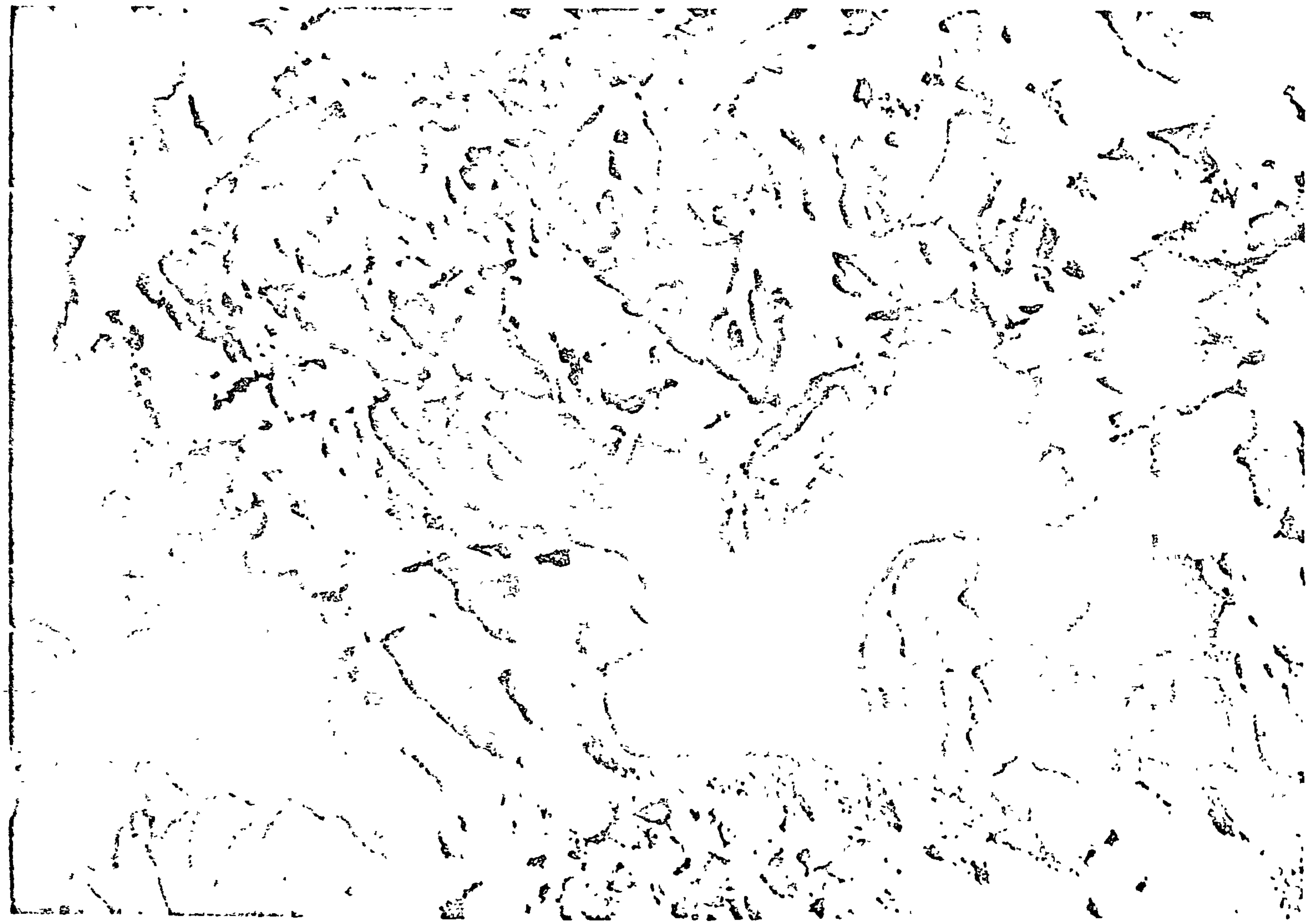


Figure 5. Terrain model carved by system.

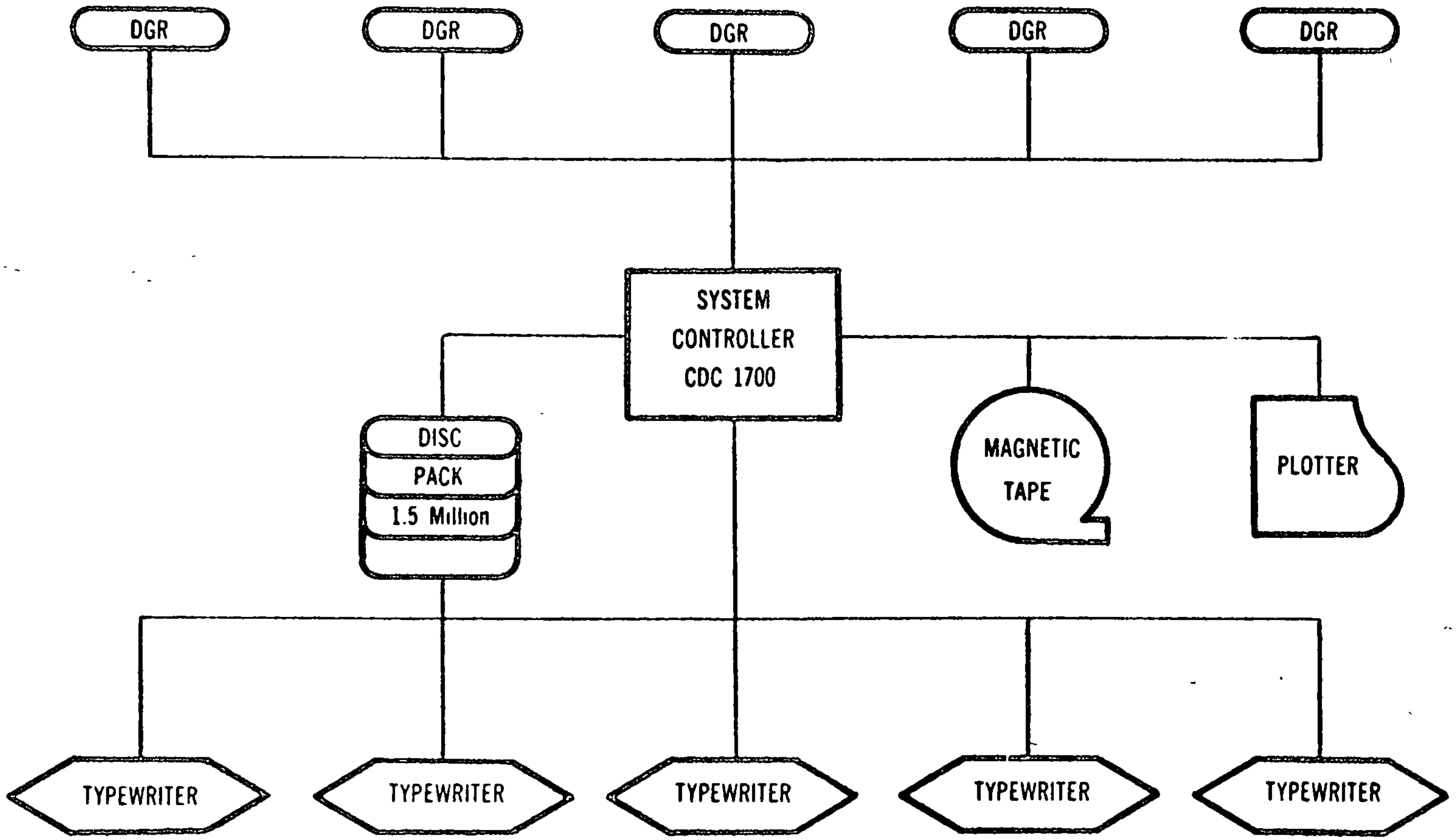


Figure 6. Digital topo data collection system

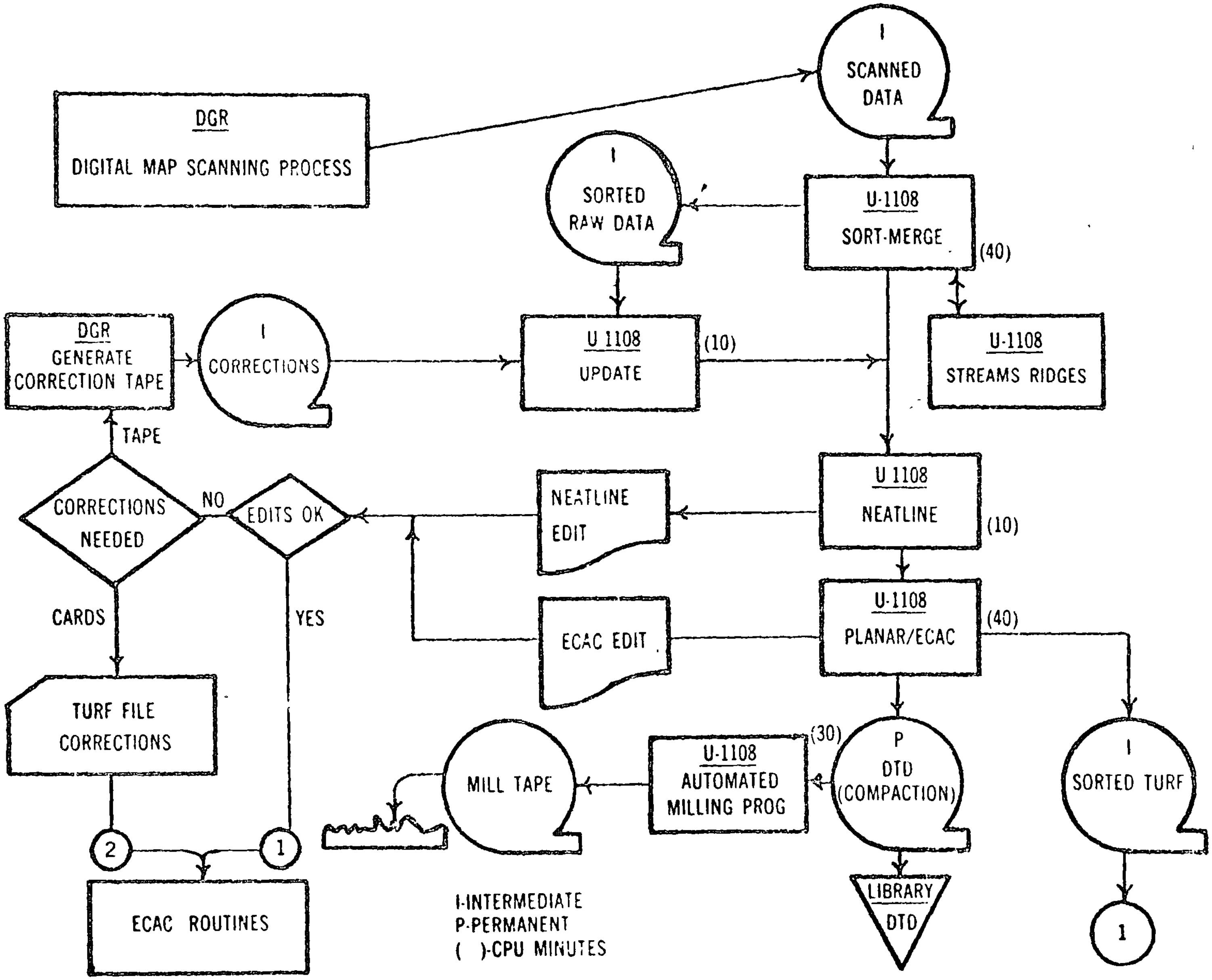
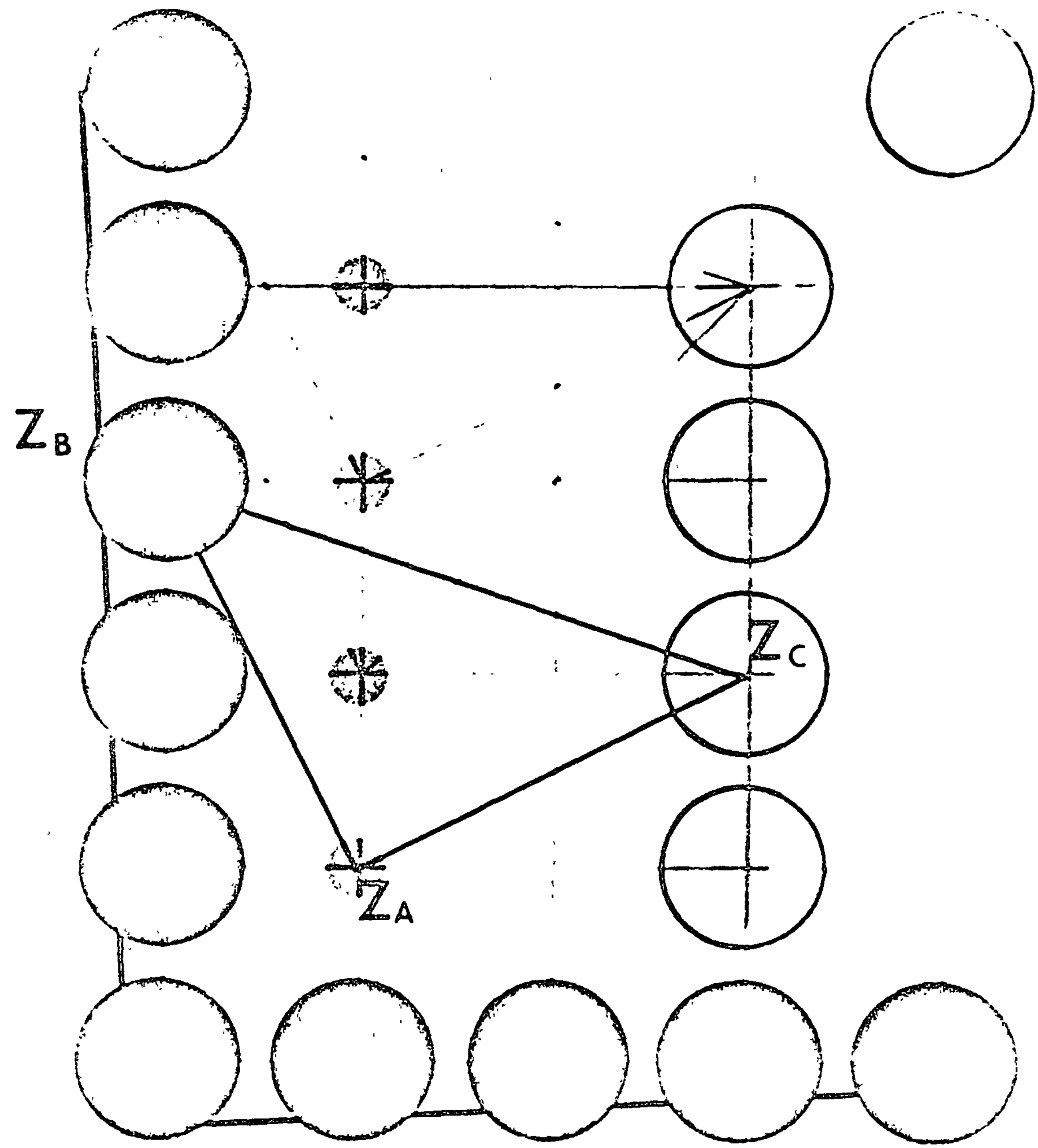


Figure 7. Digital graphic data flow.



$$Z = \frac{Z_A(y + x) + Z_B x + Z_C}{2x + y + 1}$$

Figure 8. Planar interpolation.

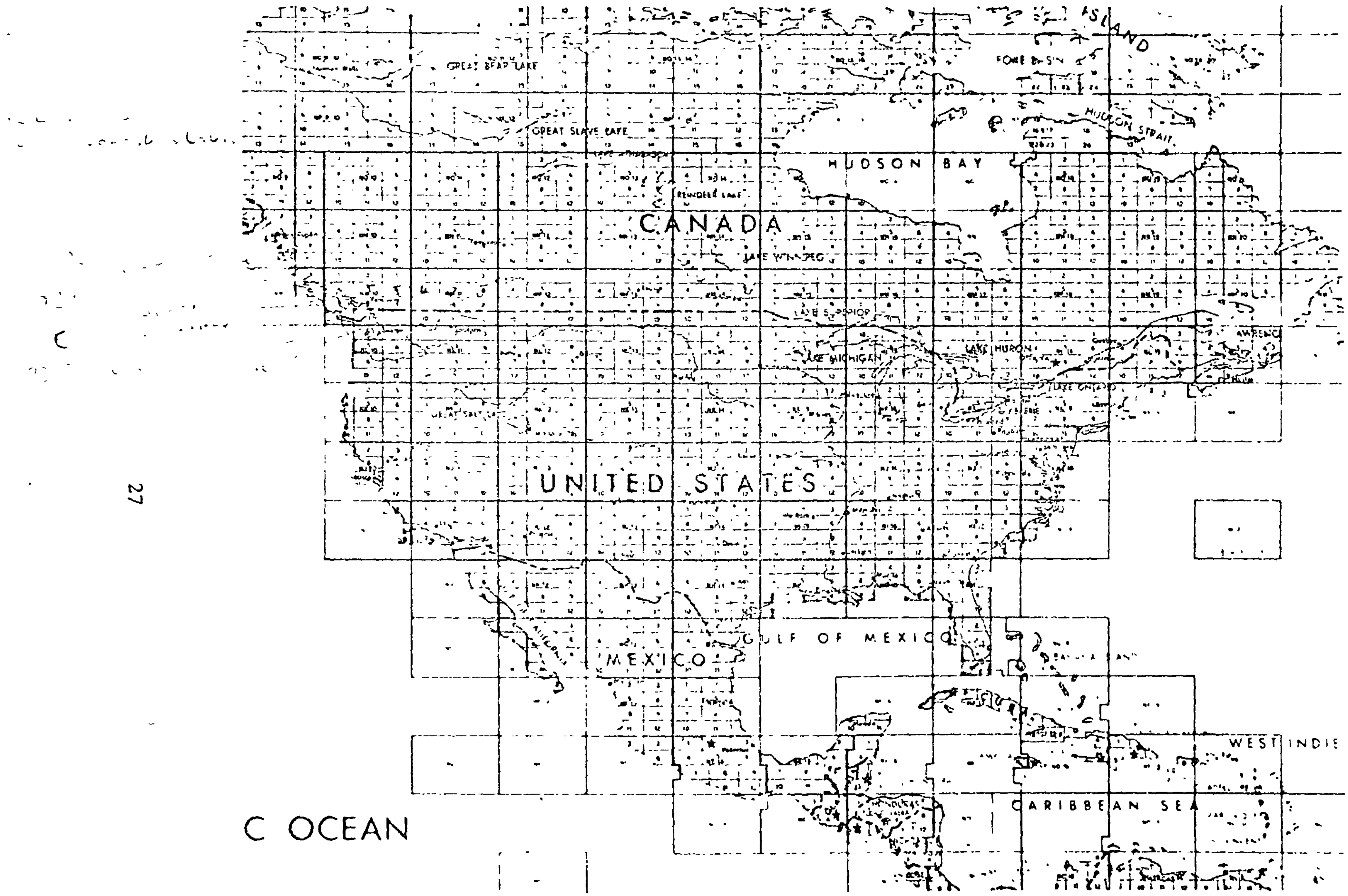


Figure 9. Index of digital elevation coverage of continental United States.

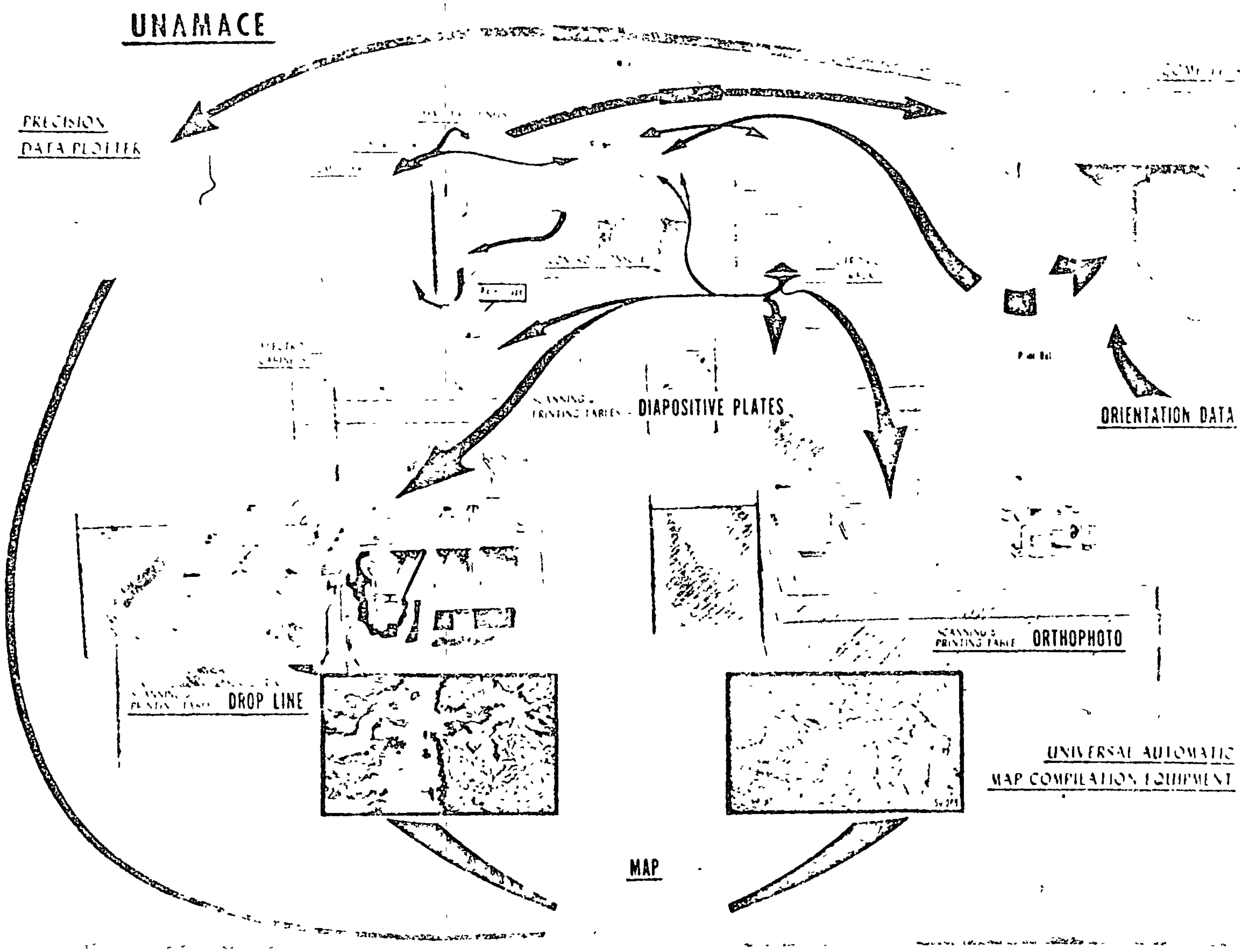


Figure 10. UNAMACE data flow model.

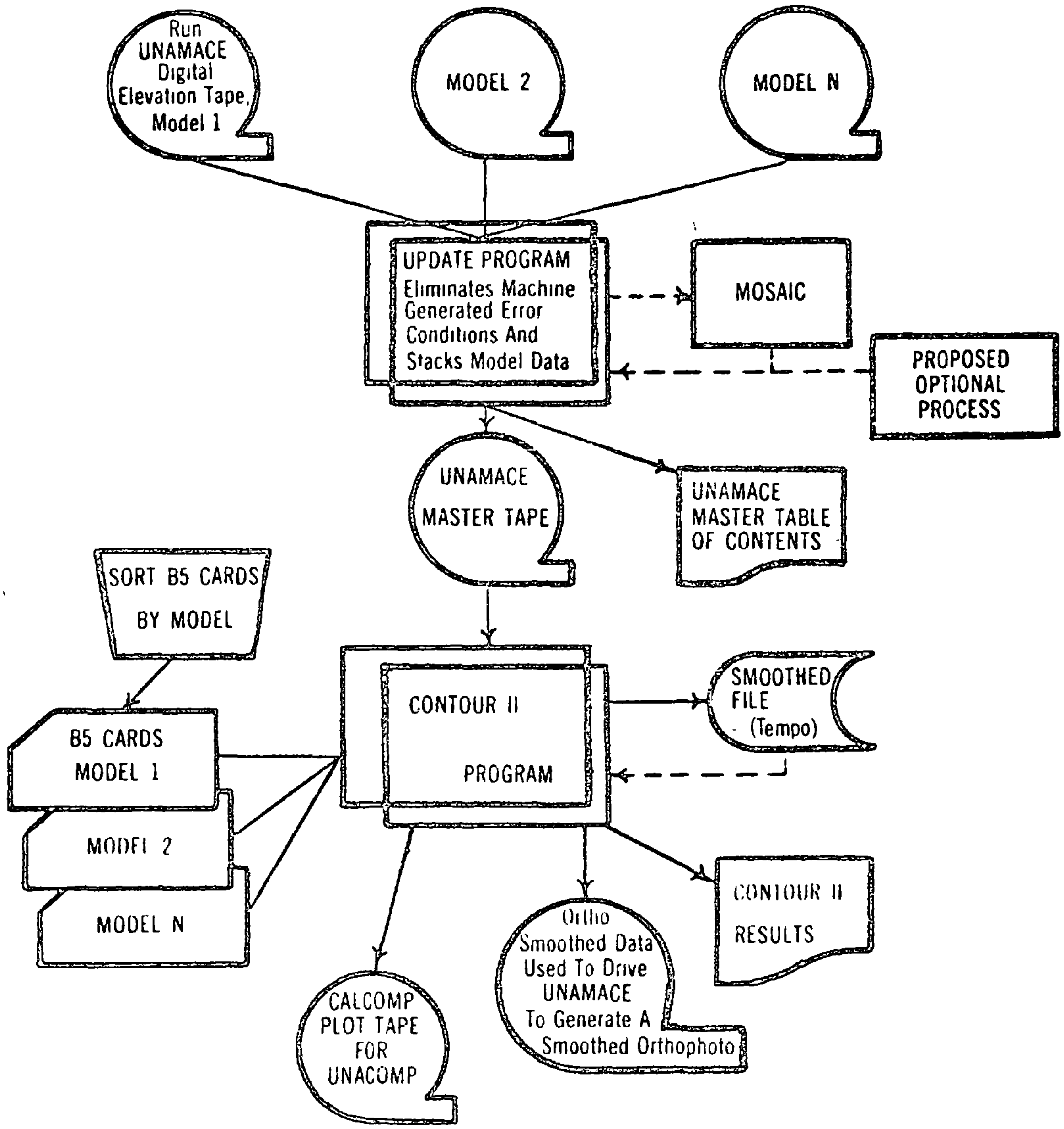


Figure 11. UNAMACE data flow chart.

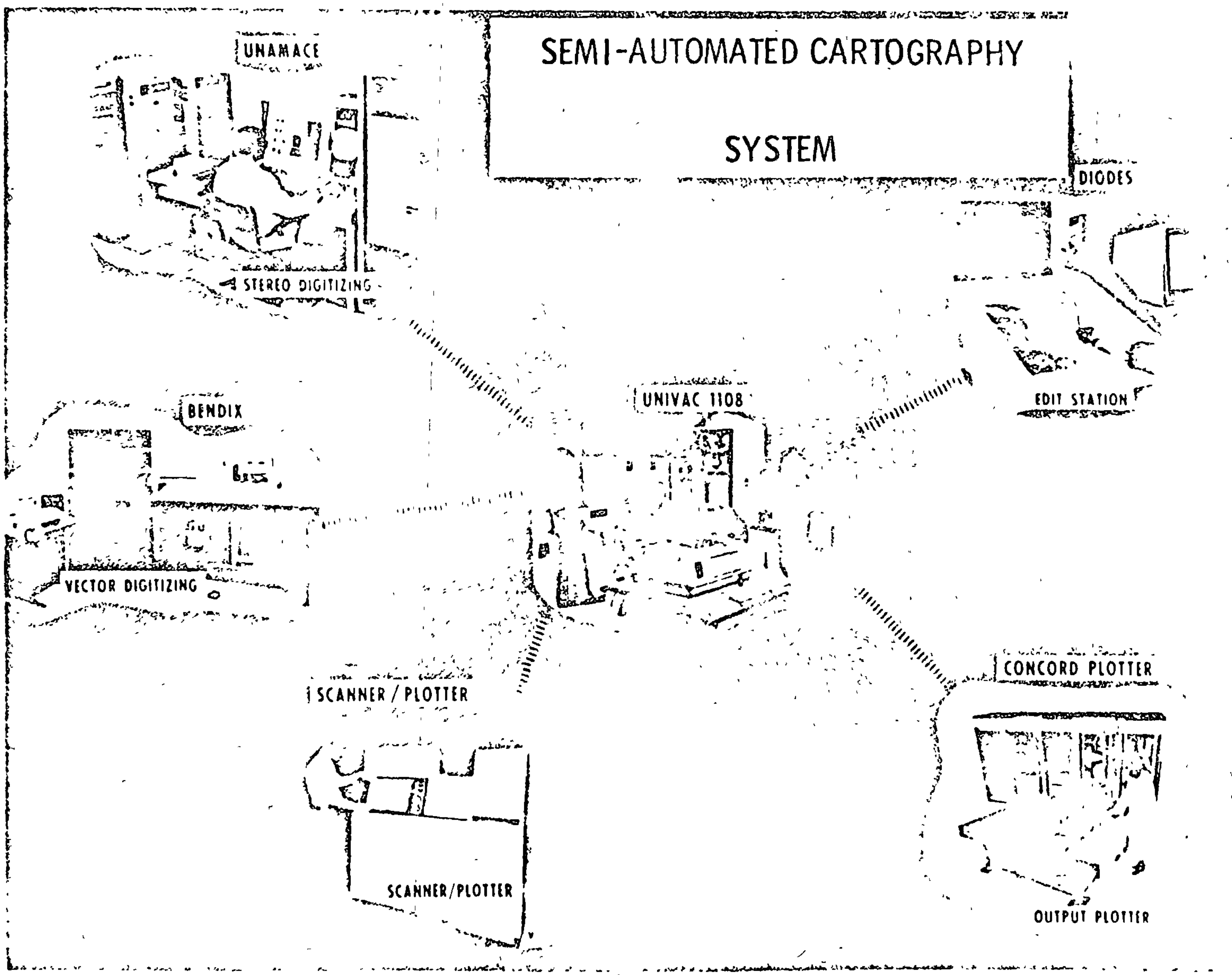


Figure 12. Semi-Automated Cartography System.

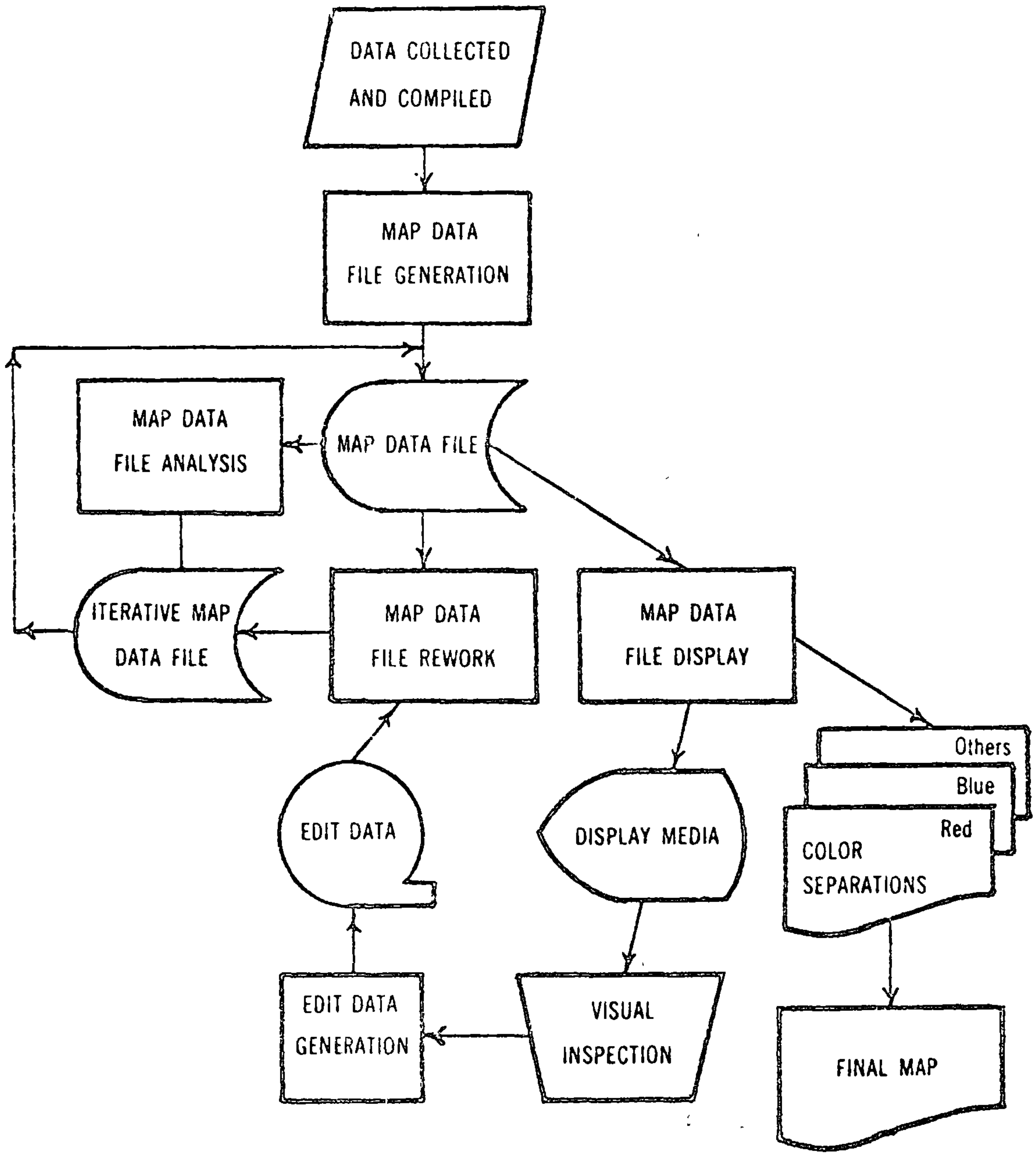


Figure 13. Graphic Improvement Software Transformation System data flow.

A Framework for Encoding Spatial Data*

Kenneth J. Ducker

K 452, Encoding, Polygons

Increasingly, geographers are encoding data for machine processing, yet replicating complex geographic patterns in machine-readable form for processing is poorly understood. This paper sets forth alternative methods for encoding geographic patterns.

Geographic information systems differ from other information systems in the explication of locational identifiers, and the importance of the locational identifiers in manipulating data. This paper accomplishes two tasks: 1) the development of a conceptual framework for encoding spatial data for incorporation into geographic information systems, and 2) the presentation of a notation for interpreting the framework. The choice among alternative ways of encoding spatial data is determined for a particular application by considering the purpose as constrained by the available hardware/software environment, quality control needs, and the magnitude of data to be encoded.

Locational Identification

Spatial data are conveniently stored in machine-readable form, not as the image itself, but in some abstract form, such as intensity values for small grid units or attributes of areal units.

The subsets of spatial data that are addressed here are those phenomena, such as land use, ownership patterns, and vegetation coverage, that can be exhaustively partitioned by type or class into regions described as polygons.

For example, assume that an aerial photograph has been exhaustively partitioned into a set of polygons that identify homogeneous land regions, i.e., residential, non-residential, agriculture, wooded areas, etc. Replication of these various sized regions in digital form usually is done as a sequence of coordinate values or equivalents describing the perimeter of each area. Problems and potentials of this mode of operation are described in more detail in the next section.

Encoding Data

Technical problems in the development of a geographic information system in the generic sense can best be approached by developing a conceptual framework for encoding geographic data, i.e., points, lines, and areas, and then evaluating alternative ways in which the encoded geographic data can be input, stored, retrieved, and output. Clearly, the major issue is the efficient manipulation of geographic data. Once that issue is determined, any number of data items related to a specific point, line or area can be included providing additional data storage is made available.

All elements of the information system including input and output must be viewed in the context of handling geographic data encoded as points, lines, or areas. For example, data might be abstracted from imagery in terms of coordinates for points, and areas encoded as a sequence of points. Data might then be stored as coordinate values for points making up areas, and the data might be displayed as line segments making up the system of areas. Clearly, the input, storage, and output encodings are not independent. Translation from one stage to another must be thought out in advance.

Figure 1 illustrates alternative ways of encoding geographic data. Depending upon ultimate needs, and the storage media and file structure environment, not all possible encoding schemes need be utilized. However, more than one is usually needed to provide a redundant coding for editing purposes to detect errors for quality control and completeness. Some encodings can also be generated from others, which provides the means for edit. Each of the procedures listed has a variety of advantages and disadvantages associated with it when viewed in light of the storage, comparison, retrieval, and output elements of the geographic information system.

The choice of an encoding method from the alternatives listed in Figure 1 must be made considering: (1) usefulness in terms of purpose, (2) ease of data capturing or encoding, (3) ability to generate other encodings for error detection, (4) ambiguities, causes by non-unique identification, (5) ease of synchronizing graphic data to descriptive data at input, and (6) storage media and file structure available. When comparing these criteria with the possible encodings, some encodings are more appropriate than others for typical uses of image data in machine-readable form. For example, areas encoded as polygons (encoded method #1) are highly useful as is, can be used to generate other encodings, and can be encoded with ease. Conversely, areas encoded as being contiguous to other areas (#2) is less useful and may not uniquely identify an area when two separate areas are wholly contained within the same larger area. Also, method #4 is difficult to encode directly, although the encoding can be generated from method #3. For error detection it is important to have two independent encodings, one of which can generate the other for comparison.

The input problem can be considered a problem of digitizing and coding a sequence of points that make up an area, or a line segment. Alternatively, the image could be scanned and encoded as small xy cells which are or are not a part of a line segment. The data storage problem is one of selecting the appropriate encoding method to meet retrieval needs and at the same time be compatible with the file structure and medium upon which the data are

*This study was supported by the Geographic Applications Program, United States Geological Survey under Contract 14-D8-001-12505. The author gratefully acknowledges the comments and suggestions of Samuel Arms, Frank Horton, Robert Shaw, and Robert Wittick.

Kenneth J. Ducker is associate professor of urban and regional planning and geography at the University of Iowa.

…ored. Pattern recognition type queries may require contiguity characteristics about small grid units be stored in a direct access mode so as to assemble homogeneous regions for comparison to a mask. Other queries might require comparison of two polygon sets, such as determining areas of vacant land use with areas of land suitable for industry. In this case the existing land use data set would be compared to the land capability data set. Thus data must be encoded to make area description and area characteristics accessible and comparable.

Converting Geographic Data to Machine-Readable Form

Specifically, three different input methods for converting geographic data to machine-readable form are utilized. The assumption here is that geo-

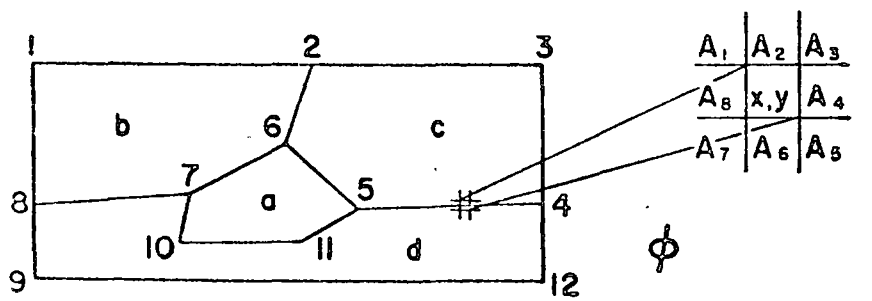


FIG. 1. Alternative Methods of Encoding Geographic Data.
Possible Encodings

1 Areas encoded as polygons made up as a sequence of points:
 $E(a) \cdot A(6, 5, 11, 10, 7, 6,)$ [Read as the entity is area *a* and attributes are a sequence of point numbers 6, 5, 11, 10, 7, and 6.]
 $E(b) . A(1, 2, 6, 7, 8, 1)$
2 Areas encoded as being contiguous to other areas (where ϕ is the area outside the area system):
 E(b) :A(c, a, d, ϕ)
 E(b) :A(c, d, b)
3 Line segments encoded in relationship to their end-points and their contiguity to areas
 $E(1,2) : A(b,\phi)$
 $E(1,8) : A(\phi,b)$
 $E(2,3) : A(c,\phi)$
 $E(2,6) : A(b,c)$ } reverse encoding for redundancy edit
 $E(6,2) : A(c,b)$
4. Points encoded as being connected to other points
 $E(1) : A(2,8)$
 $E(2) : A(1,3,6)$
5 Points encoded as being related to areas
 $E(1) : A(\ \ b)$
 $E(2) : A(\ldots,c,b)$

graphic patterns are manually identified as line segments that exhaustively partition the map or image into areas. The three ways of making the resultant polygons machine readable are:

1. point digitizing polygon vertices,
2. polygon trace with line follower digitizer, and
3. scanning for the presence and absence of lines making up polygon regions.

In all three cases the image must be partitioned into areas representing phenomena, such as land use, land capability or jurisdictional areas. The problem is to describe these areas in machine readable form for internal computation of intersections between phenomena or data sets and mapping of the original data or some derivative therefrom.

Point digitizing is the most direct way of making area data machine readable and requires the least hardware and software capability. Operating in this mode all polygonal vertices are digitized and in a separate operation all the polygon vertice numbers are coded, and sometimes for edit check the line segment, i.e., connectivity between vertices are coded. The resultant entities for this encoding process are:

 E (vertice number) : *A* (x,y) ;
 E (polygon) : *A* (an ordered sequence of vertice numbers bounding the polygon) ;

6. Small grid units encoded as whether part of line segments or not E(x,y) :A(0 or 1) absence or presence of being on a line segment.
7. Small grid units encoded as whether they are same as contiguous grid units (E(x,y) : $(A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8) = (0, 0, 0, 1, 0, 0, 0, 1)$
8. Point Connectivity Matrix

	1	2	3	4	5	6	7	8	9	10	11	12
1	0	1	0	0	0	0	0	1	0	0	0	0
2	1	0	1	0	0	1	0	0	0	0	0	0
3	0	1	0	1	0	0	0	0	0	0	0	0
4	0	0	1	0	1	0	0	0	0	0	0	1
5	0	0	0	1	0	1	0	0	0	0	1	0
6	0	1	0	0	1	0	1	0	0	0	0	0
7	0	0	0	0	0	1	0	1	0	1	0	0
8	1	0	0	0	0	0	1	0	1	0	0	0
9	0	0	0	0	0	0	0	1	0	0	0	1
10	0	0	0	0	0	0	1	0	0	0	1	0
11	0	0	0	0	1	0	0	0	0	1	0	0
12	0	0	0	1	0	0	0	0	1	0	0	0

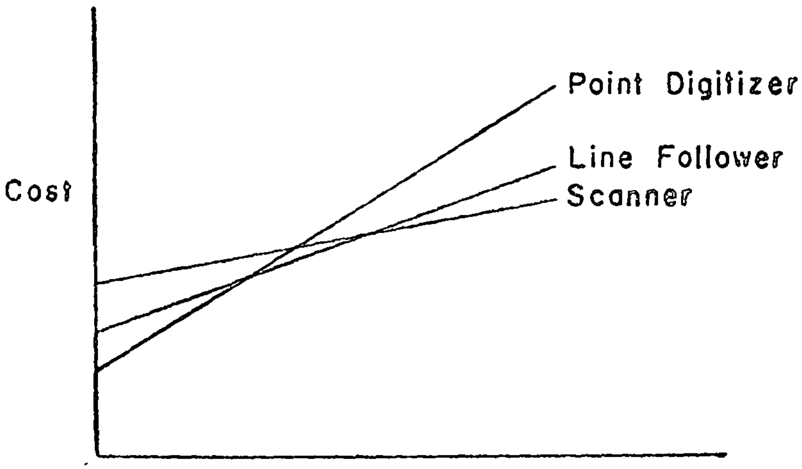
9. Area Connectivity Matrix

	a	b	c	d
a	0	1	1	1
b	1	0	1	1
c	1	1	0	1
d	1	1	1	0

10. Coordinate definition of:
 a. points
 b. area bounds
 c. area centroids

e follower mode has a higher computer processing cost reflecting proces- ig to determine points of inflection and junction. Operating in the scanner ode also involves extensive computer processing to determine end points line segments. There are high costs involved in suitable partitioning of ta, topological comprehension, and ambiguities of imagery such as rela- e positional errors.

The scanning mode and the line follower mode have higher initial costs terms of the development of algorithms to convert the data for storage and rieval, but operate more efficiently at higher volumes of data from imag- . Figure 3 represents a conceptual cost comparison of three modes. Sup-



3 Hypothetical Cost Comparison for Alternative Methods of Input Processing of ographical Data.

ing cost comparison data is not readily available though it needs to be ed.

e volume of data is a function of the size of the area being digitized, of the image, the resolution being used, the density of phenomena be- igitized, and encoding method:

Volume = f (size, scale, resolution, density, encoding method).

hen the geographic data are densely distributed on imagery and there large number of images to be processed the scanning mode becomes mically feasible. When gross or less dense patterns occur on images, encoding of area definitions and point digitizing appears most feasible.

Closing Note

s paper attempts to explicitly treat aspects of making (graphic data machine-readable that are usually implicitly considered. It stems from the author's view that independent decisions regarding one element of an information system, say input or storage, force decisions pertaining to another element, say retrieval or output, that may be contrary to the purpose of the system.

Viewing ways in which spatial data can be encoded provides the connective tissue by which integrated geographic information systems can be designed. Similarly, the encoding concepts provide a means by which systems can be compared.

Literature Cited

1. Arms, S. "The MAP/MODEL System: System Description and Users Guide," Bureau of Governmental Research and Services, University of Oregon, May, 1970.
2. Tomlinson, R. F. "An Introduction to the Geo-Information System of the Canada Land Inventory," Canada Department of Forestry and Rural Development, Ottawa, Canada, 1967.
3. U. S. Bureau of the Census. *Census Use Study. The DIME Geocoding System*, Report No. 4, Washington, D. C., 1970.

The Trade Area of a Displaced Hexagonal Lattice Point

M. F. Goodchild

Consider the classic central place system of uniform demand density on an infinite plain. Let us suppose that a single entrepreneur is, for some undefined reason, unable to locate at his optimal hexagonal lattice point. He is forced instead to locate at a distance r away in a line bisecting the angle between two of his nearest neighbours (Figure 1). We adopt a coordinate geometry in which the distance separating adjacent lattice points is 1, the optimal point (0, 0,) and the six neighbours are located at $(\pm\sqrt{3}/2 \pm 1/2)$ and (0 ± 1). Our unfortunate friend has been forced to locate at $(r,0)$.

The upper half of the modified trade area of point 0 is found by connecting the perpendicular bisectors of 0A, 0B and 0C to form $KLMN$. The lower half is symmetrical. By a little application of the principles of coordinate geometry, we find that the trade area is

$$\frac{9\sqrt{3}}{2(3-r^2)}\frac{(1-r^2)^2}{(3-4r^2)}$$

M. F. Goodchild is an associate professor of geography (he University of Western Ontario, London, Canada.

:(line segment: vertice *i*, vertice *j*) : *A* (polygon code for left adjacent area, polygon code for right adjacent area).

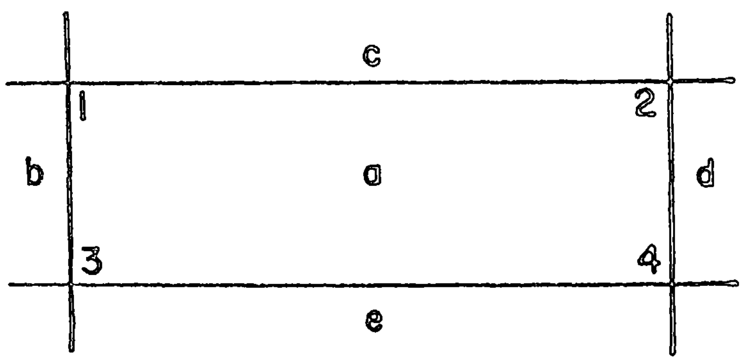
'he DIME Geocoding System (3) and MAP/MODEL (1) both use iations of this encoding scheme. Initially, only the line segments are en-ed in the DIME mode. Polygons are in effect created as the DIME al-ithm links line segments based on polygon (block) codes and redun-itly checks by matching-up the block numbers, as illustrated in Figure 2. 'he Primal edit sorts boundary segments in order by matching node num-s. The edit traverses the boundary of the block by linking the segments ether in order by matching node numbers.

'he redundant edit also traverses the boundary, with the block left and ck right fields being switched if the nodes have to be reordered to see if block number appears entirely in the block right field.

'he MAP/MODEL System encodes the points of inflection and junc-is of polygons. The resultant entity is an ordered set of vertices for each ygon:

E (polygon code) : *A* (*x,y* coordinates of points of polygon inflection and junction).

'he edit procedure used in MAP/MODEL utilizes the characteristic of method that encodes each line segment twice (because it is part of two



. 2. DIME Edits.

'IME Coding [*E*(*i* node, *j* node)]: *A* (left block; right block))

Node *i*	Node *j*	Block Left	Block Right
1	2	c	a
1	3	a	b
3	4	a	e
4	2	a	d

'rimal Edit (traverses boundary) *Redundant Edit* (block encoding)

Node *i*	Node *j*	Block Left	Block Right
1	2	c	a
2	4	d	a
4	3	e	a
3	1	b	a

polygons). The separately encoded line segments are compared and if they fall within a given tolerance are accepted; if not the records are rejected and a correction must be made.

Input from imagery via a polygon trace with a line follower digitizer is sometimes used to digitize a coastline, contour lines or irregular polygons. This eases the input problem significantly with resultant ease of instructions to the digitizing personnel. The only instruction necessary is that the digitizer stylus trace around each polygon. The resultant entity:

E (polygon code) : *A* (*x,y* coordinates sequentially ordered and at regular intervals)

In some instances, a computer algorithm must then be developed to identify junctions and points of inflection to create a polygonal file changing the attribute structure of a sequence of *x,y* coordinates to a sequence of line segments, thus simplifying the polygonal description similar to one developed by the point digitizer and manual coding. The edit employed by this method utilizes the characteristic of the method that encodes each line segment twice. The separately encoded line segments are compared and if they fall within a given tolerance are accepted. The resultant record for storage and retrieval is the same as the resultant records from point digitizing.

A drum scanner as used by the Canadian Land Inventory (2) creates machine-readable records for each resolution cell of the image, consisting of whether a line is present or absent for that cell:[1]

E (*x,y*) : *A* (presence or absence of a line segment for *x,y*)

An algorthm is necessary to connect contiguous cells having lines present to form line segments in order to proceed as above. For identification of polygons, a separately encoded polygon number and coordinate value falling within the polygon region must be digitized separately:

E (point within polygon) : *A* (*x,y*).

Again the resultant data for storage and retrieval consists of a polygon description and/or line segment description with derived coordinates for the vertices of the polygon regions.

The choice of encoding methods for imagery data is dependent upon a variety of factors, that are largely related to the magnitude of the data being encoded which in turn is a function of the size of the area and the degree to which it is being partitioned into subareas, and the availability of hardware and software. For example, the Canadian Land Inventory procedure requires redrawing maps to be scanned to eliminate all detail but the boundaries of areas.

Operating in the point digitizing mode the costs for digitizing and polygon encoding are high whereas costs for computer processing are low because the encoded data are directed to storage without intervening processing. The

[1]As the "line's" thickness exceeds the cell width the edge must first be determined before the "line" is useful.

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

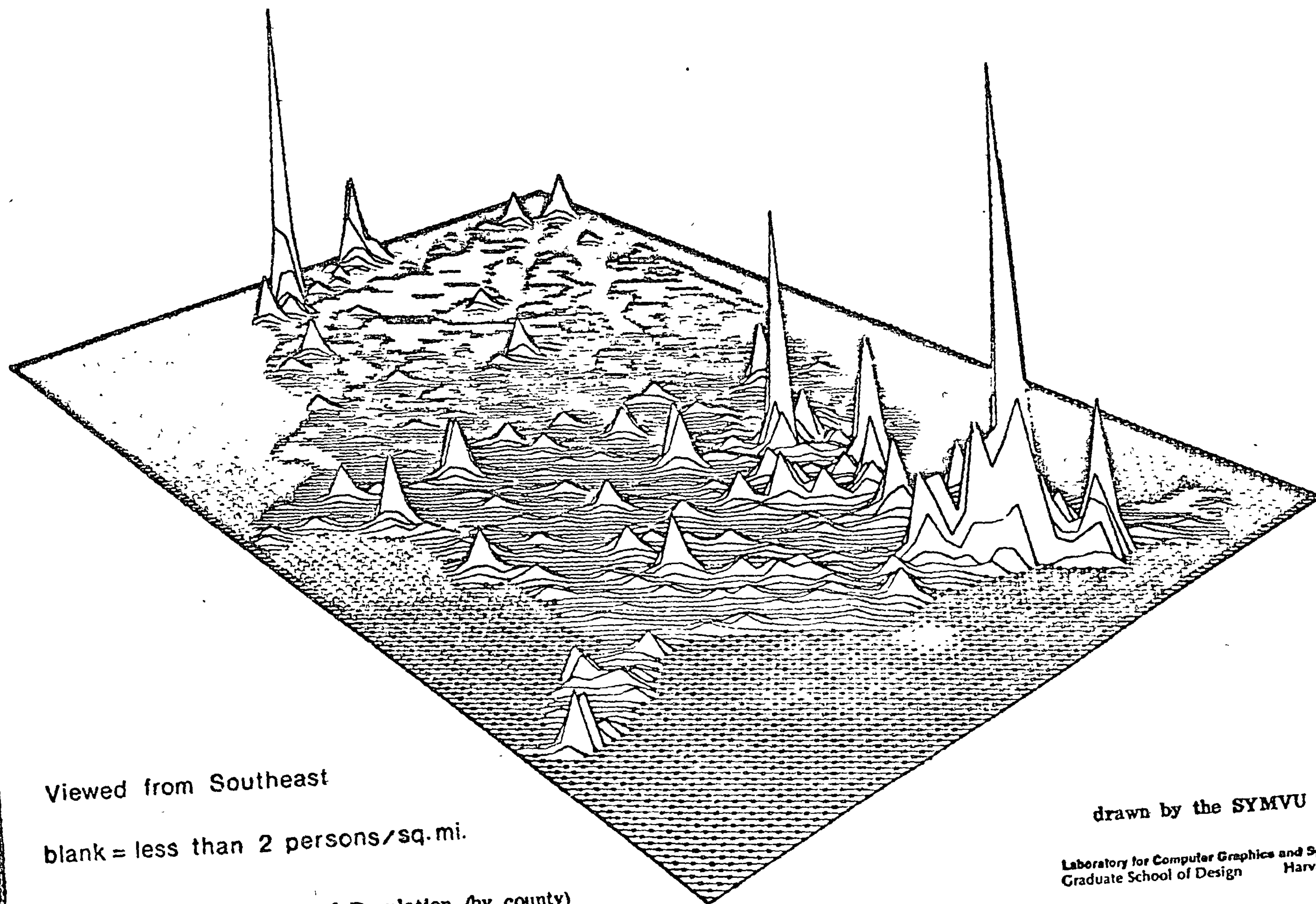
MANEJO DE SISTEMAS DE INFORMACION GEOGRAFICA EN PLANEACION

PROGRAMAS DISPONIBLES EN EL LABORATORY FOR COMPUTER GRAPHICS AND SPATIAL ANALYSIS UNIVERSIDAD DE HARVARD

NOVIEMBRE DE 1977.

Palacio de Minería Calle de Tacuba 5, primer piso. México 1, D. F.

UNITED STATES POPULATION DENSITIES, 1970



Viewed from Southeast

blank = less than 2 persons/sq.mi.

source: U.S. 1970 Census of Population (by county)

drawn by the SYMVU program

Laboratory for Computer Graphics and Spatial Analysis
Graduate School of Design Harvard University

CONTENTS

INTRODUCTION

RESEARCH DIRECTIONS

OPERATING POLICIES

PROGRAM DESCRIPTIONS

NEW PROGRAMS

CARTOGRAPHIC DATA BASES

PUBLICATIONS

ORDERING INFORMATION

Brief History of the Laboratory

The Laboratory for Computer Graphics and Spatial Analysis was established within the Graduate School of Design at Harvard University in the Spring of 1965 In December of that year it received a major grant from the Ford Foundation

The original and continuing goals of the Laboratory are

1. To design and develop computer software for the analysis and graphic display of spatial data
2. To distribute the resulting software to governmental agencies, educational organizations and interested professionals
3. To conduct research concerning the definition and analysis of spatial structure and process

The Laboratory was founded by Howard T Fisher, who served as its Director until June 1968 During this time the SYMAP, SYMVU, and GRID programs were developed and made available to practicing professionals In addition, several short courses and conferences were sponsored Succeeding Fisher, William Warntz was Director of the Laboratory until June 1970. Under his direction, the Laboratory staff focused on research in spatial structure, and began to publish their findings in a newly initiated series of Theoretical Geography Papers

In 1970 Allan Schmidt assumed responsibility for the Laboratory Under his direction several additional software packages including CALFORM and POLYVRT were introduced Extensive work was also begun on issues concerning cartographic data structures (POLYVRT and ODYSSEY) and interactive graphics (INPOM and ASPEX) In July 1976 Brian Berry became Director of the Laboratory A substantial increase in emphasis on research relating to spatial process and geographic information systems is planned

RESEAR DIRECTIONS

Research within the Laboratory covers a wide range of activities related to the analysis and display of spatially variable phenomena. Of particular interest are:

1. Theories and techniques applicable to the display of spatial data,
2. Theories intended to provide a better understanding of spatial structure and process, and
3. The distribution of resulting materials to members of planning and design professions as well as cartographers and geographers.

Display of Spatial Data (Computer Graphics)

An initial goal of the Laboratory was to develop low cost capabilities for producing computer maps. This is still an objective of the Laboratory, both in terms of updating earlier systems and developing new ones

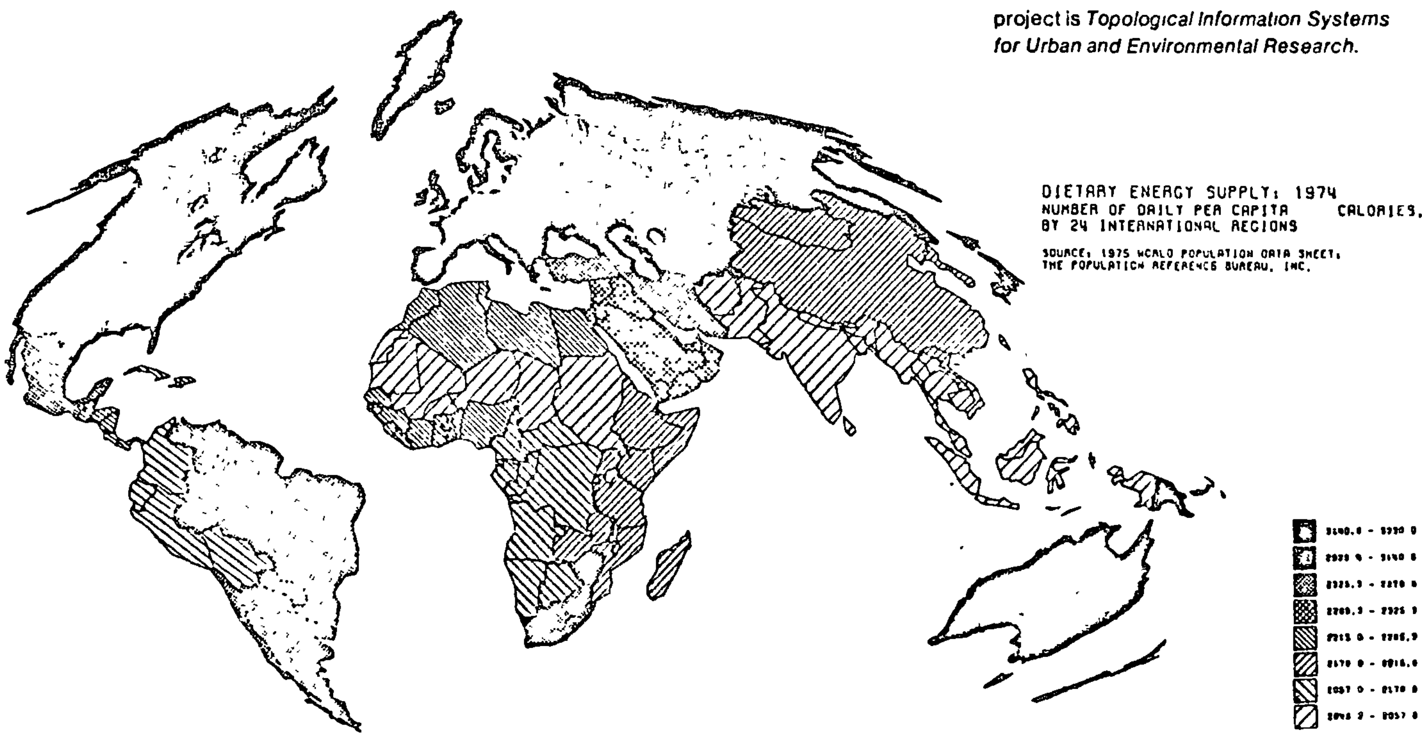
The original product, SYMAP, has continued and developed. Its system of display on a line printer led to a second generation of line plotter programs (SYMVU and CALFORM). Recently, a third generation (ASPEX and INPOM) has evolved towards an interactive cathode-ray tube environment. In the future, graphic software developed by the Laboratory will be increasingly interactive and intended for use with small to medium size computers

Understanding Spatial Structure (Spatial Analysis)

Recent progress in the display of spatial data has depended increasingly on an understanding of spatial models and their impact upon the storage, manipulation and display of x-y coordinate cartographic data bases. These files contain the basic locational information required to produce a map. If a map is to represent a spatial concept, locational attributes of that concept must first be transmitted to the display program by means of a cartographic data base.

Initial efforts in cartographic data base research resulted in design and development of the POLYVRT program. This program has the ability to translate files from one data structure and format to another — preserving the inherent topology of each file yet eliminating errors and redundant detail

Work on this project is currently being funded under a 3-year grant from the National Science Foundation The title of the project is *Topological Information Systems for Urban and Environmental Research*.



Recent Laboratory research on cartographic data structures has focused on topological analysis of spatial entities where basic relationships of connectedness provide the organization for a data structure Such a data structure provides efficient data storage, flexible data retrieval and extensive error checking capability.

A topological data structure also allows for the creation of least common geographic units (LCGUs) which result from the superimposition of two or more partitionings of a region As a result, direct overlay of arbitrary data zones in a region becomes feasible. This allows immediate comparison of different data bases without the complications of more traditional procedures for uniform gridding of a region Current activities include utilization of this approach in the fields of land use and environmental planning and thematic mapping

The Laboratory has acquired expertise in the manipulation of global cartographic files (WDB-I), U S Census DIME files for metropolitan areas and counties, U S Census Urban Atlas files of census tracts, and a number of locally generated files. This research embodies both thematic mapping concerns of display research and other issues related to automated cartography and geographic information systems

While each program can be characterized in terms of hardware requirements, these technological factors are less important than the conceptual model of spatial structure which influences its design The design of a program is largely a function of how one describes and classifies measurable entities on the earth's surface As a result, each program embodies a model of spatial structure which is reflected in the requirements and capabilities of that program

The development and dissemination of cartographic data bases by the Laboratory is intended to facilitate the preparation of maps by users of the Laboratory's computer mapping programs

Tools for the Profession

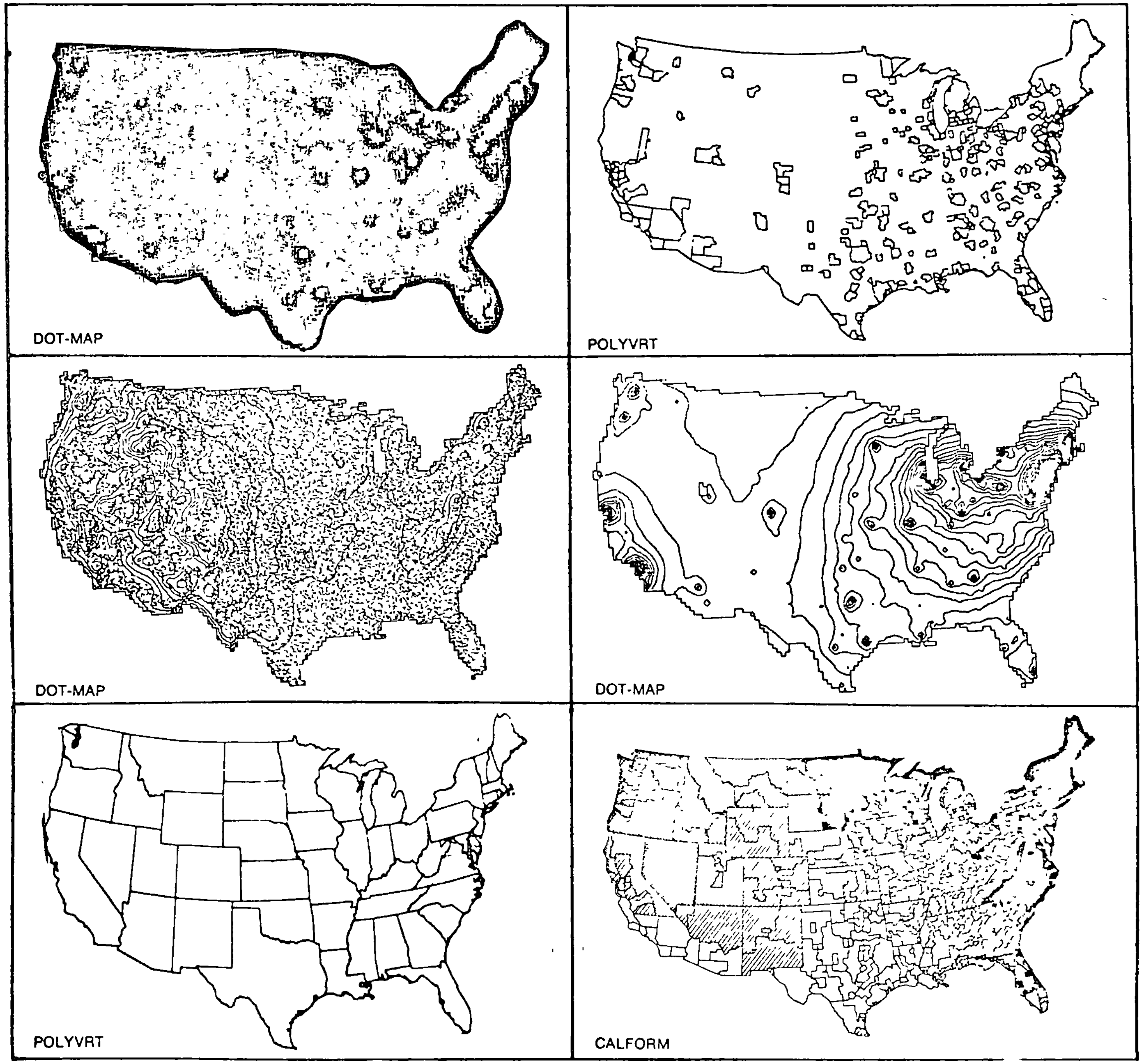
The Laboratory distributes the results of its research to academic, governmental, and commercial organizations interested in the application of spatial analytic techniques to planning and design. Available programs and publications are described in this catalog. Recent activities and developments are reported in the Laboratory's newsletter, CONTEXT, which is distributed free of charge. Summarys of research projects with appropriate illustrations are published in the Laboratory's "Red Book," a cumulative report of selected projects.

The Laboratory's applied research efforts focus upon selected projects that allow for the immediate application of new theories and techniques in a working environment. Such projects provide opportunities for evaluating and refining Laboratory products in realistic test situations. As a result, user requirements are directly reflected in the design of a spatial model as soon as possible. Feedback from our software user community also provides a significant contribution to our on-going program development efforts.

Though the Laboratory's programs are designed for use by individuals without prior training or experience in computer science, the software allows increasing sophisticated user application by those with a computer science background or automated cartography experience. The programs are user oriented with considerable flexibility and numerous options within each program. There is also a default procedure to satisfy most common requirements of a user's data base and also to minimize difficulty of obtaining an initial computer output for which sample test data is provided with the programs.

A recent addition to the list of materials available from the Laboratory are a relatively complete set of Cartographic Data Bases for use with a conformant mapping program like SYMAP, CALFORM or INPOM. Such data bases are currently available for a wide variety of geographic locations.

Selected views of the United States



The Laboratory's relationship with its user community has resulted in the establishment of operating policies concerning software distribution, support services, computer hardware, standards, research grants and contracts, the Harvard Mapping Service, and pricing.

Software Distribution

The Laboratory distributes a set of computer programs described later in this publication. In the future, new programs will be added to this list as a result of research efforts outlined above. Distribution of these programs is subject to the following agreement between the Laboratory and users of its software:

1. Neither this software, its documentation, nor adaptations thereof shall, except with prior written consent of the Laboratory, be sold, leased or otherwise distributed in any form to any individual, business entity, academic institution or governmental body whatsoever
2. This software is not to be installed at a commercial computing installation (service bureau) in a manner which would allow for its use by individuals other than employees of the purchasing organization without prior written consent of the Laboratory
3. Upon acceptance of these terms and conditions, as indicated by signature of an officer having authority to enter into such agreements, the Laboratory grants the recipient a royalty free, non-exclusive license to use the subject material at a single computer facility

Support Services

With the sale of a program, the Laboratory undertakes to provide assistance and support for the program's installation and use. Through increasing adoption of software standards and assistance of software conversion centers (see below), installation effort is minimized.

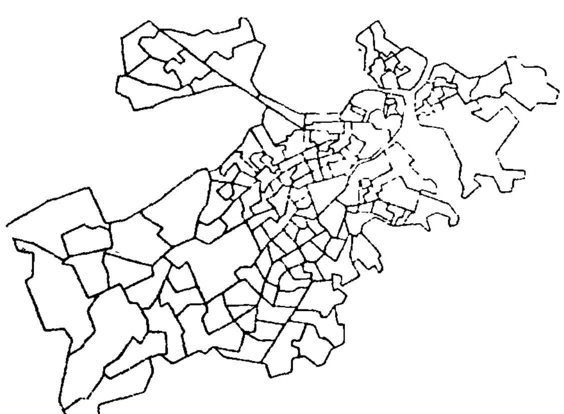
Questions concerning use of a program, however, will always exist. The utility of programs for a large user community depends on early identification and diagnosis of problems and a corresponding development of corrections and/or modifications. Such a process benefits the entire user community by eliminating errors and extending the capabilities of a total program package. Support provided free of charge by the laboratory for its programs consists of.

1. Responses to written or telephone communications concerning installation or operation of Laboratory software
2. Modifications to the program, printed and distributed to all users. Provision for a complete program replacement tape is made at a charge equal to 10% of a program's current purchase price. This option is frequently desirable when modifications to an existing program involve several hundred changes within the original source code

Computer Hardware

All of the Laboratory's programs are distributed for use on IBM hardware in FORTRAN source code. However, most of the Laboratory's programs also have been converted by certain other users to operate on a wide variety of non-IBM hardware. For most non-IBM hardware the Laboratory has available names and addresses of users who already have performed conversions and have agreed to make available a copy of their versions of the Laboratory's programs to authorized users

The Laboratory is interested in establishing formal relationships with organizations willing to act as conversion centers for the Laboratory's software with respect to a specific hardware manufacturer. Two such conversion centers have recently been established, the University of Massachusetts for all CDC users and the University of Delaware for all Burroughs users. Each designated conversion center receives a free copy of all Laboratory software. These centers prepare copies of the Laboratory's software for their respective hardware and will be able to respond to user questions concerning its installation on these machines. As outlined above, however, the Laboratory retains sole right to control program distribution by these conversion centers



Boston 1970 Census Tracts

Atlanta 1970 Census Tracts

Standards

The position of the Laboratory in the profession of automated cartography is unique in terms of standards Decisions concerning software conventions affect a large number of users and have maximum utility only if applicable to a large segment of the user community

Specific areas which the Laboratory has selected for establishing conventions for its software include

1. Glossary of terms related to automated cartography
2. Standardized subset of transferrable FORTRAN and recommended procedures for localized groupings of non-standard features For example, machine specific constructs to be avoided or input/output units to be assigned at a specific installation
3. User command language for flexible and consistent naming of program control parameters.

Research Grants and Contracts

Support for the Laboratory's activities is derived solely from research grants and contracts plus program and publication sales The Laboratory actively solicits support for its various research interests, described above Organizations which have sponsored work of the Laboratory include:

1. The Ford Foundation
2. The Office of Naval Research
3. The National Science Foundation
4. The U S Departments of Housing and Urban Development, Health, Education and Welfare, Commerce, Interior, and Defense

Research which is funded by outside organizations must satisfy criteria administered by Harvard University's Office for Research Contracts These criteria include the freedom to publish research findings subject only to established safeguards for protection of privacy or confidentiality of personal data

Harvard Mapping Service

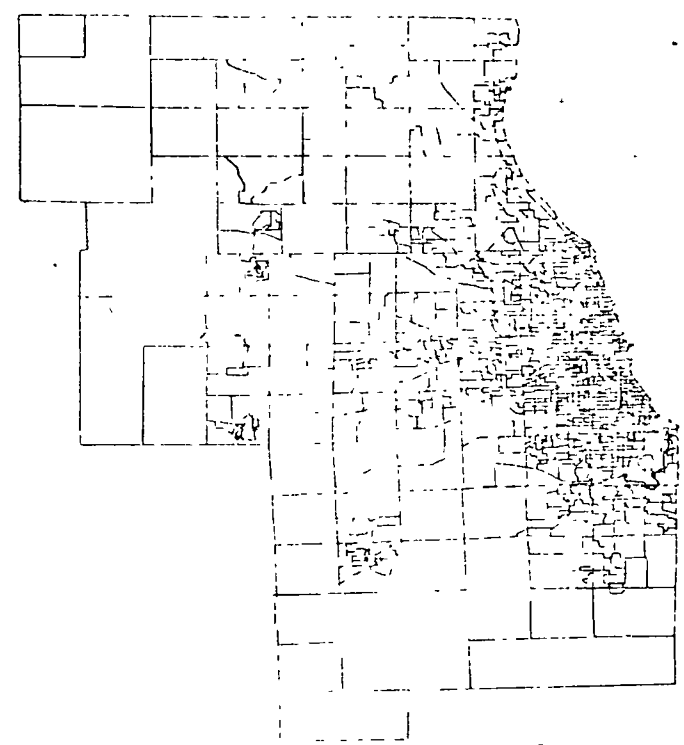
Occasionally the Laboratory has prepared maps, either by hand or by computer, to meet the needs of Harvard and MIT faculty and graduate students. When the diversity of experience thus gained can contribute to the general objectives of the Laboratory, mapping work may be undertaken for others as well Assignments of this kind have been carried out for such publications as *Life Magazine, Scientific American, The New York Times, National Geographic,* and for individual authors of various books and articles.

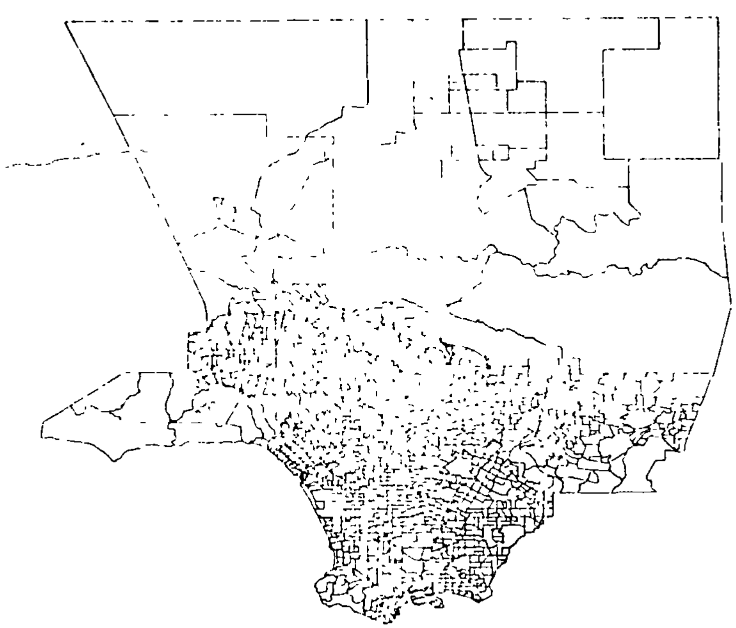
Interested organizations are encouraged to contact the Laboratory describing in as much detail as possible the nature of their needs and any applicable deadlines Enclose if practicable, copies of data tabulations, base maps, or any other material that could aid in the understanding of what is required

Pricing

Program products of the Laboratory are distributed at prices described in this catalog Substantial discounts are available to governmental agencies and educational institutions. Prices are subject to change without notice Purchase of a Laboratory program entitles a user to a copy of the current version of the program, sample data for test purposes, a user's reference manual, future updates and written instructions concerning the installation and use of the software Software prices are established to cover program development, documentation, future enhancement, user assistance plus distribution.

Chicago 1970 Census Tracts →





← Los Angeles 1970 Census Tracts

Program Name ..... **SYMAP, Version 5.20**

Computer Language .. .. FORTRAN IV
Computer Requirements
Machine .... IBM 370
Memory . ... 128K bytes
Mode .... Batch processing
Peripheral Devices
Four temporary data sets
Standard line printer

Description

SYMAP (SYnagraphic MAPping system) is a general purpose graphic display program for the mapping of spatially disposed quantitative and qualitative data It is the most widely distributed program of its kind and is used by city planners, economists, geographers, and others

SYMAP produces maps on a standard line printer Varying shades of gray representing value class intervals are created by overprinting

SYMAP creates five basic types of maps

1. *Conformant* — Maps that display data by predefined zones, e g , counties or census tracts. Input required consists of a set of x-y coordinates defining the outline of each zone and a data value for each zone.
2. *Contour* — Maps that graphically represent a continuous surface which has been computed by interpolation from data values at specified data points The value computed for each point on the surface is influenced by the values at the nearest data points and the weights assigned to these points These weights are based on $1/d^2$ where d is the distance from a data point to the point for which a value is being interpolated The weights are modified as a function of the spatial distribution of the data points around the interpolation point Input required consists of an x-y coordinate and a data value for each data point Papers describing the contouring algorithm in greater detail are available from the Laboratory
3. *TrendSurface* — Maps that graphically depict a polynomial surface of a specified order fitted to a set of data points and their data values A surface is constructed by minimizing the sum of the squares of the differences between the known data values and the computed value of the surface at each data point. Surfaces from a first order ($z=ax+by+c$) to a sixth order ($z=ax^6+bx^5y+cx^4y^2+ \cdots +fy^2+gx+hy+i$) polynomial can be calculated Input required is the same as for contour maps
4. *Residual* — Maps that describe a surface created by subtracting a trend surface from an interpolated contour surface where both surfaces were derived from the same set of data values Input required is the same as for contour maps
5. *Proximal (or Thiessen polygon)* — Maps that depict zones which are created from data points by the nearest neighbor method, i e , the value of any point on the surface is the same as the data value of the closest data point Input required is the same as for contour maps

In addition to the required input, SYMAP includes options for creating map cosmetics (legends) such as. graphic scale, place names, rivers, bodies of water, transport routes, city locations, and other point, line or area symbols There is also an option for specification of a study area outline which delineates a geographic boundary and displays interpolated data only within that outline

In SYMAP, numerous electives provide control over virtually every visual aspect of the output map For example, electives are used to specify the physical size of the map to be produced, coordinates of the display window (which allow close-ups or inset maps of sub-areas), the number, range, and symbolism of value class intervals, and other features

Electives also are used to influence the computation of the map by controlling the interpolation or calculation methods that produce proximal, contour, trend surface, residual or conformant map output Other electives modify the contour interpolation process by determining the number of data points which the program should search for within the vicinity of each interpolation point, the maximum distance to search for data points, and the amount of extrapolation to allow if the spatial distribution seems to warrant it

SYMAP reads the x-y coordinate and value data in a standard fixed format but it is possible to supply a subroutine FLEXIN (FLEXible INput) at program load time to allow for reading data in non-standard formats, extracting data from a data bank, or manipulation of data if desired

SYMAP can also produce a binary file of the data zone outlines of conformant maps in matrix format Each record consists of the zone number at each line printer character location for a given row on the output map. SYMAP can later read these files as input, allowing for the creation of inexpensive conformant maps by omitting processing required for the initial definition of zone boundaries

In addition to line printer output maps, SYMAP can also produce binary files in a matrix format The resulting files may be used as input to the Laboratory's SYMVU and ASPEX programs These files have one record for each row of symbols on the line printer output map Each record consists of the interpolated values at each line printer location (column)

Application Notes

The most common application of SYMAP involves the display of selected data from the U S Census of Population and Housing Typical subjects mapped include median family income, population density, population change, and housing quality More complex applications include use of the program for analytic, as well as descriptive purposes For example, one may wish to investigate the nature of the relationship between a number of variables using statistical methods The more highly related variables could then be mapped in an attempt to identify spatial trends or document a spatial process

SYMAP has been used in a wide variety of applications including

1. *Market Research* — To delineate patterns of ethnicity and socio-economic character in determining the most cost-effective approach for product testing and potential market penetration
2. *City Planning* — To highlight urban social and economic problems
3. *Coronary Care* — To identify inadequate hospital coronary care facilities and to determine if hospitals with coronary

Program Name	GRID, Version 3.0
Computer Language	FORTRAN IV
Computer Requirements	
Machine	IBM 370
Memory	150K bytes
Mode	Batch processing
Peripheral Devices	Standard line printer

Description

GRID is a special purpose program designed to display data values which have been collected on the basis of gridded data collection zones Since it was designed specifically for use with gridded data bases, the program is able to display such data more efficiently than would be possible with a general purpose program like SYMAP The program is frequently used to map natural resource data derived from aerial photos using a grid overlay to define data collection zones

Like SYMAP, various options may be used to control the scale, symbolism, and value range Other options allow the user to define grid cell size and shape, prescale the data, obtain grid numbering or perform dot mapping Users prepare their own data input subroutine (FLEXIN) for selecting specific subject items from a data bank FLEXIN may also be used to define a mathematical expression involving two or more variables The resulting values are then displayed by use of shading symbolism The result is a graphic matrix composed of one or more print symbols for each cell of the data grid The density of a print symbol indicates the data value or category assigned to a given cell

Although GRID normally is used with data based upon rectangular grids, there is also a method for specifying irregular outlines In the program as distributed the data grid is restricted to 10,000 cells but minor modifications to a DIMENSION statement can greatly increase the program's internal storage capability Larger data matrices also may be processed as a series of adjacent panels which the program can produce automatically

Application Notes

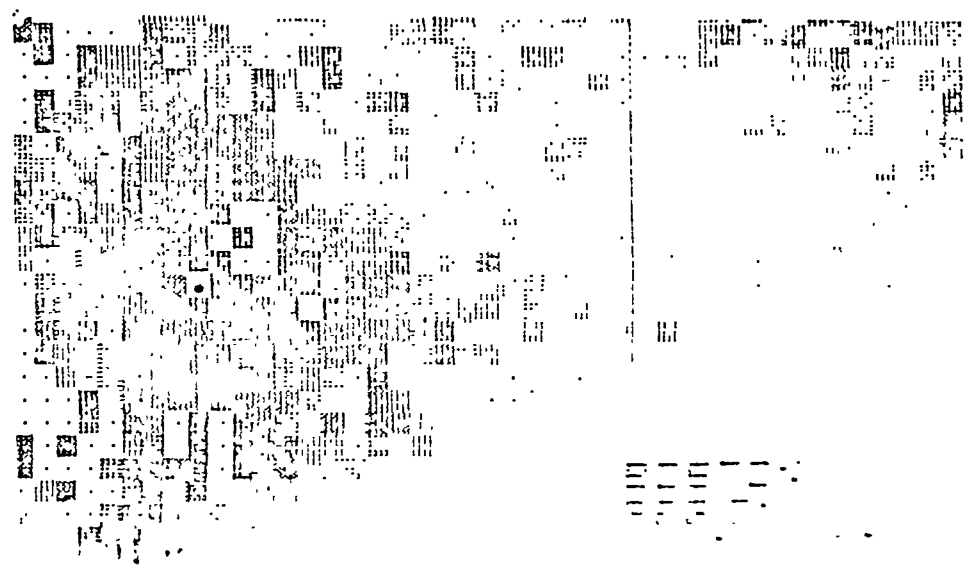
The GRID program is commonly used as [illegible] necessary to record data for a region without reference to pre-existing data zones or other boundaries Examples of such data include land use, soil type, ground water, vegetation type, zoning constraints and topography Since features of this type rarely exist neatly within predefined geographic zones, such as census tracts, it may be necessary to establish a geographic data collection unit such as a grid over a study area and then record one or more data values for each grid cell Data obtained by remote sensing techniques such as LANDSAT satellite imagery is recorded directly in a gridded (raster) format

In addition to being able to display data from a single matrix, it is also possible to read data from two or more matrices using subroutine FLEXIN and perform user defined arithmetic or logical operations upon the values within each grid cell The resulting values may then be displayed as a new matrix of values This technique is particularly helpful when evaluating a variety of alternative combinations or weighting [illegible] data matrices for several different sub[illegible]. Such an operation has been used to identify those locations (grid cells) having a maximum (or minimum) attractiveness for future development, conservation or other use GRID also allows for the evaluation of alternative sites by different analysts Each of their professional preferences may be expressed in terms of selected subjects and their weightings The resulting maps may then be interpreted as alternative solutions which reflect the judgment or preferences of each analyst in response to a given set of objectives

Materials Available

1 GRID Version 3 Computer Mapping Program
2 GRID User's Reference Manual

Copies of the FORTRAN IV source program for an IBM 370 computer are available on new, unlabeled 200 foot reels of 9 track tape written at 800 BPI



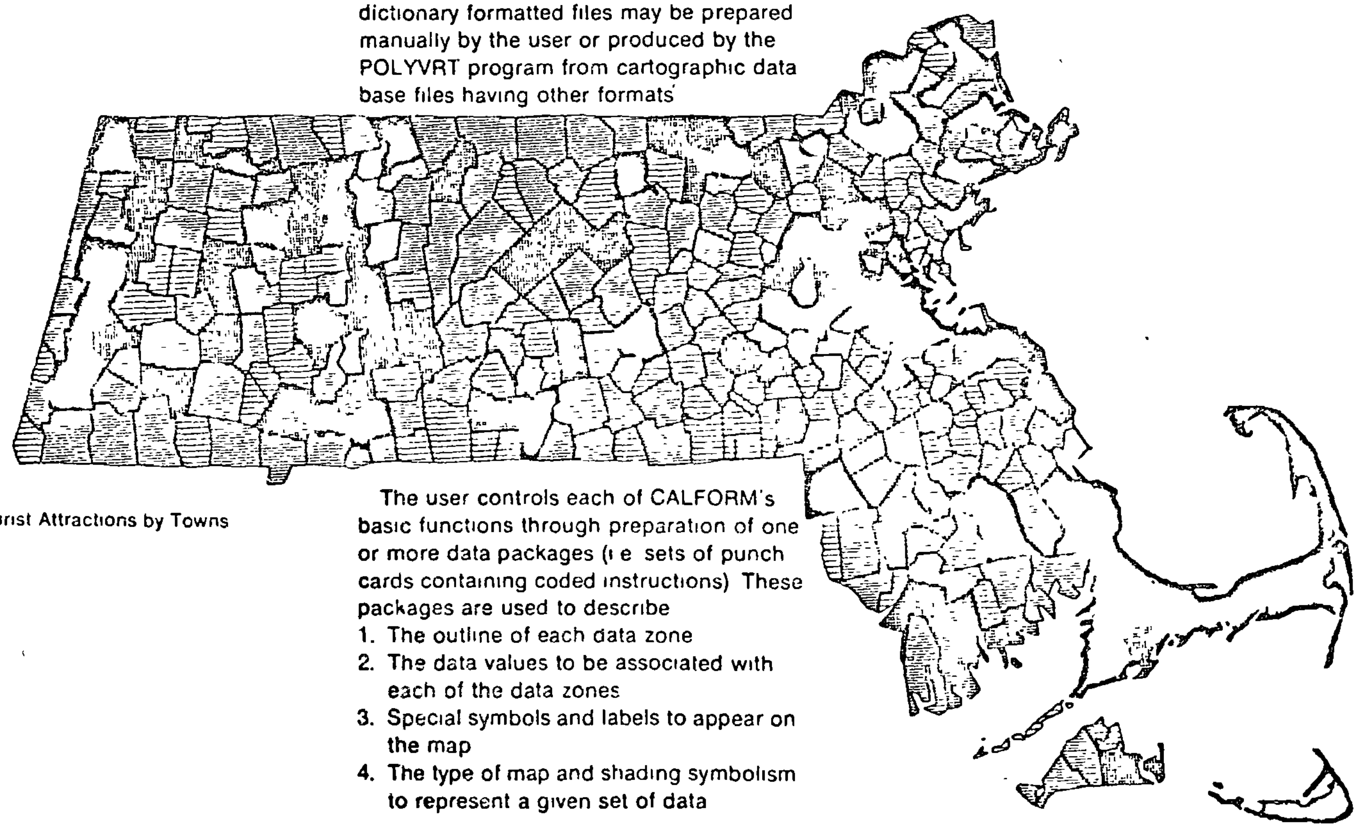
Program Name . . . **CALFORM, Version 1.2**

Computer Language . . . . . . .FORTRAN IV
Computer Requirements
Machine . . . . IBM 370
Memory . . . . .150K bytes
Mode . . . . .Batch processing
Peripheral Devices
Three temporary data sets
Pen plotter or CRT plotter

Description

CALFORM (CALcomp FORMs) is a computer program which uses a line plotter to produce a shaded map based upon data zones The program is typically used to display data by administrative zones which have been defined as irregularly-shaped polygons. The resulting conformant maps graphically depict data values for data zones such as census tracts, municipalities and counties This program may be used to produce graphics of high quality, suitable for publication.

CALFORM produces two basic types of conformant maps outline maps and shaded thematic maps A map may contain as many as 500 data zones, each of which can be described as a polygon composed of up to 600 line segments. Required input data is in a point dictionary format which ensures that each point is uniquely defined thereby eliminating graphic "sliver errors" point dictionary formatted files may be prepared manually by the user or produced by the POLYVRT program from cartographic data base files having other formats



Mass Tourist Attractions by Towns

The user controls each of CALFORM's basic functions through preparation of one or more data packages (i e sets of punch cards containing coded instructions) These packages are used to describe

1. The outline of each data zone
2. The data values to be associated with each of the data zones
3. Special symbols and labels to appear on the map
4. The type of map and shading symbolism to represent a given set of data

Once the user has defined an initial map size and type, data zone outlines, the number and range of data categories and the shading symbolism associated with these categories, numerous maps can be produced by providing a new set of data values and/or output specifications for each additional map

Application Notes

CALFORM is used to produce conformant maps of a study area which is divided into a number of data zones Each zone may have associated with it a value to be mapped By grouping the values into a number of value class intervals and selecting a line or character symbol for each interval, one or more maps may then be produced depicting the spatial pattern of the data

This program is designed to be used on a device having high resolution such as a line plotter It has a number of features to allow legends, keys and titles to be included in the map Hence, it can be used for maps of publication quality

Through the use of the FLEXIN option, the data values for each zone can be manipulated prior to their display For example, FLEXIN may be used to consolidate many small zones to a number of larger zones The values for all newly aggregated zones could be derived by weighting the contribution of each constituent zone as a function of its contribution to the new zone aggregate value

For identifying zones which are shaded according to their data value, a smaller inset outline map of each zone with its name can be included and positioned in an appropriate part of the larger map

Although the program's algorithm for creating data value intervals is designed for data measured on an interval or ratio scale, data which has an inherently ordinal or even nominal measurement scale may be mapped as well For example, symbols could be chosen to represent several different types of agricultural production The counties of a state could then be mapped by the predominant type of agriculture in each county.

Other applications of CALFORM involve mapping of data values related to point or line locations rather than areal zones By coding the points as compact areas (e g , squares) of small size, or lines as linear strips of limited width, shading may be employed to illustrate values pertaining to the points or lines

Materials Available

1. CALFORM Mapping Program
2. CALFORM User's Reference Manual

Copies of the FORTRAN-IV source program for an IBM 370 computer are available on new, unlabeled, 200 foot reels of 9 track tape written at 800 BPI. A sample set of input data is furnished as a second file

Program Name SYMVU, Version 1.2

Computer Language . . FORTRAN IV
Computer Requirements
Machine . . . . IBM 370
Memory . . . . 220K bytes
Mode . . . . . Batch processing
Peripheral Devices
Four temporary data sets including plot tape
Pen plotter

Description

The **SYMVU** computer mapping program uses a line plotter to represent gridded spatial data as a three-dimensional surface Surface locations hidden from view are not drawn SYMVU representations have been used to portray such variables as topography, income, population, air pollution and bathorhythms Users have a great deal of flexibility in specifying how the surface is to be drawn including rotation, tilt, elevation, vertical scale, base and background symbols as well as locational symbols, titles and text

SYMVU accepts a matrix of data values containing up to 130 × 130 elements The data may be generated in either of two ways If the data values are irregularly spaced, SYMAP can interpolate between the data points to produce a contour map and at the same time generate a matrix formatted tape file specifically designed to be processed by SYMVU If users furnish their own input subroutine to SYMVU (Subroutine DATA) SYMVU may be used to read a user provided matrix of data values in a non-standard input format or manipulate values according to user specifications SYMVU contains several user aids such as the automatic computation of minimum and maximum data values It also has a data smoothing routine to reduce minor fluctuations and diminish extreme variations within the data

Application Notes

Although SYMVU can represent any data provided in a matrix format (up to its dimensional capacity of 130 × 130 data units), it is most effective for displaying data which can be represented as a continuous surface Such surfaces are commonly represented on a standard two-dimensional map as a series of contours When viewed as a SYMVU perspective drawing, a three dimensional surface yields substantially greater detail and visual impact

In addition to displaying values which represent a continuous surface, SYMVU may also be used to depict a matrix of values representing a discontinuous data surface Such surfaces correspond to conformant map data for data zones such as census tracts, towns, states or countries In such cases the data zones appear as raised plateaus whose elevation corresponds to the data value of each zone The outline of each plateau retains the general shape of the data zone SYMAP may be used to generate a matrix formatted tape file of a conformant map for input to SYMVU

SYMVU surfaces have been effectively used to present data relating to such subjects as landforms (terrain elevation), population densities, air pollution concentrations, income distribution and related social statistics Other applications of SYMVU have included

1 *Animated Time-Series Films* — Spatially distributed data showing changes in population growth or air quality
2. *Stereograms* — Images achieved by plotting two views of the same surface and changing the viewing azimuth by a few degrees
3. *Cartograms* — Representations of geographic regions depicted as volumetric prisms Height is proportional to one variable (e g , population size) and the area of each data zone is defined proportional to a second variable (e g , per capita income) The volume becomes proportional to the product of the two variables (total income, in this case)
4 *Two-Color Overlays* — Combination data display in which one surface, such as topographic relief, is plotted with reference to a related data grid indicating forested areas With slight program modification, the lines over the forested portion of the grid are drawn as a separate plot The two plots are then registered and printed together in two colors using standard printing techniques

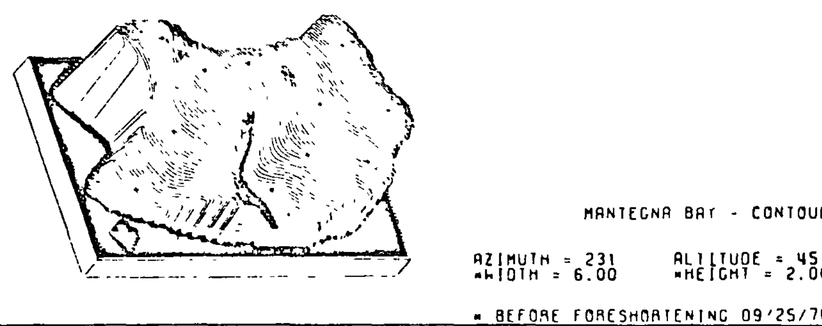
Materials Available

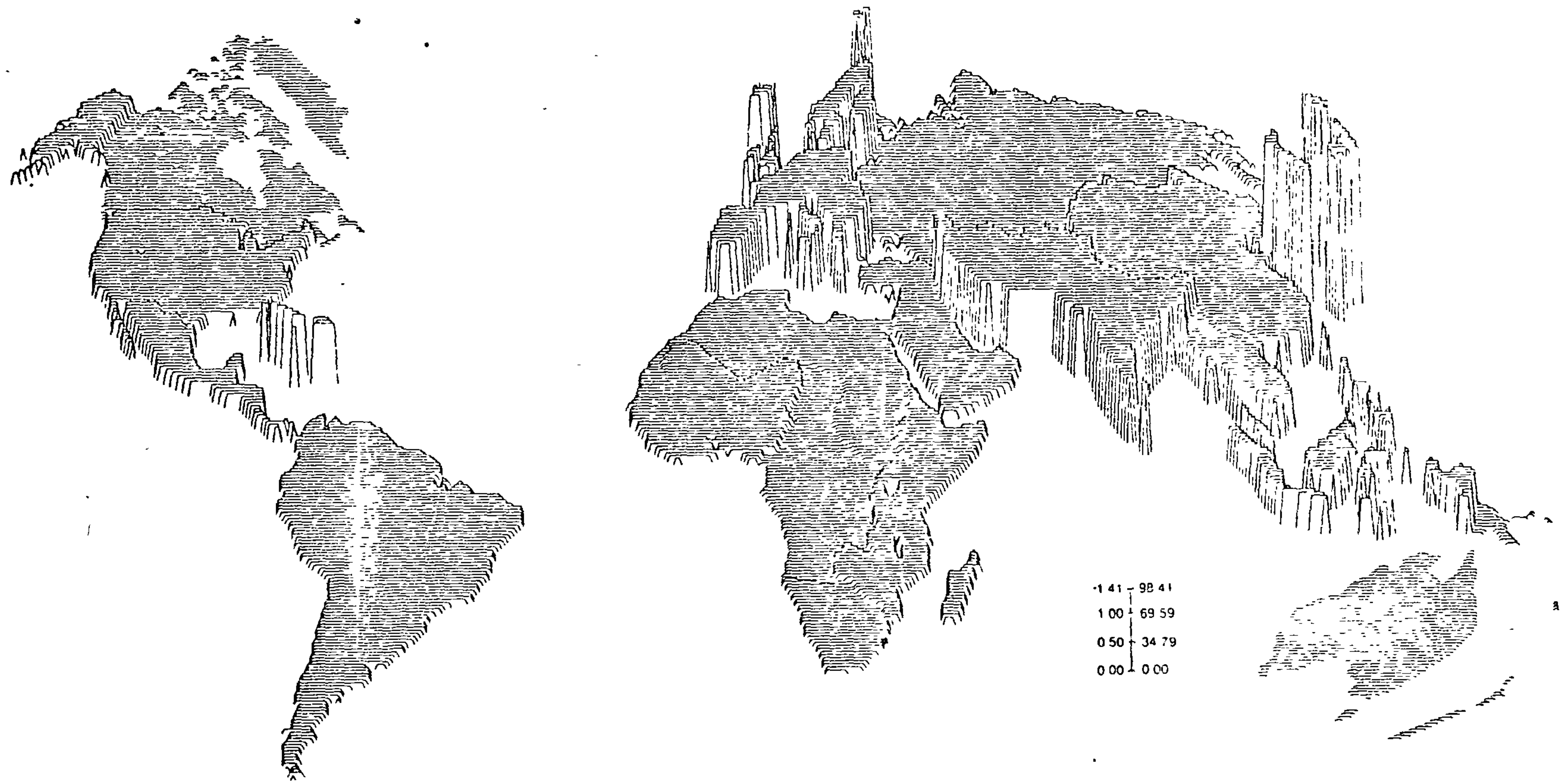
1 SYMVU Program with sample data
2 SYMVU User's Reference Manual

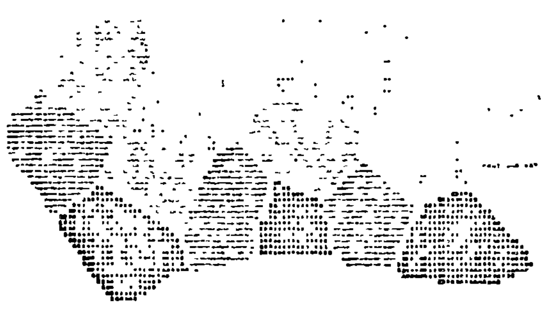
Copies of the FORTRAN-IV source program for an IBM 370 computer are available on new, unlabeled, 200 foot reels of 9 track tape written at 800 BPI

← SYMAP Contour Map

SYMVU Contour Map →

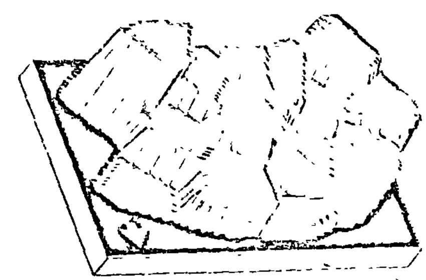






← SYMAP Conformant Map

SYMVU Conformant Map →



Program Name ... **.POLYVRT, Version 1.1**

Computer Language . . FORTRAN IV
Computer Requirements
Machine ....IBM 370
Memory . ... .160K bytes
Mode ....Batch
Peripheral Devices
Two temporary data sets
Pen plotter

Description

POLYVRT is a general purpose utility program designed for use with polygon oriented cartographic data bases (CDB's) Coordinate data describing polygons may be input and output in a variety of data structures and formats Such files are normally required for use with conformant mapping programs

POLYVRT's capabilities include

1 *Conversion of a CDB File* — Translation from one data structure or format to another
2 *Internal Topological Data Structure* — Automatic detection of errors in polygon definition
3. *Retrieval of Selected Polygons* — Retrieval from a CDB file using FORTRAN-like logical expressions
4. *Update and Correction of Coordinates* — Revision of coordinates and topological attributes of chains and polygons
5 *Map Projections* — Creation of specific projections, rotations and scalings
6. *Assignment of Levels of Line Detail* — Representation based on an automated line generalization technique.
7. *Simple Line Plotting* — Visual verification of a file.

POLYVRT relies on its topological data structure — the chain file — to provide the capabilities outlined above A "chain" consists of a series of points which form a boundary between the polygons on its left and right The first and last points of a chain are referred to as nodes Together these three elements (nodes, points, and left/right polygon identifiers) define the network of a cartographic data base The chain file concept has a generality and efficiency that offers many advantages for the creation and maintenance of cartographic data bases

Any of the following file structures may be accepted as input and converted to an internal chain file data structure In addition, the following file structures may be output from the internal chain file of POLYVRT

1 *DIME (Dual Independent Map Encoding)* Files — CDB's where each line segment is identified by its nodes, (end points) and the polygon names on each side of the line DIME files which describe the census blocks and tracts for most U S metropolitan areas plus a file of county outlines are available from the Census Bureau A county outline file for the U S in DIME format is provided with the POLYVRT program
2 *World Data Bank I* — A file containing 120,000 points which describe the outlines of the countries of the world The Laboratory has enhanced the basic WDB-I file available from the U S National Technical Information Service by adding the topological information required by a POLYVRT chain file This modified WDB-I chain file is provided along with the POLYVRT program
3. *CALFORM File* — A cartographic data base structure based on a common point dictionary with which polygons are described The CALFORM program for conformant mapping is distributed separately by the Laboratory
4 *SYMAP File* — A very simple file in which each polygon is individually defined as a string of coordinates is available as output from POLYVRT At a future date POLYVRT will be modified to also accept a SYMAP file as input
5 *LUDA (Land Use Data and Analysis)* — Files generated by the U S Geological Survey These files share the 'chain' concept with POLYVRT although they use the term 'arc ' Conversion of these files will be a new feature of POLYVRT once formats for dissemination of LUDA files have been established
6 Other file structures will be handled in response to user requests and included in future updates of the program

POLYVRT provides capabilities to perform a variety of operations upon a cartographic data base by use of optional input packages A specific subset of an input CDB may be requested through English-like IF statements Topological error detection may be used to identify and also to correct certain file errors. Correction of errors and additions to a file also may be effected by use of the program's comprehensive update package Chains, parts of chains and their topological characteristics may be added, deleted or replaced A coordinate transformation package including ten standard map projections (with provision for a user-supplied routine to define non-standard projection types) is also provided

The program includes an efficient recursive algorithm that measures the detail level (geometric significance) of each point in each chain By specifying a set of criterion bands, the user may create a line detail measure for each point and this information can be stored with the file Files may be output with any degree of line generalization making it possible to produce the smallest CDB compatible with the purpose for which it will be used Issues relevant to the selection of an appropriate level of line generalization include map size and scale plus the line drawing precision of output devices such as printers, CRT's and line plotters

Output may be displayed on a line plotter A chain structured output file can be produced for input to INPOM and CELLMAP The topological chain will be the basic data structure for much of the Laboratory's future mapping software

The distribution tape includes source code, the county DIME file, and the Laboratory revision of WDB-1 *The POLYVRT Users Manual* includes program documentation along with control card examples and instructions IBM versions include a separate file containing the Job Control Language to compile, overlay, and execute the program from the tape provided

Application Notes

Before using a conformant mapping program (such as SYMAP, CALFORM, INPOM, CELLMAP), a user must create a cartographic data base file which describes the data collection zones Generally, this description consists of x-y coordinate outlines of polygons, but there are significant differences in the record and file format requirements for various mapping programs Creating a new CDB involves the expenditure of significant amounts of time and energy.

One goal of POLYVRT is to provide a universal, archival description of data zones based upon a topological data structure POLYVRT can then use this data structure to generate any one of the record and file formats required by a specific mapping program As a result, POLYVRT provides a great deal of flexibility to the individual who wishes to create only one CDB which can be used to produce input to a variety of different mapping programs

A major obstacle to the widespread use of computer mapping software is the initial cost of preparing a CDB for a given study area This process can be avoided in many applications by using existing, publicly available files However, the conflicting formats and structures of these files reflect the lack of standards in the design of computer mapping systems Without POLYVRT a user is required to write a different program to extract useful information from a variety of sources The difficulties often outweigh the benefits

The growing library of available CDB's describe geographic units from the city block level in Metropolitan GBF-DIME files to the country level in WDB-I With POLYVRT, a user can generate a CDB specifically tailored from any one of these sources As a result, the effort required to produce a CDB and related thematic maps of an area can be significantly reduced

When an available file does not provide the information required, POLYVRT can aid in the process of constructing a CDB The chain file format offers the following advantages in the digitizing process

- Allows use of stream input modes
- Avoids ambiguity and sliver errors
- Reduces the number of keystrokes required to identify polygons

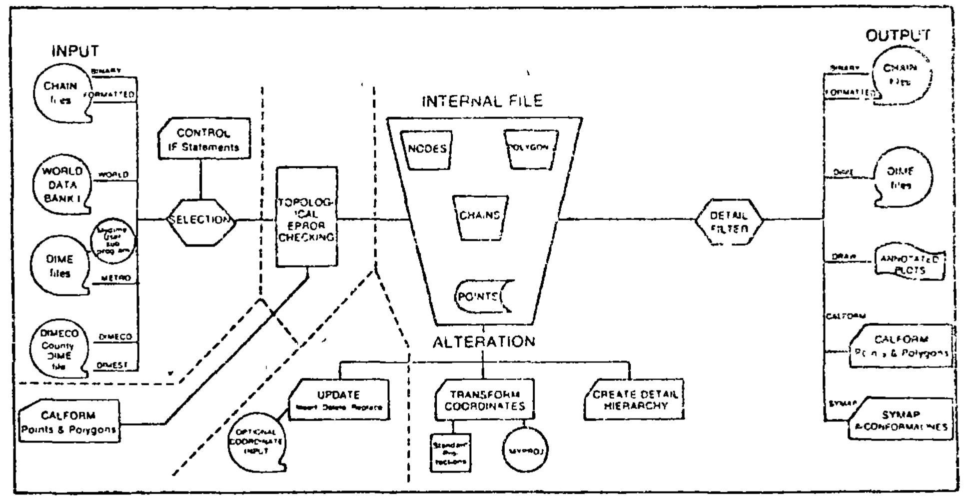
Furthermore, POLYVRT's topological error checking capability serves an important function in detecting errors The program's updating facility allows corrections to be made without redigitizing. By creating a CDB through POLYVRT, the quality of the information is likely to be quite high and the user acquires a great deal of flexibility in the subsequent uses of the file

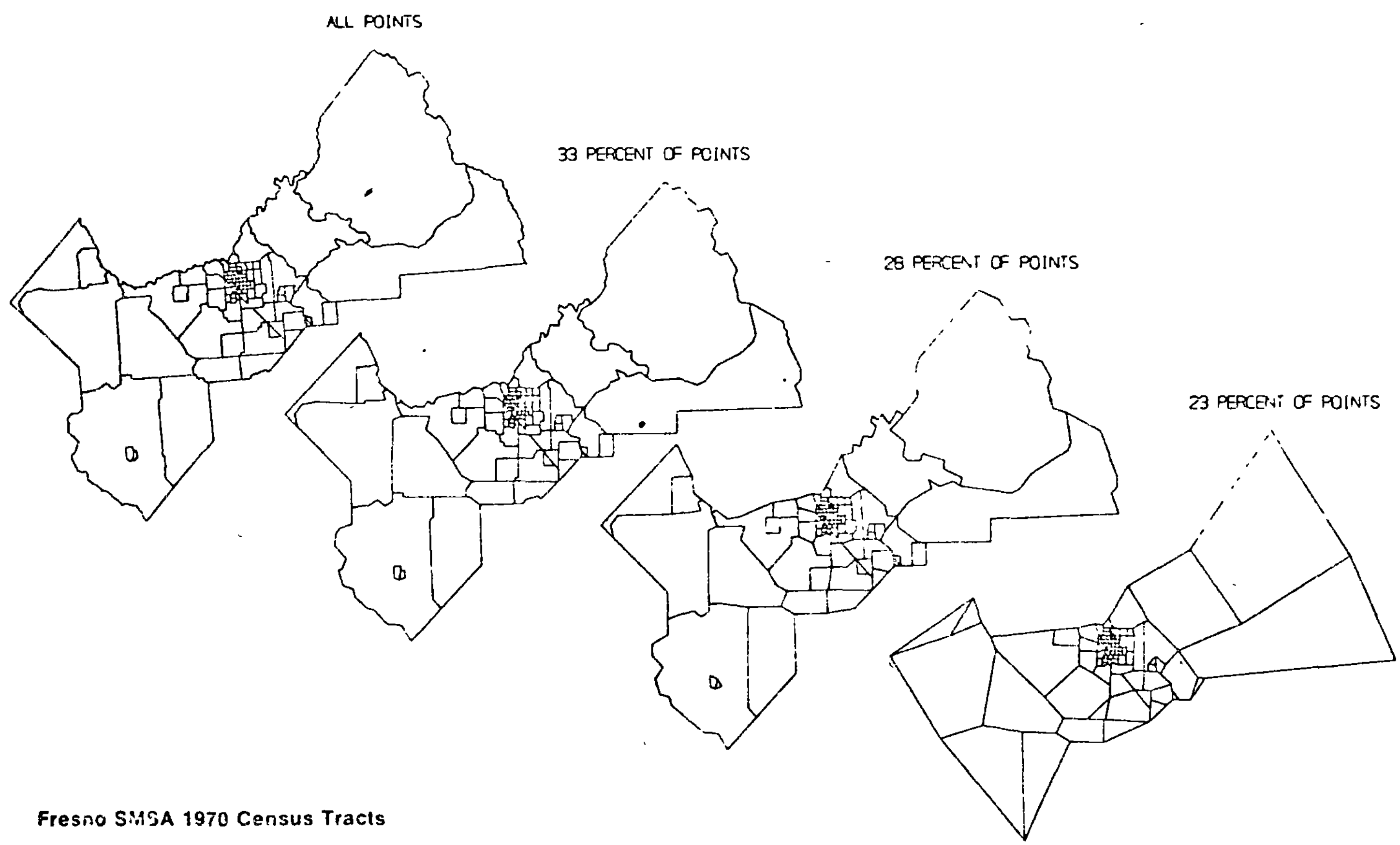
Materials Available

1. POLYVRT program with sample data
2. POLYVRT User's Reference Manual
3. U S County DIME File (as created by and available from the U S Census Bureau)
4. World Data Bank-I File (as created by the U S Government and modified by the Laboratory)

Copies of the FORTRAN-IV source program for an IBM 370 computer are available on new, unlabeled, 200 foot reels of 9-track tape written at 800 BPI

POLYVRT SYSTEM COMPONENTS





Fresno SMSA 1970 Census Tracts

Program Name . . INPOM, Version 1.0

Computer Language . . FORTRAN IV

Computer Requirements

- Machine . . . PDP-10, IBM 370
- Memory 20Kwords, 150K bytes
- Mode . Interactive or Batch

Peripheral Devices

- Seven data sets plus plot tape for offline plots (Optional)
- Tektronix 4000 series cathode ray tube
- Pen plotter

★ NOTE ★
INPOM will be available in 1977

Description

INPOM (INteractive POlygon Mapping) is a computer mapping program which produces conformant outline or shaded maps on a cathode-ray tube A study area is partitioned into data zones such as continents, countries, states, counties, census tracts, city blocks or any user-defined areal units The user provides a cartographic data base (CDB) which contains x-y coordinates describing the data zones of interest

Data values are assigned to each data zone by the user Many variables can be mapped in succession using the same CDB for a given study area INPOM may be used to extract data values from a file or they may be entered from the user's terminal and saved as a file for later use

Ten categories of graphic symbolism may be used to graphically represent the values Graphic symbols are constructed using parallel (hatched) lines, intersecting (cross-hatched) lines, or regularly spaced character symbols Angles and densities of lines, plus selection and spacing of characters are defined within the program However, they may be redefined by the user for any particular map

A user interacts with INPOM by use of an English-like command language which consists of verbs and direct objects The language may be entered in a free field format using terms which are completely spelled-out or by using any unambiguous abbreviations The user also may specify parameters affecting the size, scale, and geographical limits of a map to be produced It is also possible to enter into a dialogue with the program when a more complicated series of specifications are required, such as definitions of shading, value ranges, and titles Map output is sent directly to the user's CRT terminal in response to a 'MAP' command or to a pen plotter following a 'PLOT' command

INPOM's cartographic data base defines a study area as lists of x-y coordinates called "chains" Each chain serves the topological functions of connecting two end points (nodes) and bounding two adjacent data zones Chains usually contain other points between their two end points The actual number of points in a chain is a function of the length and complexity of the boundary represented by the chain Any cartographic data base composed of data zones can be converted by the POLYVRT program to a chain file format for input to INPOM POLYVRT can accept as input DIME files, CALFORM files or chain files (such as the World Data Bank I file supplied with POLYVRT) and produce a chain file for subsequent input to INPOM CDB coordinate data may also be entered directly into the program from the keyboard of a graphic terminal, or recorded using a digitizing tablet or other graphic input device

Once structured as an INPOM chain file, the cartographic data base must be stored on disk Using such a working file allows the program to be smaller and achieves flexibility in defining and displaying a map image Random access reading of the disk file also allows for user selection of sub-regions, legends and titles from a large data base plus a high degree of efficiency in plotting of maps

Other features of INPOM include

1 *Line Generalization* — Selectively reduces the number of points used to describe data zone outlines When displaying a map a user may select any one of 10 detail levels which have been defined for each point on the chain file at the same time it was created by POLYVRT, or INPOM will create (or re-define) detail levels according to the user's specifications

2 *Polygon Selection and Windowing* — Enables a subregion of a CDB to be isolated and mapped The area of interest may be specified by naming the polygons to be included or by providing a set of coordinates which essentially describe a window overlayed on the base map

3 *Titling Capability* — Allows any number of titles, each having up to 5 lines of text to be defined and positioned within a user-defined space on the output map Titles can be stored as files and any combination of titles from a given file can be retrieved for display on a given map

4 *Cartographic Data Base and Data Values Editing Capability* — Allows points on a chain to be modified or new chains and polygons to be created Data values can be input from the terminal or from a disk file and individually modified by the user Edited data values and CDB's can be saved offline Terminals equipped for graphic input (including tablets) may be used for recording chain formatted coordinates directly from a base map

5 *Lines* — Makes use of chains to define lines as well as polygon boundaries Lines (e g roads, rivers) may be mapped in combination with polygon outlines or displayed independently of polygon outlines

6 *Legends* — Defines and positions strings of characters as legends (e g , place names) on the maps

7 *Graphic Symbol Key* — Relates graphic symbols to each value class interval and may be displayed alone or inserted within a map Each element of the key may be positioned individually, if desired

INPOM is written in machine-independent FORTRAN IV for PDP-10 or IBM-370 computers When used at other installations it may require alteration of certain of its modules which handle random access files or address graphic terminals and plotters Graphic subroutine calls are compatible with Tektronix PLOT-10 software

Application Notes

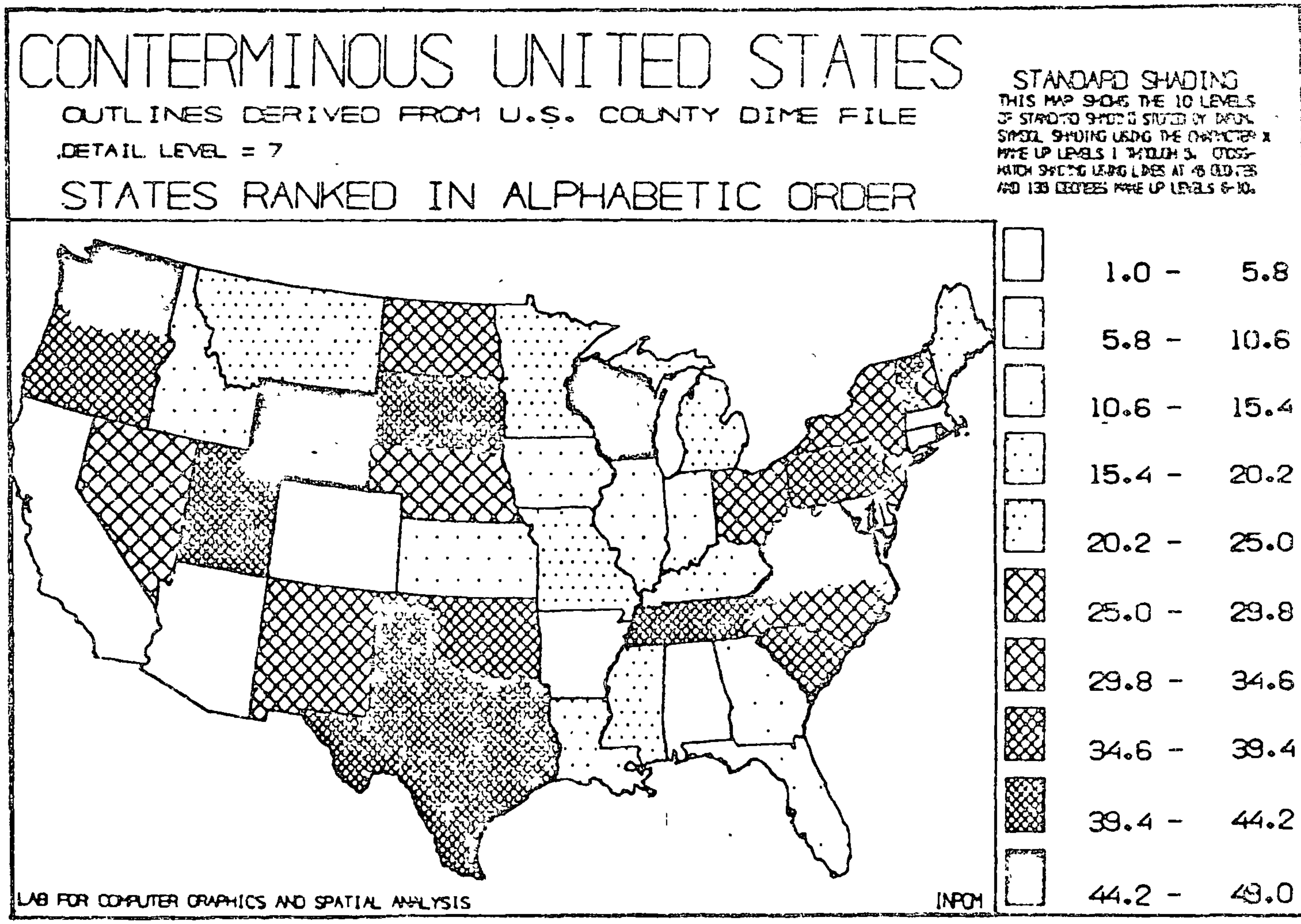
INPOM's ability to display a map on a Cathode Ray Tube or pen plotter provide it with a degree of flexibility unequalled by batch mode graphic programs A user may very quickly examine a large number of alternative maps, varying either in terms of subject matter, study area or graphic symbolism The idea of using computer generated maps as analytic as well as descriptive tools assumes new meaning when it becomes possible to produce maps within minutes rather than h[illegible] or days

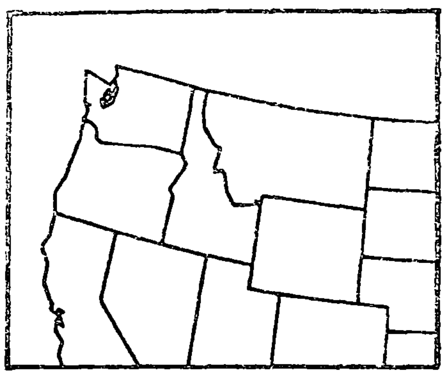
Materials Available

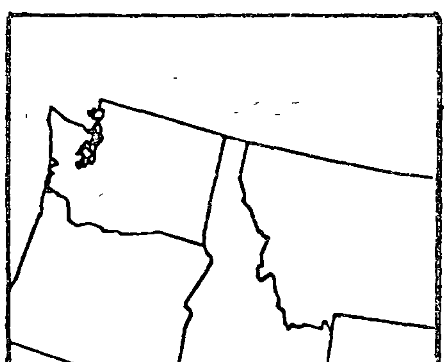
1. INPOM Version 1.0 computer mapping program
2. INPOM User's Reference Manual
3. World Data Bank — I file (as created by the U.S. Government and modified by the Laboratory).

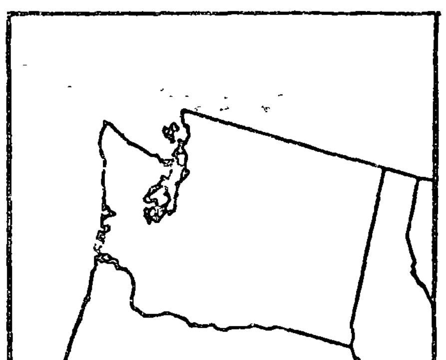
Copies of the FORTRAN-IV source program for an IBM 370 computer are available on new, unlabeled, 200 foot reels of 9 track tape written at 800 BPI.

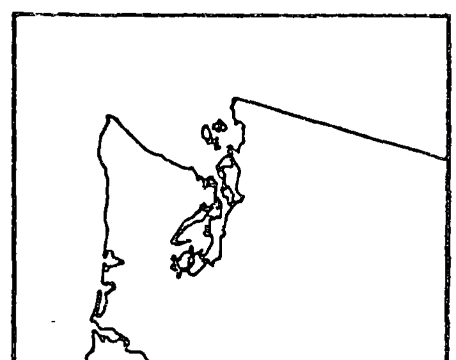
Copies of the FORTRAN-IV source program for a DEC PDP-10 computer are available on a DEC tape in compressed format with DEC specific command files and MACRO-10 files included.

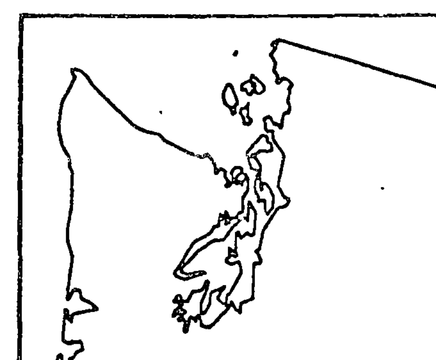












Program Name	ASPEX, Version 1.0
Computer Language	FORTRAN IV
Computer Requirements	
Machine	PDP-10, IBM 370
Memory	20Kwords, 150K bytes
Mode	Interactive or batch
Peripheral Devices	
Three data sets	
Tektronix 4000 series cathode ray tube	
Pen plotter	

Description

ASPEX (Automated Surface Perspectives) is a computer program for producing oblique perspective views of three dimensional surfaces on a cathode ray tube (CRT) The program is an outgrowth of the SYMVU program described elsewhere in this publication ASPEX, however, is a significantly more powerful tool for displaying surfaces because it operates interactively and is able to accept much larger files of input data while requiring less core memory

Input data for ASPEX consists of a matrix of values identical to that used by SMYVU ASPEX also uses the same method of relief representation, i e elevation lines drawn along rows, columns, or diagonals including an algorithm for deleting hidden line segments

ASPEX can be run either in an interactive or batch mode Interactive mode accepts input of control commands from a CRT keyboard with output to the CRT or to a plotter Batch mode involves control information being read from cards with output being produced on a line plotter

APEX's user language involves commands composed of verbs and objects A verb describes an action to be taken, and an object identifies the entity or value to be acted upon An object may be a single word, or a word set equal to a value Furthermore, a value may be a number, a set of numbers, a string of characters, or a symbol for which a value has been previously defined In interactive mode, 'HELP', 'SHOW', and 'LIST' commands are available to assist the user in selecting desired verbs and objects ASPEX currently can accept a data matrix containing up to 500x500 cells (i e , any number of rows and columns whose product is less than 250,000) Input data is processed by ASPEX and output onto a random access disk file with only a small portion of the data in core memory at any time As a result the program is able to run in a smaller amount of core memory than SYMVU while displaying larger sets of data

ASPEX allows the user to define the azimuth, altitude and viewing angle relative to the data surface The program also has an alternative mode for defining a view of the surface A user may specify two points in three-dimensional cartesian space, the eye-point of the observer and the center point of the area being viewed These points may be defined to be at any position in three-space, including positions on the surface

A flexible method for scaling and selecting a "window" of the matrix is also provided In addition, effects similar to a variable focal length (zoom) lens can be achieved by specifying 1) a "cone of vision" which determines how much of the matrix around the center of vision will be included, and 2) the "focal ratio" which determines where the picture plane (the projected image of the surface) is placed between the two viewing points (eyepoint and centerpoint)

Other features include

1 A *Subroutine FLEXIN* — To manipulate the data values of the input matrix prior to their display
2 An *Option to Generalize Relief Lines* — To reduce the number of points thereby allowing reduction of the time and cost required to produce a plot
3 A *Variety of Graphic Cosmetics* — To create endlines at the matrix edges, blocks of varying sizes at the matrix base, graphic scales for data interpretation and special background shading for regions within the square matrix but outside the study area
4 *The Ability to Produce Three Different Surface Projections* — To achieve isometric, planometric, and perspective variations

Application Notes

ASPEX is able to effectively display data surfaces of all kinds This includes data derived from digital terrain models as well as contour, conformant, proximal and trend surface maps produced by SYMAP. The resulting graphics provide a great deal of visual impact The user has a broad range of control over how a data surface is to be viewed For example, one may move across the surface by producing a series of graphic displays Such a technique would be particularly useful for purposes of film animation where each successive plot could be used to produce several seconds of viewing time on film

Individuals involved in the preparation of data surfaces frequently express a desire to select a viewing angle which provides maximum visibility of the surface This concern reflects the fact that some portion of a surface is usually hidden by "peaks" which rise up and block from view areas behind them The interactive capabilities of ASPEX minimize this problem by allowing the user to freely experiment with various viewing angles, distances, windows, and vertical scalings

A surface drawing program such as ASPEX allows thematic map data to be displayed without the necessity for data value classification inherent in traditional thematic map symbolism Since each value on the surface is shown raised to a height which is continuously proportional to the value scale, a user is free of the necessity to establish value class intervals and then select corresponding symbolism A three dimensional surface is literally classless thereby allowing minor as well as major fluctuations within the data to be shown on the same drawing There is a need to establish a vertical scale for a data surface but associated limitations are far less significant than those which result from having to classify and symbolize data on a traditional two dimensional shaded map

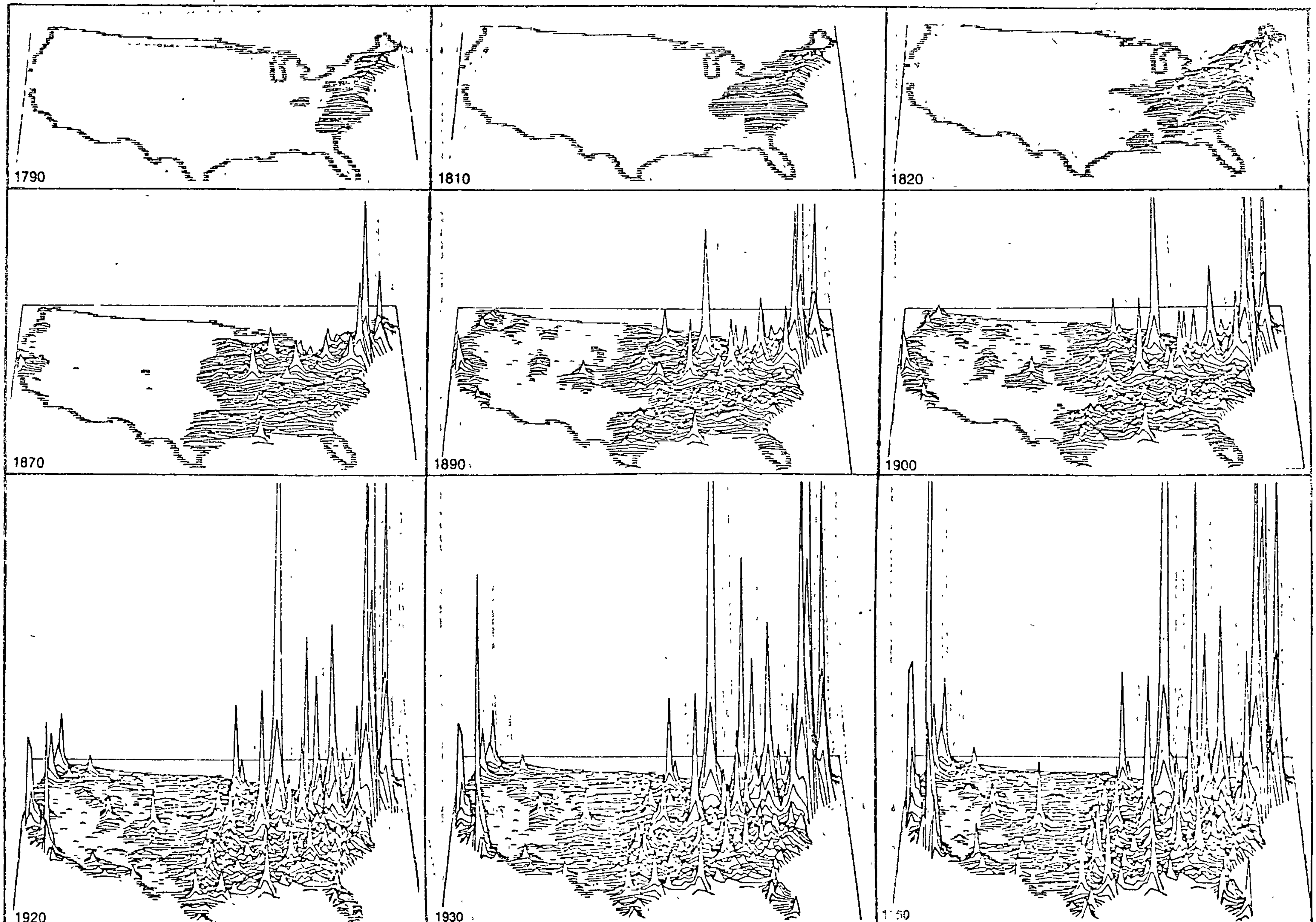
Materials Available

1 ASPEX program with sample data
2 ASPEX User's Reference Manual

Copies of the FORTRAN IV source program for an IBM 370 computer are available on new, unlabeled 200 foot reels of 9 track tape written at 800 BPI

For installation of the program on a DEC PDP-10 computer a user alternatively may request a DEC tape in compressed format with DEC specific command files and MACRO-10 files included

Selected Maps of U.S. Population Density from "Manifested Destiny"



Introduction

A cartographic data base (CDB) is a necessary part of the input data required by the Laboratory's conformant mapping programs These data bases define the x-y coordinate locations for each data zone which is to appear on a computer generated map A CDB may be digitized locally but it is usually more economical to obtain a copy of an existing CDB if one already has been prepared for the area of interest

CDBs Available

The Laboratory has available a selected set of CDBs and will consider requests to prepare others to a user's specifications CDBs distributed by the Laboratory are available in many different record and file formats, including those required by the SYMAP, CALFORM and INPOM computer mapping programs Alternatively, an individual may prefer to acquire a CDB in a topological chain format and use the Laboratory's POLYVRT software to create a variety of CDBs for use with one or more mapping programs

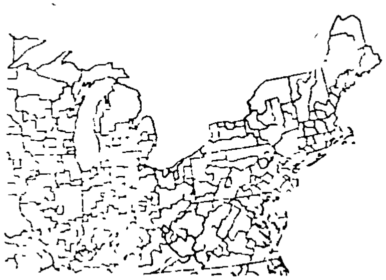
Original Data Sources

CDBs offered by the Laboratory frequently have been derived from other files such as the CIA's World Data Bank-1 and the Census Bureau's County DIME, Metropolitan DIME, and Urban Atlas files In most instances the content and/or structure of these files have been modified by the Laboratory Silver error, overlapping polygons and similar difficulties inherent in the original data have been corrected by the Laboratory

Centroid (Data Point) Files

For users of SYMAP or other programs which utilize x-y coordinate data for point locations, files of centroids (central points representative of given areal units) are available Files currently being distributed include

1. Centroids of U S counties
2. Centroids of U S States
3. Centroids of Massachusetts cities and towns



Standard Economic Areas (SEAS)

Standard Metropolitan Statistical Areas (SMSAS)

Countries of the World

Study areas and their geographic subdivision (data zones) for which CDBs are available include the following.

Study Area	Geographic Units	Standard Format	Original Source	Coordinate System	Recommended Maximum Mapping Scale	Maximum Ground Resolution per 0.1"	Maximum Ground Resolution per 1mm on map
Planet Earth	Countries	Chain	WDB-I	Radians, Lat & Long	1:10,000,000	16 mi.	10 km
United States	States	"	County DIME File	"	1:5,000,000	8.0 mi.	5 km
United States	SMSAs	"	"	"	"	"	"
United States	SEAs	"	"	"	"	"	"
United States	Counties	"	"	"	"	"	"
Any State	Counties	"	"	"	"	"	"
Any SMSA	Census tracts, 1970	"	Urban Atlas Files	Degrees, Lat. & Long	1:15,000	125 ft.	15 m
Any SMSA	Census blocks, 1970	Metro-DIME	Metro-DIME	Radians, Lat. & Long	1:5,000	40 ft.	5 m
Massachusetts	Cities & Towns	Chain	Files locally prepared	Inches, Cartesian	1:250,000	2,100 ft.	250 m

Line Generalization (Detail)

The amount of line detail (number of points) contained within a file may be specified at the time that an order is placed. Alternatively, the Laboratory will provide a degree of line detail appropriate to

1. The mapping package with which the CDB is to be used,
2. The maximum size at which all or some part of the file is to be displayed, and
3. The line resolution of the device on which the output is to appear

All of the Laboratory's CDB files contain a finite amount of detail. However, when preparing a CDB for mapping purposes, it first is desirable to generalize each line in order to retain only those points which can be seen on the final graphic output. The actual number of useful (visible) points depends upon the scale of the final map output plus the graphic resolution of the display hardware. Clearly, a line printer map can portray less detail than a cathode ray tube which in turn has less revolving power than most pen plotters for a map at a given scale. Other issues to be considered when selecting a detail level include the maximum number of points which can be processed by a given display program plus the relative cost of processing large files

The standard version of each CDB whose standard format is "chain," includes with each x,y coordinate pair a third number which is the measure of that point's deviation from the trend line. As a result, each point may be either included or excluded from a given plot depending upon the amount of line resolution desired in a given situation. This capability provides a great deal of flexibility to the user and allows for a variety of different CDB's, varying in line detail, to be produced from the original file. For a description of the algorithm which determines deviations and a discussion of their use in "detail filtering" see Douglas, David and Thomas Peucker, "Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or its Caricature" in the *Canadian Cartographer*, Vol 10, No 2, Dec. 73, pp 112-122

Record Formats

CDB's distributed by the Laboratory are available in any one of four standard formats or in a format defined by the user. Standard formats include files prepared specifically for input to the SYMAP, CALFORM or INPOM computer mapping programs. Files also may be obtained in a topological chain file format for input to the POLYVRT program and other mapping software to be distributed by the Laboratory

Record Content

CDB's prepared for input to a computer mapping program include x-y coordinates plus a geographic identification code for each polygon. Codes used are Federal Information Processing Standards (FIPS) or Census Bureau geographic identifiers. INPOM and POLYVRT files also include a detail level code for each point coordinate in the file. For CDB's other than those described which are prepared to a user's specifications, geographic codes appropriate to the locations of interest can also be provided

POLYVRT Software

All CDB's offered by the Laboratory are directly compatible with the computer mapping software for which they are intended. However, for those individuals wishing to acquire a high degree of flexibility in working with CDB's for one or more study areas, it may be advisable to request a CDB in a POLYVRT chain file format plus a copy of the POLYVRT software. By doing so one acquires the capabilities to

1. Generate CDB's in a variety of formats (e g , CALFORM, SYMAP, INPOM, etc)
2. Select only those geographic regions of immediate interest from a larger file
3. Specify any one of several map projections
4. Specify any level of line generalization desired.
5. Edit or modify the content of a CDB file in order to satisfy the users' specific requirement.

Map Projections

WDB-I, the County DIME, METRO-DIME and the Urban Atlas files contain coordinates expressed in terms of latitude and longitude. These coordinates usually are projected to a cartesian coordinate system before being used with a computer mapping package. Files describing a region no larger than a Standard Metropolitan Statistical Area (SMSA) are normally created using a standard equi-rectangular projection. For larger regions within the U S (e g one or more states) a user may wish to request a different projection. CDB files for regions as large as the entire United States will normally require projection. Projections available include Albers, Lambert conformal, Mercator, Miller, Equi-Rectangular, Sinusoidal and Orthographic. Other types can be furnished upon special request. CDB's prepared from the DIME and World Data Bank files will be provided with conic (usually Albers) or Sinusoidal projections, unless specific projection instructions are provided. Such instructions should include necessary projection parameters

Documentation

Each CDB includes documentation describing the following

1. Number of polygons and number of points
2. Number of points per polygon
3. Number of chains (for chain files)
4. Identification codes for each geographic unit (FIPS code and name)
5. Minimum and maximum x-y coordinate values for the CDB
6. Map projection parameters (when known)
7. Test plot of polygon outlines
8. Record and file descriptions

File Characteristics

Files will normally be furnished on new, unlabeled 9-track tape, written in EBCDIC mode at 800 BPI, using a blocksize of 4000 bytes or less. If this is not suitable, specify the character set, density, parity and blocking factor appropriate for your installation

Cost

The cost of a CDB is a function of the effort involved in preparing it to a user's specifications. A CDB which can be prepared from WDB-I or the County DIME files is likely to be less costly per point than are similar files derived from Urban Atlas or Metro DIME files. The reason for this is that the latter files frequently include errors which must be identified and corrected prior to their conversion to a CDB file. However, if the file for a region has already been corrected this will be reflected in its cost. Because of the many different variables and user specifications which must be taken into account, those interested in acquiring a specific CDB should request a price quotation

When requesting a CDB price quotation, the following information must be provided

1. Geographic region of interest
2. Geographic data zones desired (type, number and names as appropriate)
3. Computer mapping software package to be used. If other than a Laboratory package, specify the data format required, restrictions concerning maximum number of points or polygons and the device on which graphic output is to be displayed
4. Map projection desired (if any)
5. Line generalization required in terms of
   - Smallest scale output map
   - Positional accuracy
   - Precision of output display device
6. File characteristics (if other than as described above)

Price estimates range from a minimum of $75 for a copy of an existing file to $100+ for a file derived from WDB-I, the Urban Atlas, County DIME or METRO-DIME files

Delivery Time

CDB's derived from WDB I or the County DIME file can normally be shipped within 14 days from receipt of an order. Urban Atlas and Metro-DIME files may require somewhat longer due to additional processing involved. However all orders are normally shipped within 2-4 weeks

PUBLICATIONS

RED BOOK

This is an illustrated synopsis of projects undertaken since the Lab's organization in 1965 The "Red Book" describes the Laboratory's applications of computer graphics and spatial analysis to such fields as architecture, city planning, landscape architecture and theoretical geography Descriptions of how other research communities have applied the Lab's programs are also included "Red Book" is in 8½ x 11 loose-leaf format, chronologically organized for easy reference and convenient updating Research entries for successive years are separated by dated dividers

THEORETICAL CARTOGRAPHY PAPERS

"Issues in Thematic Map Design" Series

In 1976 the Harvard Laboratory for Computer Graphics and Spatial Analysis began publication of a series of discussion papers dealing with new concepts and problems in thematic mapping Consideration is given to maps made by hand as well as by computer Based on extensive Ford Foundation supported research begun in 1969 by Howard T Fisher and various associates, these papers are of three principal types

1 Papers dealing with fundamental cartographic principles without reference to specific mapping situations
2 Papers illustrating and discussing a variety of alternative solutions to representative mapping problems, and
3 Papers analyzing in detail alternative approaches to thematic cartography as presented by leading cartographic texts in English, French and German

The papers of the second type which assume no previous knowledge of cartography, include numerous specially prepared maps designed to show the relative advantages of different design approaches The first of these "illustrative studies presents more than 50 maps showing some 14 different basic types of symbolism as applied to the problem under consideration

As with past Harvard Papers, these are of an informal and exploratory nature with a view of eliciting reader comments It is believed that this series offers important new insights and information capable of contributing to both the theoretical and practical advance of thematic cartography To r up-dated information regarding specific papers as issued, interested persons should request placement on our mailing list Arrangements can be made for the papers in this series to be sent automatically as they become available Those requesting this service will be sent without charge a periodic abstract of such comment as may be received regarding the papers

Below is a tentative list of titles, pursuant to the first two categories mentioned

Fundamental Cartographic Principles

1. Thematic Maps What They Are and Who Needs Them?

2. Wanted An Improved Theoretical Construct to Aid in the Design of Thematic Maps

3. Hypotheses for the Mapping of Qualitative and Quantative Information

4. Thematic Map Titles — What Should They Contain?

5. Types of Study Spaces in Thematic Mapping — Distinctions with a Difference

6. Types of Locations in Thematic Mapping — Distinctions with a Difference

7. Types of Values in Thematic Mapping — Distinctions with a Difference

8. Value Keys in Thematic Mapping — What Should They Contain?

9. A Manual of Value Symbolism — Spot Type

10. A Manual of Value Symbolism — Band Type

11. A Manual of Value Symbolism — Field Type

12 A Manual of Value Symbolism — Cyclical Type

13. Multi-Subject Mapping — With Interlocking Subjects

14 Multi-Subject Mapping — With Diverse Subjects

15. The Use of Color as a Quantitative Analogue in Thematic Mapping

16. A Glossary of Terms Used in Thematic Mapping

Alternative Solutions to Representative Problems

1. Foursquare I — An Introductory Problem in Single-Layer Thematic Mapping

2. Foursquare II — Illustrating Basic Symbolisms for the Representation of Values

3 Foursquare III — Value Curves and Classing Procedures

4. Sparsely Populated France — An Illustrative Problem with Alternative Solutions

5. Densely Populated France — An Illustrative Problem with Alternative Solutions

6. All France I — An Illustrative Problem with Alternative Solutions

7. All France II — An Illustrative Problem with Alternative Solutions

8. The Washington-Boston Corridor — An Illustrative Problem with Alternative Solutions

9. U.S Standard Metropolitan Statistical Areas (SMSAs) With Over One Million Inhabitants — An Illustrative Problem with Alternative Solutions

10. The White House Area, Washington, D C — An Illustrative Problem with Alternative Solutions

11. Twelvesquare I — An Introductory Problem in Multi-Layer Thematic Mapping

12. Twelvesquare II — The Thematic Mapping of Non-Layered Space

AUTOMATED CARTOGRAPHY PAPERS

NOTE The first 12 papers report on a Laboratory study that demonstrates the role of computer graphics in presenting data related to the source, distribution, and effect of air pollution

1. SUMMARY REPORT 150 pp , $9 00 There are 3 sections 1) a summary of the entire project, 2) a discussion of each individual study area, 3) a technical report and a discussion of computer graphic outputs

2. SUMMARY GRAPHICS 162 pp , $9 75 This is a complete set of graphics for all individual case studies

3. COMPUTER EQUIPMENT AND PROGRAMS 83 pp , $5 00 This report discusses computing machinery (hardware), computer programs (software), plus types of maps and data for graphic display Also, the introduction considers computer mapping criteria, operating principles, and current and future applications

SYMAP INTERPOLATION CHARACTERISTICS 70 pp , $4 20 There are four sections 1) the SYMAP interpolation algorithm, 2) sensitivity and accuracy of interpolation, 3) use of electives to affect interpolation and 4) use of barriers to affect interpolation A comparison of SYMAP's algorithm with other methods of interpolation is also included

CASE STUDY REPORTS These are detailed descriptions of individual case study plus technical reports relating to other research undertaken Relevant graphics are included with each report

5. ST LOUIS REGION CASE STUDY 3 vol , 326 pp , $19 75 This includes comparisons between air quality and socio-economic data, examination of diurnal variations in air quality, modeling of air quality from power plant emissions and statistical examination of the data

6. KANSAS CITY REGION CASE STUDY 112 pp , $6 75 Air quality data is manipulated and mapped with particular emphasis on background values at the periphery of the study area These include defining coherent areas for mapping, time-averaging of data and comparison with air quality standards Also, emission densities are mapped and air quality is compared with socio-economic data

7. MONTREAL REGION CASE STUDY 59 pp , $3 75 An Index of Atmospheric Purity (I A P) is determined and mapped in this study The coverage and extent of various lichens and mosses are statistically examined and mapped to achieve the best interpretation for the limited amount of data gathering In addition, the I A P was compared to measured air quality

8. CONNECTICUT STUDY 44 pp , $3 00 Air quality mapped and compared at various time periods Comparisons are then made using mathematical models

9. SOUTHERN NEW ENGLAND 55 pp , $3 50 Measured air quality supplemented with estimated background values is compared to socio-economic surfaces

10 PUGET SOUND REGION 32 pp , $2 00 Air pollution emissions from future transportation and urban development options are mapped Transportation related air qualities reflect the influence of mountain barriers to pollutant dispersion

11. CALIFORNIA REGION 49 pp , $3 00 Investigations were made of carbon monoxide concentrations, the total annual suspended dust concentrations and air quality relative to statewide air quality standards

12. BOSTON REGION 51 pp , $3 25 This report illustrates differences in statistical surfaces based on daily, monthly, seasonal and 'y averages

13 The Use of Computer Graphics in Planning, by Howard Fisher, presented at the 1970 National Conference of the American Society of Planning Officials April 6, 1970 3 pp , $1 00
Those involved in urban planning are not likely to be able to come to sound conclusions and be able to make sound recommendations unless the facts regarding the human community under study are readily available in easily understood form '

14. Maps ala Computers, reprint of a March 11, 1970 article from the Christian Science Monitor, $1 00
"It s hard to know where to begin solving such amorphous challenges as crime and air pollution in cities A group at Harvard with the help of a computer is mapping concentrations of these problems These computer maps lend a clarifying visual perspective "

15 Computer Cartography, by Thomas K Peucker Commission on College Geography Resource Paper No 17, Association of American Geographers 1972 75 pp $3 00
A valuable tutorial covering such topics as theory of data surface processing and representation of surfaces The publication also contains numerous illustrations involving over twenty computer mapping programs

16 Thematic Cartography, by Philip Muehrcke Commission on College Geography Resource Paper No 19 Association of American Geographers 1972 66 pp $3 00
This report describes the various steps involved in thematic mapping data collection, mapping and analysis Specific subjects discussed include spatial sampling measurement, symbolic representation, visual analysis and quantitative map analysis

17 Management by Computer Graphics, by Kenneth Shostack and Charles Eddy, Harvard Business Review, November December 1972, 12 pp $1 00
"The tremendous output of the high-speed computer has far outstripped the modern executive s ability to examine absorb, and use all the information generated in his day to-day decision making But this imbalance between the machine and the manager is now being corrected through the development of computer graphics "

18 Color in Art by Howard T Fisher and James M Carpenter Fogg Art Museum Harvard University 1974 134 pp $11 00
An excellent introduction to color and color theory presented from the special viewpoint of those concerned with the use of color in art and design work based upon a study originally undertaken with specific reference to the use of color in thematic map design Includes among its many black and white and color illustrations numerous charts and diagrams of which four of particular reference value are in color

19 Manifested Destiny A Graphic Account of the Settlement and Growth of America 1790-1970, by Geoffrey Dutton, Laboratory for Computer Graphics and Spatial Analysis Graduate School of Design, Harvard University 1975 30 pp , $2 00
Surfaces of population distribution for the conterminous United States are displayed (using the ASPEX program) for each census date from 1790 through 1970 Total population by county were used as input data allocated to a grid for display The spreading of settlement and the growth of cities are dramatically illustrated by the changing population surface

20 Programming for Transportability A Guide to Machine Independent FORTRAN, by Nicholas Chrisman and Denis White Laboratory for Computer Graphics and Spatial Analysis Graduate School of Design Harvard University, March 1976 40 pp $2 50
This paper addresses the problem of creating a "transportable" subset of FORTRAN Four aspects of the problem are dealt with

1 Syntactic limitations
2 Semantic problems
3 Necessary evils
4 Suggestions about good practice

This publication is based on experience with IBM CDC PDP, UNIVAC Honeywell and Burrough s operating systems

21 Computer Cartography World-Wide Technology and Markets, by Eric Teicholz and Julius Dorfman, International Technology Marketing, Newton Massachusetts May 1976 427 pp , $375 00
The study covers all aspects of automated cartography including the automated cartography process users and sources of automated cartography in the United States, Canada Europe Latin America and Japan the status and forecast of automated cartography technology and a market survey and forecast

THEORETICAL GEOGRAPHY PAPERS

1 Concepts and Applications—Spatial Order, by William Warntz and Michael Woldenberg 16 May 1967, 196 pp $12 00

2 A Nomographic Representation of the Geoid, by Walter Messcher, 28 August 1967, 13 pp , $1 75

3 Implicit Map Projections in Computer Print-Outs, by William Warntz, 15 September 1967, 24 pp $2 25

4 Out of Print

5. Superceded by paper No 8.

6 Superceded by paper No 23.

7 The Geometry of Mixed Hexagonal Hierarchies in the Context of Central Place Theory, by C Ernesto S Lindgren 22 December 1967 44 pp $3 00

8 Energy Flow and Spatial Order, with Special References to Mixed Hexagonal Central Place Hierarchies, by Michael J Woldenberg 3 January 1968 46 pp $3 25

9 The Continent Problem — Geography and Spatial Variance, by Christopher W Warntz, 32 pp , $2 25

10 Distances and Land Values as Data for Introducing Problems Associated with Spatially Continuous Fields of Correlation Coefficients, by William Warntz 29 January 1968, 18 pp , $1 50

11 Space Straightening and Flattening, by C Ernesto S Lindgren, 2 February 1968, 6 pp , $1 50

12 Out of Print

13 Spatial Order in Fluvial Systems Horton's Laws Derived from Mixed Hexagonal Hierarchies of Drainage Basin Areas, by Michael J Woldenberg 14 February 1968, 37 pp , $3 80

14 A Nomographic Representation of Trajectories, by Walter Messcher 8 March 1968, 11 pp $1 50

15 A Two-Dimensional Interpolation Function for Computer Mapping of Irregularly Spaced Data, by Donald Shepard 20 March 1968, 20 pp $1 50

16 N-Dimensional Spatial Analysis and Computer Graphics Part A, by C Ernesto S Lindgren, 14 April 1968, 60 pp , $4 20

17 Hyper-surfaces and Geodesic Lines in 4-D Euclidean Space, by C Ernesto S Lindgren 17 May 1968, 15 pp , $1 50

18 A Note on Stream Ordering and Contour Mapping, by William Warntz 1 July 1968 30 pp , $2 00

19 Hierarchical Systems Cities, Rivers, Alpine Glaciers, Bovine Livers and Trees, by Michael J Woldenberg, 8 July 1968 160 pp $10 00

20. A Study of Rivers and Other Branching Systems, by Gordon Howie 31 July 1968 19 pp , $1 50

21. Geography and an Existence Theorem A Cartographic Computer Solution to the Localization on a Sphere of Sets of Equal-Valued Antipodal Points for Two Continuous Distributions with Practical Applications to the Real Earth, by Stephen E Selkowitz, 5 August 1968, 64 pp $4 50

22 A Mathematical Representation of the Altitude Relationships on the Surface of the Earth, Using Spherical Functions to the 16th Order, by Albert Prey, translated by Bonnie Binder August 1968 62 pp , $4 50

23 Plane Globe Projection — A Linnean System of Map Projection, Translation of Min Rat Prof Dr Hans Maurer s Ebene Kugelbilder forwarded and edited by William Warntz, 19 August 1968, 337 pp $18 00

24 The Descriptive Geometry (or Representative) of a Collection of Points Fixed by N Coordinate Numbers or of N-Dimensions, by Felippe dos Santos Reis translated by C Ernesto S Lindgren 30 September 1968 31 pp $2 25

25 The Law of Travel and Its Application to Rail Traffic, by Eduard Lill translated and forwarded by Thomas K Peucker 10 January 1969, 123 pp $7 75

26 Minimum Time Paths and the Migration of the Arctic Tern, by Robert C Eckhardt 7 February 1969, 26 pp $2 00

27 Notes on the Methodology for Generation of the Representative of a Set, by C Ernesto S Lindgren, 14 February 1969 39 pp , $2 75

28 A Minimum Path Problem Reconsidered by C Ernesto S Lindgren, 21 February 1969, 11 pp, $1 50

29 The Use of the Geodesic Curvature in the Determination of Geodesic Lines, by C Ernesto S Lindgren, 28 February 1969, 13 pp , $1 50

30 Numerical-Geometrical Techniques for Information Storage and Retrieval, by C Ernesto S Lindgren, 1 October 1969 8 pp , $1 50

31 Some Reflections on Concepts Based on Three-Dimensional Geometry, by C Ernesto S Lindgren 8 October 1969, 13 pp , $1 50

32 Algorithms and Models Based on Projective Transformations in Spatial Location, Regional Planning, and Central Place Theory, by C Ernesto S Lindgren, 17 October 1969 14 pp $1 50

33 Graphical Representation of a Matrix with Applications in Spatial Location, by C Ernesto S Lindgren and Carl Steinitz, 30 October 1969 53 pp $3 75

34 Some Thought on Optimal Mapping and Coding of Surfaces, by Thomas K Peucker Ph D , 3 November 1969 14 pp $1 50

35 Homological Transformations in Four-Dimensional Space, by C Ernesto S Lindgren 25 November 1969 20 pp , $1 50

36 A Study of the Movement of a Point on a Plane and in Space, by C Ernesto S Lindgren, 7 December 1969 16 pp , $1 50

37 Set of Equal-Value Antipodal Points for Two Continuous Distributions, by C Ernesto S Lindgren 12 December 1969 11 pp , $1 50

38 An Outline for the Theory of Man-Made Space — Essays in Urbanology, Number One, by Kozmas Balkus, 26 February 1970 55 pp $4 00

39 A Structural Taxonomy of Spatial Hierarchies, by Michael J Woldenberg, 24 March 1970, 53 pp , $4 00

40 The Determination of Fixed-Points in Finite Dimensional Spaces, by C Ernesto S Lindgren 3 June 1970, 17 pp , $1 50

41. Law and Order in the Human Lung, by Michael J Woldenberg, Gordon Cumming, Keith Harding, Keith Horsfield, Keith Prowse and Shiam Singhal, 29 July 1970 58 pp , $4 00

42. The Hexagon as a Spatial Average, by Michael J Woldenberg, 15 October 1970, 26 pp $2 00

43 The Two Dimensional Spatial Organization of Clear Creek and Old Man Creek, Iowa, by Michael J Woldenberg, 23 June 1971, 34 pp , $2 50

44 The Sandwich Theorem — A Basic One for Geography, by William Warntz, C Ernesto S Lindgren, Katharine Kiernan, Louisa Bonfiglioli, Eduardo Lozano, 30 June 1971, 88 pp , $5 25

45 Relations Between Horton's Laws and Hydraulic Geometry as Applied to Tidal Networks, by Michael J Woldenberg, 10 June 1972, 40 pp , $3 00

46 The Two-Dimensional Spatial Analysis of the Pecatonica River in Wisconsin, by Michael J Woldenberg and Larry Onesti, 1 July 1972, 57 pp , $4 00

47. A Computer Program for Mixed Hexagonal Hierarchies, by Rachel Thurston, Michael Woldenberg and David Barer, 15 June 1975, 214 pp , $13 00

48 Fred K Schaefer and the Science of Geography, by William Bunge, Detroit Geographical Expedition, I, Detroit, Michigan, 1 November 1968 26 pp , $1 50

49 Sparking Potential, Personal Interaction and Social Distance Directions for a Theory, by Geoffrey Dutton, Graduate Student, Department of City and Regional Planning, Harvard University Cambridge Massachusetts 1 August 1969, 11 pp $1 50

50 Macroscopic Aspects of Metropolitan Evolution, by Geoffrey H Dutton, 11 March 1970, 116 pp $7 25

51 Notes on the Friction of Distance in the U S Telephone Network, 1935-1995, by Geoffrey H Dutton, November 1971, 11 pp , $1 50

52 Tabulations of Data on Area, Population, Income and Certain Derived Quantities for the 3070 Counties of the 48 Conterminous States of the United States, 1967, by Geoffrey Dutton Katharine Kiernan Douglas Kingsbury and William Warntz 19 May 1971, 324 pp , $19 25

53 The Geographical Distribution of Income in the Conterminous United States, 1967-68, and the Income Fronts by States, by William Warntz, 24 May 1971 20 pp , $1 50

54 A Description of the 1967-68 United States Income Potential Surface, by Douglas Kingsbury, 20 June 1971 13 pp $1 50

55 National and Regional Parameters of Growth and Distribution of Urban Population in the United States, 1790-1970, by Geoffrey H Dutton 9 July 1971, 30 pp $2 25

56 Allometric Growth in Social Systems, by Michael J Woldenberg, 15 pp , $1 50

57 Allometry in Micro-Environment Morphology, by Ranko Bon, Graduate Student, Department of City and Regional Planning, Harvard University, June 1972 32 pp $2 50

OTHER PAPERS

1 The Potential of Video Tape Recorders for the Design Professions, Eric Teicholz, Laboratory for Computer Graphics and Spatial Analysis, Graduate School of Design, Harvard University, July 1972, 18 pp , $1 50

Video tape recorder technology is one of the most rapidly developing communication technologies This paper looks at aspects of new technology, talks about the differences between film and video tape, discusses the relationship between VTR and computer technology and suggests some potential applications to the design professions.

2 CPED Computer Programs in Environmental Design, Kaiman Lee, Center for Environmental Research, Boston, Massachusetts 1974, 1300 pp , $210 00

CPED is a 5 volume set of reference manuals containing documented and illustrated computer program abstracts A reliable design and planning data source for architects, researchers, and others concerned with computer technology Contents include feasibility studies, architectural programming, site planning, relational planning, two and three dimensional graphics, cost control, environmental control, circulation analysis, text manipulation, project control, office management and evaluation

3 Computer Aided Space Planning by Eric Teicholz, a paper presented at the First Bi-National (U S /Australian) Urban Systems Symposium September 1974 23 pp $1 50

This paper discusses general approaches to using a computer for space planning aid in an architectural, urban and regional context Programs are classified and various techniques for space planning are described and illustrated

4 Interactive Mapping of Urban Data by Eric Teicholz, a reprint of a paper that appeared in the Proceedings of the Second General Assembly of the *World Future Society*, June 1975, 12 pp , $1 50

The paper describes several factors contributing to the explosion of statistical and geographical data related to urban areas It also deals with corresponding interest in automated procedures for the input, analysis and display of spatial data INPOM and ASPEX, the Laboratory's two and three-dimensional interactive mapping programs, are briefly described

5 Interactive Graphics Comes of Age by Eric Teicholz, a reprint of the *Datamation* article appearing in December 1975, 4 pp , $1 50

The paper is a brief history of interactive computer graphics that shows technical trends in industries and universities There is a comparison of the turnkey integrated interactive graphics systems along with market figures and application areas for this newly emerging field

6 Computer Graphics A Perspective by Eric Teicholz, a reprint of a paper appearing in *Biosciences Communications*, S Karger AG, Basel, Switzerland, January 1976, 17 pp , $1 50

The paper surveys basic ideas and approaches to computer graphics in general and to interactive computer graphics in particular It is an historical overview including speculations on the future of interactive graphics with reports on various hardware systems and their programming implications

7 How to Talk to a Computer Applications Consultant by Eric Teicholz, a paper presented at the National Computer Conference New York City, June 1976, 8 pp , $1 00 A discussion of the first (and last?) ten years of computer aided architectural design in relation to evolving technologies and architectural practices

8. Graphic Technology and the Display of Spatial Data by Eric Teicholz, a paper presented at the Wescon Conference, Los Angeles, Calif , November 1976, 10 pp $1 00 Describes the unique characteristics of spatial data as a subset of graphics Hardware and software considerations related to the display of spatial are also discussed

EDUCATIONAL AIDS

Self-Study

Self-study instructional materials are available concerning use of the SYMAP program They are designed for students having neither knowledge of geography nor experience with statistical and quantitative mapping The materials include five lessons to acquaint the user with the various capabilities of SYMAP Each lesson successively requires greater use of various input packages and mapping electives plus sample problems to be completed by the reader

Video Tapes

Through a grant from the Alfred P Sloan Foundation, the Laboratory distributes half-inch BAJ-standard video tapes of the SYMAP, SYMVU, and GRID programs which introduce the prospective user to specific computer mapping programs Each tape presents the various output devices used to produce computer maps, and shows the step-by step preparation of a computer map using the program Because of their general nature, the tapes are intended only as supplements to the user reference manuals which are available for each program

Order Form

Computer Programs and Related Materials

Program Distribution — General Information

ssemination of Computer Programs
An organization which requests a program from the Laboratory must agree in writing that the program only will be used for its own internal operations and will not be made available to others without prior written consent of the Laboratory A University computing installation and all its users are defined as a single user, a commercial service bureau or time sharing company offering use of the Laboratory's programs is *not* considered to be a single user A commercial vendor such as a service bureau or time sharing company wishing to provide access to the Laboratory's programs for its customers should contact the Laboratory for leasing arrangements.

Postage and Handling
Computer programs and other items recorded on magnetic tape or punched cards are shipped to locations within the U S by first class mail, postage and handling prepaid Documentation materials normally provided with a program are included as part of the shipment All other publications are sent as third class mail, printed matter

Computer programs and publications sent outside the U S are shipped via surface mail unless other arrangements have been made in advance. Price estimates for air mail shipment of specific items to another country will be sent upon request Alternatively, additional funds to cover the cost of air mail shipment may be enclosed with an order Any funds in excess of the amount required will be refunded

Telephone Orders
Telephone orders will be accepted from organizations within the U S A purchase order number is required Orders may be placed during the hours of 9 30 AM to 4 30 PM by calling area code (617) 495-2526

Payment
Payment should accompany an order unless a purchase order number has been included Payment must accompany all orders of less than $5 00 Orders to be shiped outside the U S must include advance payment in the form of a bank draft payable in U S dollars

Prices
Prices marked with an asterisk (*) apply to educational institutions and governmental agencies only

Computer Program	Quantity	Cost (US$)	*	Amount
SYMAP				
1 Mapping Program incl items 2 through 6 below	____	$1,000 00	$665 00*	____
2 User's Reference Manual	____	10 00		____
3 Self Study Course (5 lessons)	____	25 00		____
4 Interpolation Characteristics Report	____	5 00		____
5 Grid Sheet	____	1 00		____
6 SYMAP Ruler	____	50		____
7 SYMAP Replacement Tape	____	100 00		____
8 SYMAP Video Tape	____	40 00		____
SYMVU				
1 Mapping Program incl User's Reference Manual	____	715 00	$475 00*	____
2 User's Reference Manual	____	5 00		____
3 SYMVU Replacement Tape	____	71 50		____
4 SYMVU Video Tape	____		40 00	____
GRID				
1 Mapping Program incl User's Reference Manual	____	575 00	$380 00*	____
2 User's Reference Manual	____	4 00		____
3 GRID Replacement Tape	____	57 50		____
4 GRID Video Tape	____	40 00		____
CALFORM				
1 Mapping Program incl User's Reference Manual	____	715 00	$475 00*	____
2 User's Reference Manual	____	5 00		____
3 CALFORM Replacement Tape	____	71 50		____
POLYVRT				
1 Cartographic Data Base (CDB) Utility Program incl User's Reference Manual plus CDB's for U S Counties and nations of the world	____	1,100 00	$735 00*	____
2 User's Reference Manual	____	6 00		____
3 POLYVRT Replacement Tape	____	110 00		____
ASPEX				
1 ASPEX Interactive mapping program incl User's Reference Manual	____	$1,200 00	$800 00*	____
2 User's Reference Manual	____	5 00		____
3 ASPEX Replacement Tape	____	120 00		____
INPOM TO BE AVAILABLE IN 1977				
1 INPOM Interactive mapping program incl User's Reference Manual and a CDB for nations of the world	____	$1 200 00	$800 00*	____
2 User's Reference Manual	____	5 00		____
3 INPOM Replacement Tape	____	120 00		____

Computer programs and data bases for use on IBM computers are distributed on unlabeled 9-track magnetic tape recorded in EBCDIC mode (odd parity) at 800 BPI Logical record length is 80 and physical record length is 8 000

If a non-IBM computer has been listed on side two, the program to be sent will be,

(1) a version which has been converted by others for use on the computer specified, or
(2) a standard IBM version plus a list of users with similar hardware who have expressed a willingness to assist others in conversion.

For further information concerning the current availability of programs for specific non-IBM hardware, contact the Laboratory

(over)

CUT ON DOTTED LINE

Order Form for Computer Programs & Related Materials/Side 2

Those requesting a program must complete the following:

CONTACT PERSON to receive future program changes and announcements (sent free of charge).

Name ______________________ Title ______________________

Organization ______________________ Phone ______________________

Address ____

____

HARDWARE INFORMATION

Computer to be used ______________ Manufacturer ______________ Model ______________

TERMS OF AGREEMENT

It is the policy of the Laboratory for Computer Graphics and Spatial Analysis, Graduate School of Design, Harvard University, to make its programming systems available only under the following conditions

1 Neither this software, its documentation, nor adaptations thereof shall, except with prior written consent of the Laboratory, be sold, leased or otherwise distributed in any form to any individual, business entity, academic institution or governmental body whatsoever.

2 Upon acceptance of these terms and conditions, as indicated by signature of an officer having authority to enter into such agreements, the Laboratory grants the recipient a royalty free, non-exclusive license to use the subject material at a single computer facility.

I, an authorized official, understand and accept the stated conditions for use of the material to be provided by the Laboratory.

(Signature) ______________________________ (Date) ______________

(Name, printed) ______________________________

(Title) ______________________________

(Organization) ______________________________

(Address) ______________________________

____

(Telephone) ______________________________

PAYMENT

Checks should be made payable to Harvard University, Lab for Computer Graphics.

Total Amount $ __________ Check Enclosed __________ Purchase Order #: __________ Date: __________

Orders should be sent to. The Lab for Computer Graphics, 520 Gund Hall, Harvard University, Cambridge, Mass 02138.

SHIP TO:

Name ______________________

Organization ______________________

Address ______________________

____

BILL TO:

Name ______________________

Organization ______________________

Address ______________________

____

CUT ON DOTTED LINE

Order Form

Publications

Publication Distribution — General Information

Postage and Handling

Publications are sent as third class mail, printed matter via surface mail unless other arrangements have been made in advance Price estimates for air mail shipment of specific items to another country will be sent on request Alternatively, additional funds to cover the cost of air mail shipment may be enclosed with an order Any funds in excess of the amount required will be refunded

Telephone Orders

Telephone orders will be accepted from organizations within the U S A purchase order number is required Orders may be placed during the hours of 9 30 AM to 4 30 PM by calling area code (617) 495-2526

Payment

Payment should accompany an order unless a purchase order number has been included Payment must accompany all orders of less than $5 00. Orders to be shipped ouside the U S must include advance payment in the form of a bank draft payable in U S dollars.

AUTOMATED CARTOGRAPHY PAPERS	Quantity	Price	Amount
1 Summary Report		$ 9 00	
2 Summary Graphics		$ 9 75	
3. Computer Equipment and Programs		$ 5 00	
4. Symap Interpolation Characteristics		$ 4 20	
5 St Louis Region Case Study		$ 19 75	
6 Kansas City Region Case Study		$ 6.75	
7. Montreal Region Case Study		$ 3.75	
8. Connecticut Case Study		$ 3 00	
9 Southern New England Case Study		$ 3 50	
10 Puget Sound Region Case Study		$ 2 00	
11 California Region Case Study		$ 3 00	
12 Boston Region Case Study		$ 3 25	
13. The Use of Computer Graphics in Planning		$ 1 00	
14 Maps a la Computers		$ 1 00	
15. Computer Cartography		$ 3 00	
16. Thematic Cartography		$ 3 00	
17. Management by Computer Graphics		$ 1 00	
18 Color in Art		$ 11 00	
19 Manifested Destiny A Graphic Account of the Settlement, etc.		$ 2 00	
20 Programming for Transportability, etc.		$ 2 50	
21 Computer Cartography World-Wide etc.		$375 00	

THEORETICAL GEOGRAPHY PAPERS	Quantity	Price	Amount
1 Concepts and Applications—Spatial Order		$12 00	
2 A Nomagraphic Representation of the Geoid		$ 1 75	
3 Implicit Map Projections in Computer Printouts		$ 2 25	
7 The Geometry of Mixed Hexagonal Hierarchies, etc.		$ 3 00	
8 Energy Flow and Spatial Order, etc.		$ 3 25	
9. The Continent Problem, etc		$ 2 25	
10 Distances and Land Values, etc		$ 1.50	
11 Space Straightening and Flattening		$ 1 50	
13 Spatial Order in Fluvial Systems, etc		$ 3 80	
14 A Nomographic Representation of Trajectories		$ 1 50	
15 A Two-Dimensional Interpolation Function, etc.		$ 1.50	
16 N-Dimensional Spatial Analysis, etc.		$ 4 20	
17. Hyper-surfaces and Geodesic Lines, etc.		$ 1 50	
18 A Note on Stream Ordering and Contour Mapping		$ 2 00	
19 Hierarchical Systems Cities, Rivers, etc.		$10 00	
20. A Study of Rivers and Other Branching Systems		$ 1 50	
21 Geography and an Existence Theorem, etc.		$ 4 50	
22. A Mathematical Representation, etc		$ 4.50	
23. Plane Globe Projection, etc.		$18 00	
24 The Descriptive Geometry of a Collection, etc		$ 2.25	
25 The Law of Travel and Its Application to Rail Traffic		$ 7.75	
26 Minimum Time Paths and the Migration of Arctic Tern		$ 2 00	
27 Notes on the Methodology, etc		$ 2.75	
28 A Minimum Path Problem Reconsidered		$ 1 50	
29 The Use of the Geodesic Curvature, etc.		$ 1.50	

(over)

CUT ON DOTTED LINE

30 Numerical-Geometrical Techniques, etc		$ 1 50	
31 Some Reflections on Three-Dimensional Geometry Concepts		$ 1 50	
32 Algorithms and Models Based on Projective Transformations, etc		$ 1 50	
33 Graphical Representation of a Matrix with Applications, etc		$ 3 75	
34 Some Thought on Optimal Mapping and Coding of Surfaces		$ 1 50	
35 Homological Transformations in Four-Dimensional Space		$ 1 50	
36 A Study of the Movement of a Point on a Plane, etc		$ 1 50	
37 Set of Equal-Value Antipodal Points, etc		$ 1 50	
38 An Outline for the Theory of Man-Made Space, etc		$ 4 00	
39 A Structural Taxonomy of Spatial Hierarchies		$ 4 00	
40 The Determination of Fixed-Points, etc		$ 1 50	
41 Law and Order in the Human Lung		$ 4 00	
42 The Hexagon as a Spatial Average		$ 2 00	
43 The Two Dimensional Spatial Organization, etc		$ 2 50	
44 The Sandwich Theorem — A Basic One for Geography		$ 5 25	
45 Relations Between Horton's Laws and Hydraulic Geometry, etc		$ 3 00	
46 The Two-Dimensional Spatial Analysis, etc		$ 4 00	
47 A Computer Program for Mixed Hexagonal Hierarchies		$13 00	
48 Fred K Schaefer and the Science of Geography		$ 1 50	
49 Sparking Potential, Personal Interaction and Social Distance, etc		$ 1 50	
50 Macroscopic Aspects of Metropolitan Evolution		$ 7 25	
51 Notes on the Friction of Distance		$ 1 50	
52 Tabulations of Data on Area, Population, Income, etc		$19 25	
53 The Geographical Distribution of Income, etc		$ 1 50	
54 A Description of the 1967-68 United States Income, etc		$ 1 50	
55 National and Regional Parameters of Growth and Distribution, etc		$ 2 25	
56 Allometric Growth in Social Systems		$ 1 50	
57 Allometry in Micro-Environment Morphology		$ 2 50	

OTHER PAPERS	Quantity	Price	Amou
1 The Potential of Video Tape Recorders, etc		$ 1 50	
2 CPED Computer Programs in Environmental Design		$210 00	
3 Computer Aided Space Planning		$ 1 50	
4 Interactive Mapping of Urban Data		$ 1 50	
5 Interactive Graphics Comes of Age		$ 1 50	
6 Computer Graphics A Perspective		$ 1 50	
7 How to Talk to a Computer Consultant		$ 1 00	
8 Graphic Technology and the Display etc		$ 1 00	

PAYMENT

Checks should be made payable to Harvard University, Lab for Computer Graphics.

Total Amount $________ ________ Check Enclosed ____________ Purchase Order # ____________ Date ____________________

Orders should be sent to The Lab for Computer Graphics, 520 Gund Hall, Harvard University, Cambridge, Mass 02138.

SHIP TO:

Name ____________________

Organization ____________________

Address ____________________

____

BILL TO:

Name ____________________

Organization ____________________

Address ____________________

____

CUT ON DOTTED LINE

Order Form

Publications

Publication Distribution — General Information

Postage and Handling
Publications are sent as third class mail, printed matter via surface mail unless other arrangements have been made in advance Price estimates for air mail shipment of specific items to another country will be sent on request Alternatively, additional funds to cover the cost of air mail shipment may be enclosed with an order Any funds in excess of the amount required will be refunded

Telephone Orders
Telephone orders will be accepted from organizations within the U S A purchase order number is required Orders may be placed during the hours of 9 30 AM to 4 30 PM by calling area code (617) 495-2526

Payment
Payment should accompany an order unless a purchase order number has been included Payment must accompany all orders of less than $5 00 Orders to be shipped ouside the U S must include advance payment in the form of a bank draft payable in U S dollars

AUTOMATED CARTOGRAPHY PAPERS	Quantity	Price	Amount
1 Summary Report		$ 9 00	
2 Summary Graphics		$ 9 75	
3 Computer Equipment and Programs		$ 5 00	
4 Symap Interpolation Characteristics		$ 4 20	
5 St Louis Region Case Study		$ 19 75	
6 Kansas City Region Case Study		$ 6 75	
7. Montreal Region Case Study		$ 3 75	
8 Connecticut Case Study		$ 3 00	
9 Southern New England Case Study		$ 3 50	
10 Puget Sound Region Case Study		$ 2 00	
11. California Region Case Study		$ 3.00	
12 Boston Region Case Study		$ 3 25	
13 The Use of Computer Graphics in Planning		$ 1 00	
14. Maps a la Computers		$ 1 00	
15 Computer Cartography		$ 3 00	
16 Thematic Cartography		$ 3 00	
17 Management by Computer Graphics		$ 1 00	
18 Color in Art		$ 11 00	
19 Manifested Destiny: A Graphic Account of the Settlement, etc.		$ 2 00	
20 Programming for Transportability, etc.		$ 2 50	
21 Computer Cartography World-Wide etc.		$375 00	

THEORETICAL GEOGRAPHY PAPERS	Quantity	Price	Amount
1 Concepts and Applications—Spatial Order		$12 00	
2 A Nomagraphic Representation of the Geoid		$ 1 75	
3 Implicit Map Projections in Computer Printouts		$ 2 25	
7 The Geometry of Mixed Hexagonal Hierarchies, etc.		$ 3 00	
8 Energy Flow and Spatial Order, etc.		$ 3 25	
9. The Continent Problem, etc		$ 2 25	
10 Distances and Land Values, etc		$ 1 50	
11 Space Straightening and Flattening		$ 1 50	
13 Spatial Order in Fluvial Systems, etc.		$ 3 80	
14 A Nomographic Representation of Trajectories		$ 1 50	
15 A Two-Dimensional Interpolation Function, etc		$ 1 50	
16 N-Dimensional Spatial Analysis, etc		$ 4 20	
17 Hyper-surfaces and Geodesic Lines, etc.		$ 1 50	
18 A Note on Stream Ordering and Contour Mapping		$ 2 00	
19 Hierarchical Systems Cities, Rivers, etc.		$10 00	
20 A Study of Rivers and Other Branching Systems		$ 1 50	
21 Geography and an Existence Theorem, etc		$ 4 50	
22 A Mathematical Representation, etc		$ 4 50	
23 Plane Globe Projection, etc		$18 00	
24 The Descriptive Geometry of a Collection, etc.		$ 2 25	
25 The Law of Travel and Its Application to Rail Traffic		$ 7 75	
26 Minimum Time Paths and the Migration of Arctic Tern		$ 2 00	
27 Notes on the Methodology, etc.		$ 2 75	
28 A Minimum Path Problem Reconsidered		$ 1 50	
29 The Use of the Geodesic Curvature, etc		$ 1 50	

(over)

CUT ON DOTTED LINE

		Quantity	Price	Amount
30	Numerical-Geometrical Techniques, etc		$ 1.50	
31	Some Reflections on Three-Dimensional Geometry Concepts		$ 1 50	
32	Algorithms and Models Based on Projective Transformations, etc		$ 1 50	
33	Graphical Representation of a Matrix with Applications, etc.		$ 3 75	
34	Some Thought on Optimal Mapping and Coding of Surfaces		$ 1 50	
35	Homological Transformations in Four-Dimensional Space		$ 1 50	
36	A Study of the Movement of a Point on a Plane, etc		$ 1 50	
37	Set of Equal-Value Antipodal Points, etc		$ 1 50	
38	An Outline for the Theory of Man-Made Space, etc		$ 4 00	
39	A Structural Taxonomy of Spatial Hierarchies		$ 4 00	
40	The Determination of Fixed-Points, etc.		$ 1 50	
41	Law and Order in the Human Lung		$ 4 00	
42	The Hexagon as a Spatial Average		$ 2 00	
43	The Two Dimensional Spatial Organization, etc		$ 2 50	
44	The Sandwich Theorem — A Basic One for Geography		$ 5 25	
45	Relations Between Horton's Laws and Hydraulic Geometry, etc		$ 3 00	
46	The Two-Dimensional Spatial Analysis, etc.		$ 4 00	
47	A Computer Program for Mixed Hexagonal Hierarchies		$13 00	
48	Fred K Schaefer and the Science of Geography		$ 1 50	
49	Sparking Potential, Personal Interaction and Social Distance, etc		$ 1 50	
50	Macroscopic Aspects of Metropolitan Evolution		$ 7 25	
51	Notes on the Friction of Distance		$ 1 50	
52	Tabulations of Data on Area, Population, Income, etc		$19 25	
53	The Geographical Distribution of Income, etc		$ 1 50	
54.	A Description of the 1967-68 United States Income, etc		$ 1 50	
55	National and Regional Parameters of Growth and Distribution, etc		$ 2 25	
56	Allometric Growth in Social Systems		$ 1 50	
57	Allometry in Micro-Environment Morphology		$ 2 50	

OTHER PAPERS		Quantity	Price	Amount
1	The Potential of Video Tape Recorders, etc.		$ 1 50	
2	CPED Computer Programs in Environmental Design		$210 00	
3	Computer Aided Space Planning		$ 1 50	
4	Interactive Mapping of Urban Data		$ 1 50	
5	Interactive Graphics Comes of Age		$ 1 50	
6	Computer Graphics A Perspective		$ 1 50	
7	How to Talk to a Computer Consultant		$ 1 00	
8	Graphic Technology and the Display etc		$ 1 00	

PAYMENT

Checks should be made payable to Harvard University, Lab for Computer Graphics.

Total Amount $________ Check Enclosed ________ Purchase Order # ________ Date: ________

Orders should be sent to The Lab for Computer Graphics, 520 Gund Hall, Harvard University, Cambridge, Mass 02138.

SHIP TO:

Name ________

Organization ________

Address ________

BILL TO:

Name ________

Organization ________

Address ________

CUT ON DOTTED LINE

ILLUSTRATION CREDITS

(Where there are two or more illustrations, titles read upper to lower, left to right)

Page	Subject	Data Source	Computer Program	Prepared By
Front Cover	Per Capita Rate of GNP Growth. 1965-1972	WDB-I 1974 World Bank Atlas	CALFORM	David Sheehan Geoffrey Dutton Peter Corbin
2	Dietery Energy Supply: 1974	1975 World Population Data Sheet	CALFORM	Geoffrey Dutton
3	Upper Left· 1970 U.S. Population Density Middle Left· U.S. Land Elevations Lower Left. U.S. State Outlines Upper Right. 1970 SMSA Outlines Middle Right· 1960 U.S. Population Potential Lower Right: % Homes With Oil Heat	U.S. Census Standard Topographic Map County DIME File County DIME File U.S. Census Arthur D. Little, Inc.	DOT-MAP* DOT-MAP* POLYVRT POLYVRT DOT-MAP* CALFORM	Geoffrey Dutton Geoffrey Dutton Nichoias Chrisman Dennis White and Nicholas Chrisman Geoffrey Dutton Nicholas Chrisman
4	Boston 1970 Census Tracts Atlanta 1970 Census Tracts	Urban Atlas Files Urban Atlas Files	CALFORM CALFORM	Denis White Denis White
5	Los Angeles 1970 Census Tracts Chicago 1970 Census Tracts	Urban Atlas Files Urban Atlas Files	POLYVRT POLYVRT	Denis White Denis White
7	Locally prepared SYMAPS	Boston Local Maps	SYMAP	David Sheehan
8	Locally prepared GRID	Boston Local Maps	GRID	David Sinton
9	Massachusetts Tourism Attraction by Towns	Mass. Dept of Commerce	CALFORM	David Sheehan
10	Montegna Bay SYMAP and SYMVU Contour maps	Symap Self Study Course	SYMAP and SYMVU	David Sheehan
11	1970 Population by Nations (top) Montegna Bay Symap Symvu Conformant maps	United Nations Yearbook Symap Self Study Course	SYMVU SYMAP and SYMVU	David Sheehan Geoffrey Dutton

*Program under development, for further information contact the Laboratory.

ILLUSTRATION CREDITS

(Where there are two or more illustrations, titles read upper to lower, left to right)

Page	Subject	Data Source	Computer Program	Prepared By
13	Diagram of POLYVRT Components Fresno SMSA 1970 Census Tracts	POLYVRT Manual Urban Atlas	POLYVRT POLYVRT	Nicholas Chrisman Denis White
15	Conterminous United States	County DIME File	INPOM	Geoffrey Dutton
17	Selected U.S. Population Densities from "Manifested Destiny"	U.S. Census of Population 1790-1970	ASPEX	Geoffrey Dutton
18	Left: Standard Economic Area (SEA) Center. Standard Metropolitan Statistical Areas (SMSA) Right: Countries of the World	County DIME File County DIME File WDB-I	POLYVRT/INPOM POLYVRT POLYVRT	Nicholas Chrisman Nicholas Chrisman Nicholas Chrisman
Inside Front Cover	U.S. 1970 Population Density	U.S. 1970 Population Census	SYMVU	Geoffrey Dutton James Little

Staff Members of the Laboratory for Computer Graphics & Spatial Analysis

Brian J.L. Berry, Director
Allan H. Schmidt, Assoc. Director
Eric Teicholz, Assoc. Director
Howard T. Fisher, Director Emeritus

Richard Carling, Computer Programmer
Nicholas Chrisman, Sr. Research Analyst
James Dougenik, Research Analyst
Geoffrey Dutton, Research Analyst
Randolph Franklin, Research Assistant
Denis White, Research Analyst
David Sheehan, Consultant

Janice Kostka, Admin. Asst.
Barbara Petersen, Program Distribution

MANEJO DE SISTEMAS DE INFORMACION GEOGRAFICA EN PLANEACION

MANUAL RESUMIDO DEL PROGRAMA GRID

ARQ. ALEJANDRO VILLANUEVA EGAN

NOVIEMBRE DE 1977.

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 y 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 horizontales por renglón y de izquierda a derecha en cada renglón.

El tamaño y forma de la malla debe ser especificada por el 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 o 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 la localización de escala están también disponibles -- en GRID.

1.2 ALIMENTACION AL PROGRAMA

Para obtener una representación gráfica (o un mapa), el usuario debe proporcionar tres conjuntos de instrucciones y tiene la opción de pro-- porcionar 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 o 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 área de estudio en el caso de que 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 EL 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 o 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 el GRID como la "etapa de trabajo final " en una serie de " etapas " que están compuestas de rutinas para manipula-- ciones estadísticas complejas..

En este caso, los datos utilizados para crear la representación gráfica se transfieren al programa GRID en la forma de arreglos binarios, por un arreglo (o registro lógico) por cada renglón del mapa. El programa espera 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 o 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 - construído dentro del programa permite graficar al usuario tantas - células dato como sean necesarias para un mapa. Para lograr tal obje tivo, se pueden dividir los archivos que se salgan del limite en conjun tos 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 contormo - - irregular se organicen de tal modo que el primer conjunto este - - referido a la sección superior del mapa y al 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 especifica 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.

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) Columas 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.

campos 3 y 4 : el tamaño de cada célula de malla en términos del número de caracteres en la dirección vertical (campo 3) y horizontal (campo 4). (Recordando que un caracter mide 1/8 ''verticalmente 1/10'' horizontalmente.

Los números que van en los cuatro campos descritos deban 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 múltiples, especifique en el campo 3 el número de conjuntos de datos a ser mapeados.

Es estandar es un conjunto de datos y alimentación de datos con la opción B.

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 o intervalos de clase en los - que se divide el valor del rango total (desde 2 hatas 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 o intervalo. El valor del rango total de los datos (modificado por el mínimo y el má ximo de las 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 o intervalo. (vêr 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.

Columnas 1-10 son usadas para especificar el simbolismo general para cada nivel (columna 1 para el símbolo 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 área 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 númerica que es impresa con los niveles. El estandar es no histograma ó diagrama de barras y la inclución 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 puede 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 ENOTEXT 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 simbolo . El -- rango de los datos puede dividirse en 20 niveles (1a. si se especificó - un valor máximo en la opción). 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 de la opción 1 y en el número de niveles la opción 3.

Para especificar simbolismo PUNTUAL, perfore 1.0 en el campo 1. Para restablecer el simbolismo de la escala 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 a partir de las - cuales son numerados todos los renglones y columnas.

Si las coordenadas de la CRM no son especificadas, 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 número 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 o tercios y usarse un sistema de numeración no-contínuo. 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 numérica 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 o 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 mas fácil de codificar es un cero o blanco a menos que el cero sea un valor válido. Por simplicidad, puede codificarse directamente al valor de - 999999.0.

4.2 Paquete de Contorno irregular.

Para simplificar el manejo de contornos irregulares se ha construído una pequeña rutina dentro del programa, de tal forma que el -- usuario puede especificar la forma del contorno sin tener que llenar el rectángulo 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 número 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 11-16 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 (o renglones especificados en las columnas 1-5) superior, la segunda tarjeta al segundo renglón (o 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 reemplazado por un nuevo paquete o 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 un 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 o en un archivo de datos en cinta o disco; o,

ii). Los análisis estadísticos que van a realizarse sobre una variable, o variables, para derivar el valor a ser mapeado.

Esta subrutina es llamada por el programa principal una vez para cada célula dato que va a ser mapeada. Cada vez que es llamada, lee la tarjeta de datos o 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 va 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.

Al 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 o la última tarjeta de datos, si los datos están en tarjetas.
(Ver ejemplo 9)

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

MANEJO DE SISTEMAS DE INFORMACION GEOGRAFICA EN PLANEACION

SELECTED COMPUTER PROGRAMS

FROM THE DEPARTAMENT OF GEOGRAPHY.

UNIVERSITY OF MICHIGAN
ANN ARBOR

NOVIEMBRE, 1977.

S E L E C T E D

C O M P U T E R   P R O G R A M S

Second Printing
1973

A Michigan Geography Publication

Edited by

Waldo R. Tobler

Department of Geography

University of Michigan

Ann Arbor

1970

References:

D. Boettner, ed., Michigan Terminal System, Computing Center, University of Michigan, Ann Arbor, 1967.

C. Dinger, et al, Computer Programs in Oceanography, National Oceanographic Data Center, Washington D.C., 1967, 95 pp.

W.J. Dixon, BMD Biomedical Computer Programs, University of California Press, Los Angeles, 1968, 600 pp.

E. Fronczak, The University of Michigan Digital Plotting System, U.M. Computing Center, Ann Arbor, September 1966, Revisions + 140 pp.

W.L. Garrison, and W.C. Krumbein, Environmental Sciences Computer Study, Northwestern University, Evanston, 15 reports, 1962-1965.

D.F. Marble, Some Computer Programs for Geographic Research, Department of Geography, Northwestern University, Evanston, 1967.

D. Merriam, ed., Computer Contributions, Kansas State Geological Survey, Lawrence, 35 reports, 1966-1970.

NASA, Computer Program Abstracts, GPO, Washington D.C., quarterly.

G. Rushton, et al, Technical Reports, Computer Institute for Social Research, Michigan State University, East Lansing, various dates.

F. Sarno, Polygraphics Manual, Polytechnic Institute of Brooklyn, New York, 1968.

J. Sonquist, and J. Morgan, The Detection of Interaction Effects, Survey Research Center, ISR, University of Michigan, Ann Arbor, 1964.

J.R. Tarrant, Computers in Geography, University of East Anglia, Norwich, Norfolk, England, 1970.

Introduction

The program listings and documentation presented here have been assembled from among those in use at the Geography Department at the University of Michigan. They are obviously the result of the efforts of many individuals, foremost of whom is Frank Rens, for several years one of my graduate assistants. More recently H. Moellering and D. Rhynsburger undertook the documentation task and prepared the programs for publication. None of this work would have been possible without the splendid cooperation from many individuals connected with the University Computing Center. Although the programs have been tested extensively (some for eight years) it is necessary to make the usual disclaimer.

The programs are in FORTRAN (IV, G level) and normally are used under control of the Michigan Terminal System on the IBM 360/67, a very large system by current standards. Several programs make use of the 763 Calcomp plotter via the *PLOTSYS program as described in the plotting manual; some have also been used with the SC 4020 CRT display after relatively minor changes.

Not all of the programs available to the Geography Department are included here. Specifically, the multivariate analysis programs from the BMD system are available for call from teletype terminals; many of the programs listed in the Kansas Geological Survey series, the Northwestern University series, the Michigan State series, et cetera, have been converted for use on the MTS system. The Institute for Social Research and the Computing Center also maintain large files of programs. Persons interested in these programs should consider obtaining the appropriate references from those listed on the preceding page.

Waldo R. Tobler
Professor of Geography
June 1970

Geographical Publications of Related Interest

B. Berry and D. Marble, Spatial Analysis, Englewood Cliffs, Prentice Hall, 1966.

William Bunge, Theoretical Geography, Second Revised and Enlarged Edition, Lund, Gleerup, 1966.

R.J. Chorley, et al, Models in Geography, New York, Barnes and Noble, 1967.

R. Chorley and P. Haggett, Network Models in Geography, London, 1969.

J. Cole and C. King, Quantative Geography, New York, J. Wiley, 1968.

W. Garrison and D. Marble, eds, Quantative Geography, Vols. 1 & 2, Evanston, Northwestern University, 1967.

P. Haggett, Locational Analysis in Human Geography, New York, St. Martin's, 1966.

L. King, Statistical Analysis in Geography, Englewood Cliffs, Prentice Hall, 1969.

L. King, editor, Geographical Analysis, a quarterly international journal of theoretical geography published by the Ohio State University Press, Columbus, Ohio.

J. Nystuen, ed., Discussion Papers, 1-12, Michigan Interuniversity Community of Mathematical Geographers, 1967; available as OP-33067 from University Microfilms.

A. Wrobel, and B. Berry, eds., Economic Regionalization and Numerical Methods, Geographia Polonica Nr. 15, Warszawa, 1968.

LIST OF PROGRAMS

Program Availability

Card decks or magnetic tape copies of the programs listed in this document (and others)are now available from the Geography Program Exchange, Computer Institute for Social Research, Michigan State University, East Lansing, Michigan, 48823, U.S.A. : telephone, 517-353-2042, attn. Prof. R. Wittick.

Disclaimer

Although the programs contained herein have been tested extensively no warranty, expressed or implied, is made by the University of Michigan or individuals as to the accuracy and functioning of the program and the related program descriptions.

DETERMINATION OF GEOGRAPHICAL ORIGINS

Purpose: The program estimates sets of source coordinates from empirical geographical distributions.

Description: The program is a simple modification of an earlier program by Casetti and Semple. The modification consists of (a) generalization of the types of trends which can be estimated, (b) addition of a contouring subroutine for automatic plotting, and (c) minor improvements the program flow and output formats. A complete description is given in the paper by Casetti and Semple, which is available. In brief, the procedure may be considered to be an inverse diffusion operator in two dimensions. The contouring subroutine is general and can be used with other programs.

Deck Make up: 1) Title Card

2) Format Card

3) Control Card

4) First Contouring Controls (optional)

5) Second Contouring Controls (optional)

6) Third Contouring Controls (optional)

7) Data Cards

Cards (1) through (7) may be repeated as many times as desired. All are read from unit 5, printing is on unit 6, and punching of output on unit 7.

Card Format: 1) Title Card

Any title in columns 1 through 72.

2) Format Card

The format for one observation consisting of a 4 character name, x coordinate, y coordinate, and weight z, in that order. The name should be specified as an A4 field, the remaining variables as E or F fields. The format is punched in columns 1 through 72.

3) Control Card (6I3)

column	
1-3	N the number of observations.
4-6	IKOT a switch to define the independent variable.
	1 1.0/distance
	2 Log_e (distance)
	3 exp (distance)
	4 distance
	5 distance squared.
7-9	ILOG
	1 implies take natural log of dependent variable z.
	0 default, no transformation of z.
10-12	KOPT printing switch
	1 print all correlation matrices
	2 print correlation map
	3 print both map and matrix.
13-15	IPUN.GT.0 to punch results
16-18	NN The size of the correlation matrix. Default = 11

4) First Contour Control Card

Use only if KOPT.GT.1.

Blank card yields default cases.

column

column	name	description
1-2	CON	number of contour intervals Default = 10
3	TOUR	Type of contour interval 0 = values calculated (Default case) 1 = values specified 2 = variable interval specified 3 = standard deviation units.
4	LINES	type of contour map 0 continuous bands 1 alternate bands 2 lines
5-9	INCHES	(F5.2) width of contour map, up to 12.7 inches.

5) Second Contour Control Card

Use only if KOPT.GT.1 and TOUR = 1

column

column	description
1-10	Lowest contour value (F10.0)
11-20	Highest contour value (F10.0)

6) Third Contour Control Card

Use only if KOPT.GT.1 and TOUR = 2

Up to 15 variable contour levels punched as (15F4.0).

7) Data Cards

name, x, y, z punched as described on the Format Card.

References: E. Casetti, and R.K. Semple, A Method for the Stepwise Separation of Spatial Trends, MICMOG paper #11, 1968. Available from University Microfilms in OP - 33067.

```
C     PROGRAM "GEOFIT"
C
C     THE PROGRAM ESTIMATES SOURCE COORDINATES FOR SPATIAL TRENDS
C     BASED ON MICMOG PAPER # 11,1968, BY CASETTI AND SEMPLE
C     REVISION BY L. TREVILLIAN ACCORDING TO INSTRUCTIONS FROM
C     W. TOBLER, GEOGRAPHY DEPARTMENT, UNIVERSITY OF MICHIGAN
C     READS 5, WRITES 6, PUNCHES 7
      INTEGER CON,TOUR,RX,RY,GX,GY,HX,HY
      REAL INCHES,INCON
      REAL*8 ZVA,ZAV,XAV,XVA,YAV,YVA,DAV,DVA,ZB,RAV,RVA(10)
      DIMENSION INCON(15),DRANG(25),FCN(6),TITL(18),EXVT(10)
      DIMENSION X(500),Y(500),Z(500),NAME(500),SLOPE(10),A(500),D(500),
     1 ZZ(500),ZI(500),XORG(10),YORG(10),R(50,50),H(50,50)
     2,G(50,50),SAVIT(500),CFPT(10),TIT(18),FMT(18)
      DATA TITL/'MAP OF CORRELATION COEFFICIENTS',10*' '/
C     SUBROUTINES
C        CORR COMPUTES THE CORRELATION COEFFICIENTS AND FINDS
C        THE MAXIMUM COEFFICIENTS.
C        STATS FINDS MEAN, MAX, MIN, VARIANCE OF AN ARRAY
C        COMFCN FINDS AND TRANSFORMS DISTANCES
C        SCON PRINTS CONTOUR MAP OF CORRELATIONS
C
C     INPUT VARIABLES
C      N IS THE NUMBER OF DATA POINTS
C      IKOT DEFINES THE INDEPENDANT VARIABLE
C      IF IKOT=1, 1.0/DISTANCE
C      IF IKOT=2, LOG(DISTANCE)
C      IF IKOT=3, EXP(DISTANCE)
C      IF IKOT=4, DISTANCE
C      IF IKOT=5, DISTANCE SQUARED
C     NON-ZERO ILOG YIELDS LOG TRANSFORM OF DEPENDANT VARIABLE
C      I.E., ILOG=1, IKOT=4 YIELDS EXPONENTIAL DECAY
C      I.E., ILOG=1, IKOT=5 YIELDS GAUSSIAN CURVE
C      KOPT, A SWITCH
C        IF KOPT=1, PRINT ALL CORRELATION MATRICES
C        IF KOPT=2, A CONTOUR MAP WILL BE PRINTED
C        IF KOPT =3, BOTH THE MATRIX AND THE MAP WILL BE PRINTED
C      IPUN .GT. 0 TO PUNCH RESULTS
C      NN IS THE SIZE OF THE CORRELATION MATRIX.
C
1       READ(5,507,END=999) (TIT(I),I=1,18)
        READ(5,507) (FMT(I),I=1,18)
507     FORMAT(18A4)
        READ(5,1000) N,IKOT,ILOG,KOPT,IPUN,NN
1000    FORMAT(6I3)
        WRITE(6,1000) IKOT,ILOG,KOPT,IPUN,NN,N
C
C      IF A CONTOUR MAP IS DESIRED, READ PARAMETERS NECESSARY FOR
C      CALL ON SCON. BLANK CARD YIELDS VALID DEFAULT OPTIONS.
        IF(KOPT.LE.1) GO TO 30
        READ(5,1010) CON,TOUR,LINES,INCHES
        IF(TOUR.GT.3 .OR. TOUR .LE. 0)  GO TO 30
        GO TO (10,20,30),TOUR
10      READ(5,1020) VL,VU
        GO TO 30
```

```
      IF(RY.GE.NN) YYY=SMAY
      HA=(2.*VA)/(NN-1.)
C
C     SECOND ITERATION.  GRID SIZE HA
C
      CALL CORR(N,Z,ZAV,ZVA,NN,HA,H,HBIG,HX,HY,X,XXX,Y,YYY,IKOT)
      XXXX= XXX+(HX-2.)*HA
      YYYY=YYY+(NN-HY-1.0)*HA
      IF(HX.LE.1)XXXX=XXX
      IF(HX.GE.NN)XXXX=XXX+(NN-3.)*HA
      IF(HY.LE.1) YYYY=YYY+(NN-3.0)*HA
      IF(HY.GE.NN) YYYY=YYY
      GA=(2.*HA)/(NN-1.)
C
C     THIRD ITERATION.  GRID SIZE GA
C
      CALL CORR(N,Z,ZAV,ZVA,NN,GA,G,GBIG,GX,GY,X,XXXX,Y,YYYY,IKOT)
C
C    FIND THE MAP COORDINATES OF MAXIMUM CORRELATION AND RECOMPUTE THE
C    FUNCTION
C
      XORG(ITER)=XXXX+((GX-1.)*GA)
      YORG(ITER)=YYYY+(NN-RY)*GA
      CALL COMFCN(X,Y,N,D,XORG(ITER),YORG(ITER),IKOT)
      BABA=0.
      SLOPE(ITER)=0.
      CALL STATS(N,IZ,D,DAV,DVA,DMAX,DMIN,500,1)
      DO 90 I=1,N
      D(I)=D(I)-DAV
      BABA=BABA+D(I)*Z(I)
90    SLOPE(ITER)=SLOPE(ITER)+D(I)*D(I)
      SLOPE(ITER)=BABA/SLOPE(ITER)
      CEPT(ITER)=ZAV-SLOPE(ITER)*DAV
C
C    FIND THE EXPLAINED VARIANCE AND CALCULATE THE RESIDUALS
C
      DO 100 I=1,N
      D(I)=SLOPE(ITER)*D(I)
      SAVIT(I)=SAVIT(I)+D(I)
      ZZ(I)=Z(I)-D(I)
100   CONTINUE
      CALL STATS(N,IZ,ZZ,RAV,RVA(ITER),RMAX,RMIN,500,1)
      CALL STATS(N,IZ,D,DAV,DVA,DMAX,DMIN,500,1)
      EXVAR=DVA
      UNEXP=RVA(ITER)
      EXVT(ITER)=DVA/ZB*100.
      EXVC =EXVC+EXVT(ITER)
C
C    SECTION TO PRINT OUT DATA AND RESULTS
C
      ZVA=ZVA*ZVA
      WRITE(6,501) ITER
      WRITE(6,500) (NAME(I),X(I),Y(I),Z(I),D(I),ZZ(I),I=1,N)
      WRITE(6,620) XAV,YAV,ZAV,DAV,RAV,XVA,YVA,ZVA,DVA,RVA(ITER)
      WRITE(6,2030) ITER
      K=1
```

```
      RY=NN-RY+1
      WRITE(6.590) K,RX,RY,RBIG,VA
      IF (KOPT-2) 520,530,520
520   WRITE(6.601)
      DO 521 I=1,NN
521   WRITE(6.600) (R(I,J),J=1,NN)
530   K=2
      HY=NN-HY+1
      WRITE(6.590) K,HX,HY,HBIG,HA
      IF(KOPT-2) 540,550,540
540   WRITE(6.601)
      DO 541 I=1,NN
541   WRITE(6.600) (H(I,J),J=1,NN)
550   K=3
      GY=NN-GY+1
      WRITE(6.590) K.GX,GY,GBIG,GA
      IF(KOPT-2) 560,570,560
560   WRITE(6.601)
      DO 561 I=1,NN
561   WRITE(6.600) (G(I,J),J=1,NN)
570   WRITE(6.610) XORG(ITER),YORG(ITER),EXVAR,UNEXP
      WRITE(6.2020) EXVT(ITER),EXVC
      IF(KOPT.EQ.1) GO TO 250
C
C   CALL SCON WITH ARGUMENT R
C
      CALL SCON(R.NN,NN,TITL,CON.TOUR,LINES,INCHES,DRANG,VL,VU,
     & SMAX,SMAY,VA)
250   CONTINUE
      I=ITER-1
      IF(I.LE.0) GO TO 209
      IF(EXVT(ITER).GT.EXVT(I)) ITER=I
      IF (ITER.EQ.I) GO TO 210
209   CONTINUE
      IF((EXVT(ITER).LE.1.).OR.(ITER.GE.5)) GO TO 210
      CALL STATS(N,IZ,ZZ,ZAV,ZVA,ZMAX,ZMIN,500,1)
      DO 200 I=1,N
200   Z(I)=ZZ(I)-ZAV
      ZAV=0.
      ZVA=DSQRT(ZVA)
      GO TO 80
210   CONTINUE
      EVEC=0.0
      WRITE(6.2000)
      I=0
      BABA=100.*RVA(1)/(100.-EXVT(1))
      WRITE(6.1999) I,BABA,EVEC,EVEC
      DO 220 I=1,ITER
      EVEC=EVEC+EXVT(I)
220   WRITE(6.2010) I,XORG(I),YORG(I),RVA(I),EXVT(I),EVEC,CEPT(I),
     & SLOPE(I)
      WRITE(6.630)
      IF (ILOG.NE.0) WRITE(6,640)
      GO TO (631,632,633,634,635),IKOT
631   WRITE(6.641)
      GO TO 650
```

```
632     WRITE(6,642)
        GO TO 650
633     WRITE(6,643)
        GO TO 650
634     WRITE(6,644)
        GO TO 650
635     WRITE(6,645)
650     CONTINUE
        WRITE(6,508) (TIT(I),I=1,18)
        WRITE(6,505) ITER
505     FORMAT(' THE FINAL RESULTS OF THE ' I2,' ITERATIONS ARE:'/)
        DO 300 I=1,N
        IF (ILOG.EQ.0) GO TO 299
        ZI(I)=EXP(ZI(I))
        SAVIT(I)=EXP(SAVIT(I))
299     ZZ(I)=ZI(I)-SAVIT(I)
        IF (IPUN.GT.0) WRITE(7,506) NAME(I),X(I),Y(I),ZI(I),SAVIT(I),
      & ZZ(I)
300     CONTINUE
        WRITE(6,500)(NAME(I),X(I),Y(I),ZI(I),SAVIT(I),ZZ(I),I=1,N)
500     FORMAT('0',48X,'DATA'/'0NAME',15X,'X',15X,'Y',15X,'Z',11X,
      1 'APPROX',6X,'RESIDUALS'//(' ',A4,10X,5(F10.4,5X)))
501     FORMAT('1THE RESULTS OF PASS NUMBER ',I2,' ARE AS FOLLOWS:')
506     FORMAT(' ',A4,5(F10.4,1X))
508     FORMAT('1',18A4,//)
590     FORMAT('0MAXIMUM CORRELATION ON ITERATION ',I1,
      1 ' OCCURRED AT ',I2,',',I2/' THE VALUE OF THE CORRELATION ',
      1 'COEFFICIENT WAS ',F8.5/' THE GRID SIZE WAS  ',F10.5)
600     FORMAT(20(' ',F6.4))
601     FORMAT('0MATRIX OF CORRELATION COEFFICIENTS'//)
610     FORMAT(//'0THE MAP COORDINATES OF HIGH CORRELATION ARE '
      1 ,2(F10.5,2X)/' THE EXPLAINED VARIANCE IS    ',E11.4/
      2 ' THE UNEXPLAINED VARIANCE IS  ',E11.4)
620     FORMAT('0MEAN',10X,5(E11.4,4X)/' VARIANCE',6X,5(E11.4,4X))
630     FORMAT(///' THE FUNCTION IS '//)
640     FORMAT(' BASE E LOGARITHM OF ')
641     FORMAT(' Z= A + B*(1.0/DISTANCE) ')
642     FORMAT(' Z = A + B*LOGE(DISTANCE) ')
643     FORMAT(' Z = A + B*EXP(DISTANCE) ')
644     FORMAT(' Z = A + B*DISTANCE ' )
645     FORMAT(' Z = A + B*(DISTANCE**2) ')
1010    FORMAT(I2,2I1,F5.2)
1020    FORMAT(2F10.0)
1030    FORMAT(15A4)
1050    FORMAT('1',24X,'ORIGINAL DATA'//' NAME',17X,'X',16X,'Y',16X,'Z'/
      &//(' ',A4,5X,3(5X,F12.4)))
1060    FORMAT('0MEAN',5X,3(5X,E12.5)/' VARIANCE ',3(5X,E12.5))
1999    FORMAT(5X,I1,36X,E12.5,5X,2(F7.3,8X))
2000    FORMAT('1SUMMARY'//71X,'CUMULATIVE'/' ITERATION',13X,'ORIGIN',
      & 15X,'RESIDUAL',6X,'EXPLAINED',5X,'EXPLAINED',5X,'CONSTANTS'/
      & '   NUMBER',11X,'X',12X,'Y',11X,'VARIANCE',5X,'VARIANCE(%)',
      & 3X,'VARIANCE(%)'//)
2010    FORMAT(5X,I1,7X,F10.4,5X,F10.4,4X,E12.5,5X,2(F7.3,8X),E12.5,
      & 2X,E12.5)
2020    FORMAT('0THE PERCENTAGE OF THE TOTAL VARIANCE EXPLAINED ON THIS'
      & ,' ITERATION IS ',F6.3,'%.'/' THE CUMULATIVE EXPLAINED VARIANCE'
```

```
      END
      SUBROUTINE CORR(N,Z,ZAV,ZVA,NN,VA,R,RBIG,RX,RY,X,
     & XMIN,Y,YMIN,IKOT)
      DIMENSION R(100,100),A(500),Z(500),X(500),Y(500)
      INTEGER RX,RY
      DOUBLE PRECISION ZVA,ZAV,AVA,AAV
C
C
C     CORR COMPUTES A MATRIX OF CORRELATION COEFFICIENTS AND FINDS THE
C     MAXIMUM OF THE COEFFICIENTS
C
C     A FUNCTION OF THE DISTANCES BETWEEN POINTS ON A GRID (DEFINED
C     BY XMIN,YMIN, AND VA) AND THE COORDINATES OF THE Z VECTOR
C     IS COMPUTED AND CORRELATED WITH THE ACTUAL VALUES OF Z.
C
      XX=XMIN
      YY=YMIN
C
C     N IS THE NUMBER OF DATA POINTS IN THE VECTOR,Z, WHICH IS TO BE
C     CORRELATED WITH THE VECTOR A.  ZAV AND ZVA ARE THE MEAN AND
C     STANDARD DEVIATION OF Z.
C     NN IS THE NUMBER OF ROWS(OR COLUMNS) OF THE COEFFICIENT MATRIX,R.
C     RBIG IS THE LARGEST CORRELATION COEFFICIENT, THE COORDINATES
C     OF RBIG ARE RX AND RY.
C     X AND Y ARE THE COORDINATES OF THE POINTS IN THE Z VECTOR.
C     XMIN AND YMIN ARE THE MINIMUM VALUES OF THESE COORDINATES.
C     IKOT DEFINES THE FUNCTION TO BE USED IN THE FIT.
C     SCALE IS A CONSTANT SCALE FACTOR.
C     VA IS THE GRID SIZE.
C
      RA=N
      DO 90 LL=1,NN
      L=NN-LL+1
      DO 80  M=1,NN
C
C     COMPUTE THE DISTANCES BETWEEN THE Z COORDINATES AND (XX,YY)
C
C
C     COMPUTE THE FUNCTION OF THE DISTANCES AS SPECIFIED BY IKOT
C
      CALL COMFCN(X,Y,N,A,XX,YY,IKOT)
60    IAV=1
C
C     FIND THE MEAN AND VARIANCE OF THE FUNCTIONAL POINTS
C
      CALL STATS(N,IAV,A,AAV,AVA,AMAX,ZMIN,500,1)
      DO 65 I=1,N
65    A(I)=A(I)-AAV
      AAV=0.
      AVA=DSQRT(AVA)
      R(L,M)=0.
C
C     FIND THE CORRELATION BETWEEN THE FUNCTIONAL AND THE OBSERVED VALUES
C
      DO 70 I=1,N
70    R(L,M)=R(L,M)+A(I)*Z(I)
```

```
      R(L,M)=R(L,M)/(RA*AVA*ZVA)
C
C    INCREMENT THE GRIC VALUES
C
80    XX=XMIN+M*VA
      XX=XMIN
90    YY=YMIN+LL*VA
C
C    FIND THE MAXIMUM COEFFICIENT AND ITS COORDINATES
C
      RBIG=ABS(R(1,1))
      RX=1
      RY=1
      DO 110 I=1,NN
      DO 110 J=1,NN
      IF(ABS(R(I,J))-RPIG) 110,110,100
100   RBIG=ABS(R(I,J))
      RX=J
      RY=I
110   CONTINUE
      RBIG=R(RY,RX)
      RETURN
      END
      SUBROUTINE SCON(Z,ROWS,COLS,TITL,CON,TOUR,LINES,INCHES,DRANG,
     & ZMIN,ZMAX,XMIN,YMIN,VA)
      DIMENSION FORM(10),TITL(18)  ,SYMTAB(21),PSYM(131),Z(100,100),
     1 INCON(15),DRANG(25),TABE(25),ZPREV(131),SYM(131)
      DIMENSION CUMH(27)
      REAL*8 ZBAR,ZSIG
      INTEGER ROWS,COLS,CON,TOUR,COLMAX,COL,PSYM,PS,GRID,TOUT,TLOW,
     1 SYM,BLANK,PLUS,MINUS,STAR,SYMTAB
      INTEGER XSTOP,YSTOP
      REAL MINZ,MAXZ,INCHES,INCON
      DATA SYMTAB/'1','2','3','4','5','6','7','8','9','A','B','C',
     1 'D','E','F','G','H','I','J','K'/,BLANK,PLUS,MINUS,STAR/' ',
     2 '+','-','*'/
C    ****************************************
C   ZERO COUNTERS AND ARRAYS USED IN FREQUENCY TABLE
C ******************************************
      WRITE(6,1290)
1290  FORMAT('1THE CONTOUR MAP WILL BE DEVELOPED FROM THE  ',
     & 'FOLLOWING DATA'///)
      DO 4 I=1,ROWS
4     WRITE(6,1300) (Z(I,J),J=1,COLS)
1300  FORMAT(15(' ',F7.4))
      WRITE(6,1241) XMIN,YMIN,VA
1241  FORMAT(///' THE RELATIVE ORIGIN IS ',F10.4,' , ',F10.4,' .'/
     & 'THE GRIC SIZE IS ',F10.4,' .')
      JJ=0
      MM=0
      SUM=0.0
      TLOW=0
      TOUT=0
      DO 5 I=1,25
5     TABE(I)=0.0
      N=ROWS*COLS
```

```
C     ***********************************************
C      N = NUMBER OF DATA POINTS
C     ***********************************************
       WRITE(6,1030) TITL,N,COLS,ROWS
1030    FORMAT(1H1,18A4//' MAP IS DEVELOPED FROM',I5,' GRIDDED VALUES',
      1 ' CONSISTING OF ',I3,' COLUMNS BY ',I3,' ROWS.'//)
        CALL STATS(ROWS,COLS,Z,ZBAR,ZSIG,MAXZ,MINZ,100,100)
        ZSIG=DSQRT(ZSIG)
C     ***********************************************
C   DETERMINE CONTOUR LINE ELEVATIONS
C   TOUR VALUE DETERMINES METHOD USED IN DEFINING CONTOURS
C     ***********************************************
       IF (TOUR .GT. 3 .OR. TOUR .LE. 0) GO TO 60
       GO TO (30,40,50),TOUR
30     WRITE (6,1060)
1060   FORMAT (25H CONTOUR LEVELS SPECIFIED)
       GO TO 70
40      WRITE(6,1080)
1080   FORMAT (37H VARIABLE CONTOUR INTERVALS SPECIFIED)
       ZMIN=DRANG(1)
       ZMAX=DRANG(CON)
       GO TO 70
50     ZMIN=ZBAR-3.0*ZSIG
       ZMAX=ZBAR+3.0*ZSIG
       CON=13
       WRITE (6,1090)
1090    FORMAT(30H CONTOUR INTERVAL = 0.5 SIGMA
       GO TO 70
60     WRITE (6,1100)
1100   FORMAT (26H CONTOUR LEVELS CALCULATED)
        ZMAX=MAXZ
       ZMIN=MINZ
70     WRITE(6,1110) ROWS,COLS,ZBAR,ZSIG,MAXZ,MINZ
1110   FORMAT (13H GRIDDED DATA/,42H LINEAR INTERPOLATION WITHIN GRID SQU
      1ARES /,17H OBSERVED VALUES /,  8H ROWS = ,I3,10H   COLS = ,
      2I3,/,8H ZBAR = ,F10.3,10H   ZSIG = ,F10.3,10H   MAXZ = ,F10.3,10H
      3  MINZ = ,F10.3)
C     ***********************************************
C   CALCULATE MAP WIDTH IN CHARACTERS (COL)
C     ***********************************************
       IF (INCHES.LE.0.0) COL=100
       IF (INCHES .GT. 12.7) COL=127
       IF (INCHES.GT.0.0.AND.INCHES.LE.12.7) COL=INCHES*10.0
       PS=(COL-COLS)/(COLS-1)
C     ***********************************************
C   TEST GRID SIZE--IF GRID SMALLER THAN 2 CHARACTERS BETWEEN COLUMNS,
C   AN ERROR STATEMENT IS PRINTED AND PLOT ABORTED
C     ***********************************************
       IF (PS.GT.1) GO TO 80
       WRITE (6,1120)
1120   FORMAT (25H TOO MANY COLUMNS IN GRID)
       GO TO 999
C
C   CALCULATE REQUIRED CONSTANTS
C
80     COLMAX=PS*(COLS-1)+COLS
```

```
       YINC=1.0/((PS+1.0)*0.6)
      GRID=PS+1
       XINC=1.0/GRID
      RANGZ=ZMAX-ZMIN
      RHIGH=ZMAX+RANGZ
      RLOW=ZMIN-RANGZ
      IF (CON.EC.0.OR.CON.GT.19) CON=5
       FCCN=CON
      CI=RANGZ/FCON
      IF (TOUR.EQ.2) GO TO 100
      JJJ=CCN+1
      DC 90 I=1,JJJ
90     DRANG(I)=ZMIN+(I-1)*CI
100   CONTINUE
      WRITE (6,1130)
1130  FCRMAT (1H /,16H VALUES EMPLOYED)
      WRITE (6,1140) ZMIN,ZMAX,CI,GRID,XINC,YINC,CON,PS,COLMAX
1140  FORMAT (8H ZMIN = ,F10.3,10H   ZMAX = ,F10.3,8H   CI = ,F10.3/,8H
     1GRID = ,I5,10H   XINC = ,F10.3,10H   YINC = ,F10.3,/,7H CCN = I5,
     2 8H   PS = ,I5,12H   COLMAX = ,I5)
C     *****************************************
C   WRITE OUT CONTOUR SYMBOL TABLE
C     *****************************************
      WRITE (6,1150)
1150  FORMAT (35H4CONTOUR SYMBOL TABLE IS AS FOLLOWS)
      IF (LINES.EC.1) WRITE (6,1160)
1160  FORMAT (43HO(ONLY EVEN SYMBOLS PRINTED IN BODY OF MAP) )
      WRITE (6,1170)
1170  FORMAT (7HOSYMBOL,11X,5HVALUE)
      IF (LINES.NE.0) GO TO 110
      WRITE (6,1180)
1180  FORMAT (5HO   *,6X,7HNO DATA,/,5H    -,4X,10HBELOW  LOW)
      WRITE (6,1190) (SYMTAB(I),DRANG(I),I=1,CON)
1190  FCRMAT (4X,A1,2X,F12.4)
      WRITE (6,1200)
1200  FORMAT (4X,1H+,4X,10HABOVE HIGH)
      GC TO 120
110   WRITE (1,1210)
1210  FORMAT(5HO   *,25X,7HNO DATA,/,5H    -,23X,10HBELOW  LOW)
      WRITE (6,1220) (SYMTAB(I),DRANG(I),DRANG(I+1),I=1,CON)
1220  FORMAT (4X,A1,2X,F12.4,6H  TO  ,F12.4)
      WRITE (6,1230) CRANG(CON+1)
1230  FCRMAT (4X,1H+,2X,F12.4,10H  OR  OVER)
120   WRITE(6,1030) TITL,N,COLS,ROWS
      XSTOP = GRIC
      GRID=COLS-1
      CCL=3
      DO 140 J=1,GRID
      DO 130 I = 2,XSTOP
      SYM(COL)=BLANK
130   CCL=CCL+1
      SYM(CCL)=PLUS
140   CCL=COL+1
      N=COLMAX+2
      SYM(2)=PLUS
      SYM(1)=BLANK
```

```
      GO TO 208
207   PSYM(COL)=STAR
208   TOUT=TOUT+1
      ZPREV(COL)=ZEST
      ZLAST=ZEST
      GO TO 480
325    CONTINUE
       IF(TCUP.NE.2) GO TO 312
       DO 311 LMN=2,JJJ
       I=LMN-1
       IF(ZEST.LT.DRANG(LMN)) GO TO 313
311    CONTINUE
312    I=((ZEST-ZMIN)/CI)+1.
313    TABE(I)=TABE(I)+1.
      LVAL=LINES+1
      IF(ITOP.EQ.1.OR.MM.EQ.IROW.OR.JLFT.EQ.1)LVAL=3
      GO TO (326,327,328) ,LVAL
326   DO 309 I=1,CON
      ZSTEP=DRANG(I)
      SW1=1.
      SW2=1.
      SW3=1.
      SW4=1.
      IF (ZSTEP.GE.ZLAST.OR.ZSTEP.GE.ZPREV(COL)) SW1=0
      IF (ZSTEP.LE.ZEST) SW2=0.
      IF (ZSTEP.LT.ZLAST.OR.ZSTEP.LT.ZPREV(COL)) SW3=0.
      IF (ZSTEP.GE.ZEST) SW4=0.
      IF((SW1.EQ.0..AND.SW2.EQ.0.).OR.(SW3.EQ.0..AND.SW4.EQ.0.))GO TO310
309   CONTINUE
      PSYM(COL)=BLANK
      GO TO 320
310   PSYM(COL)=SYMTAB(I)
320   ZLAST=ZEST
      ZPREV(COL)=ZEST
      GO TO 124
327   K=I/2
      IF((K*2).NE.I) PSYM(COL)=SYMTAB(I)
      IF((K*2).EQ.I) PSYM(COL)=BLANK
      GO TO 124
328   PSYM(COL)=SYMTAB(I)
      ZPREV(COL)=ZEST
124   SUM=SUM+1.0
  480 XP=XP+XINC
      Z1=Z2
330   JLFT=0
      COL = COL + 1
      IF(Z(M,COLS).GE.ZMIN.AND.Z(M,COLS).LE.ZMAX) GO TO 333
      IF(Z(M,COLS).GT.ZMAX) GO TO 335
      IF(Z(M,COLS).LE.RLOW) GO TO 336
      PSYM(COL)=MINUS
      TLOW=TLOW+1
      GO TO 334
335   IF(Z(M,COLS).GE.RHIGH) GO TO 336
      PSYM(COL)=PLUS
      TABE(CON+1)=TABE(CON+1)+1.0
      GO TO 334
```

PLOTTING OF BIVARIATE STANDARD DEVIATIONS

Purpose: This program may be used to provide some summary measures of geographical distributions.

Description: The program first calculates the weighted mean center of the input point distribution:

$$\bar{X} = \frac{\sum X_i W_i}{\sum W_i} \quad \text{and} \quad \bar{Y} = \frac{\sum Y_i W_i}{\sum W_i}$$

where $\sum W_i$ is the sum of weights (all weights may be one) and X_i and Y_i are the coordinates of the respective points. The variance, standard deviation, and standard distance are then calculated in the normal statistical manner. The eigenvalues of the covariance matrix are the squares of the semi-major and semi-minor axes of the standard deviational ellipse. The counterclockwise angle of inclination of the major axis from East-West is given, as well as the total area of the ellipse. The coefficient of circularity (the ratio of the minor axis to the major axis) is also calculated. The ellipse is then plotted; dots are used for the center and end points of the principle axes. The data are plotted as circles with radii given by $r = aW^b$, where r is the radius of the circle, W is the weight while a and b are empirically defined coefficients. For the United States if W represents population then a = 0.0219 and b = 0.44, and these are the values used in the program.

Comments: The program is currently dimensioned to plot up to 2000 observations per run. This may be changed by altering the value of X in the dimension statement. Multiple data sets may be run by repeating the data deck. The program reads the observation deck from unit 4, reads controls from 5, writes onto 6, and plots on 9. Output is planned for 30" CALCOMP plotter and standard CALCOMP plotting library is assumed. The origin for the X,Y coordinates is, as usual, defined to be the Southwest (lower left) of the study area. This program is designed so that it may be easily used at a remote terminal which has access to shared files.

Data Deck: 1) Control Card

2) Title Card

3) Variable Format Card

4) Observation Deck

Data Card Composition: 1) Control Card

column

1-4	NO	Number of observation (≤2000) I4
5-11	XMIN	Minimum X value (F7.2)
12-18	XMAX	Maximum X value (F7.2)
19-25	YMIN	Minimum Y value (F7.2)
26-32	YMAX	Maximum Y value (F7.2)
33-39	SCAL	Denominator of the representative fraction of the map scale (F12.0)

2) Title Card

Any title (≤72 characters)

3) Variable Format Card

E or F type FORTRAN format for one observation.

4) Observation Deck

Observations as described by the variable format card.

The input data must be: X coordinate, Y coordinate, Weight.

Programmer: Waldo R. Tobler

References: Lefever, D. Weltz, "Measuring Geographic Concentration by Means of the Standard Deviational Ellipse," American Journal of Sociology, Vol. 32, #1, July 1926, pp. 88-94.

Lee, Douglass B., Jr., Analysis of Residential Segregation, Division of Urban Studies, Center for Housing and Environmental Studies, Cornell University, Ithaca, New York, 1966.

```
      DO 3 I=1,NO
      READ(4,FMT) X(1,I),X(2,I),X(3,I)
      IF(X(3,I)) 5,5,6
5     X(3,I)=1.0
6     S(1)=S(1)+X(1,I)*X(3,I)
      S(2)=S(2)+X(2,I)*X(3,I)
      S(3)=S(3)+X(3,I)
      WRITE(6,1004) I,X(1,I),X(2,I),X(3,I)
3     CONTINUE
      B(1)=S(1)/S(3)
      B(2)=S(2)/S(3)
      B(3)=S(3)/FNO
      SXY=0.0
      DO 7 I=1,NO
      W=X(3,I)
      DO 8 J=1,2
      DEV(J)=X(J,I)-B(J)
      SX(J)=SX(J)+DEV(J)*W
8     SX2(J)=SX2(J)+DEV(J)*DEV(J)*W
7     SXY=SXY+DEV(1)*DEV(2)*W
      SX2(3)=SXY
      WRITE(6,1005)
      DO 9 I=1,3
9     WRITE (6,1006) I,S(I),B(I),SX2(I)
      WRITE(6,1007)
      DO 10 I=1,3
      S2(I)=SX2(I)/S(3)
      S(I)=SQRT(ABS(S2(I)))
10    WRITE(6,1008) I,S2(I),S(I)
      BB=S2(1)+S2(2)
      SDIS=SQRT(BB)
      WRITE(6,1009) SDIS
      C=-(S2(1)*S2(2)-S2(3)*S2(3))
      ROOT=SQRT(ABS(BB*BB+C*4.0))
      E1=0.5*(BB+ROOT)
      E2=0.5*(BB-ROOT)
      WRITE(6,1010)
      WRITE(6,1008) NO,E1,E2
      E1=SQRT(E1)
      E2=SQRT(E2)
      WRITE (6,1008) NO, E1,E2
      AREA=E1*E2*FI
      DEN=SX2(1)-SX2(2)
      ANG=0.5*ATAN(2.0*SX2(3)/DEN)
      IF(DEN)30,30,32
30    ANG=ANG+FI2
32    WRITE(6,1016)
      IF(ANG-FI) 34,34,33
33    ANG=ANG-FI
34    ANG2=ANG+FI2
C CALCULATE END POINTS
      XA=B(1)+E1*COS(ANG)
      YA=B(2)+E1*SIN(ANG)
      WRITE(6,1008) NO,XA,YA
      XB=B(1)+E2*COS(ANG2)
      YB=B(2)+E2*SIN(ANG2)
```

PELTO'S D-FUNCTION AND RELATIVE ENTROPY

Purpose: This program calculates two measures of the degree of mixing in multicomponent systems.

Description: If one assumes that each observation consists of V variables, then the procedure is restricted to those situations in which it is meaningful to make the assumption that the totality of variables, for each observation, sum to 100 percent.

The D-function measures the strength of dominance of variables of components; the lower the D value (and closer to the theoretical minimum), the stronger the dominance. The minimum of the D-function is 100 x (1 - (1/CLASS)). The relative entropy measure may be taken to be an indicator of mixing of the components and would be 100 when all components appear in equal proportions. Since entropy may be taken as a measure of mixing (uncertainty), zones of high entropy should form boundaries between regions, whereas zones of low entropy should form regional areas. Relative entropy (H_r) is calculated by the following formula and expressed as a percent:

$$100\ H_r = \frac{-100\ P_i \log P_i}{H_m}$$

where P_i are the individual proportions and H_m is the theoretical maximum for the entropy. This of course assumes that the probability of occurance is the same for all components.

Comments: The program will handle up to 999 observations for each set of data. Since each observation is processed individually, data sets larger than 999 may be processed by splitting the set and run separately. Many sets of data may be run by repeating the data deck. As written the program will process observations of up to 20 components. This may be expanded by increasing the size of the arrays of D, P, MEMBER, RD, and PD. Their sizes should always be the number of components + 5. The components need not be in proportions; these will be calculated by the program. The results for each observation may be punched out if desired. Input data may include X,Y coordinates for subsequent map plotting. The order of input data is X coordinate, Y coordinate (if coordinates are specified), then the components 1 through NCOMP.

Data Deck:
1) Control Card
2) Variable format card
3) Observation Deck

Data Card Composition:
1) <u>Control Card</u>

column		
1-3	NOBS	Number of observations
4-5	NCOMP	Number of components (up to 20)
6	PRO	One (1) if proportions have not yet been calculated, zero (0) otherwise (if in doubt set equal to one).
7	PUN	One (1) if results are to be punched, zero (0) otherwise.
8-9	COORD	If X,Y coordinates are to be read and punched for output in other programs

they must be scaled to fit an output format of F6.2. Enter power of ten to scale coordinates appropriately: (-9 to +9) otherwise leave blank (no coordinates read).

10-69 any title

2) Variable Format Card

E or F type FORTRAN for each observation ($\leq$ 80 characters)

3) Observation Deck

The observations as described in the format.

Reference: The program is a verbatim translation of: C.R. Pelto, "Mapping of Multicomponent Systems," Journal of Geology, 62 (1954), pp. 501-511; Bobbs Merrill Geography reprint #G-178.

```
   25 WRITE (6,98) NO,(P(I),I=1,NCOMP)
C CALCULATE PROPORTIONS
      IF(PRO.NE.1) GO TO 40
      SUM = 0.
      DO 30 I=1,NCOMP
   30 SUM = SUM + P(I)
      DO 35 I=1,NCOMP
   35 P(I) =P(I)/SUM
C RANK PROPORTIONS FROM GREATEST TO LEAST
   40 DO 45 I=1,NCOMP2
   45 D(I) =-1.0
      MEMBER (1)=1
      P(NCOMP+1) =0.0
      LOOK = NCOMP +1
      DO 60 I=1,LOOK
      DO 55 K=1,I
      IF(P(I).LE.D(K)) GO TO 55
      L=I-K+1
      DO 50 JL=1,L
      J=I-JL+1
      M=J+1
      D(M)=D(J)
   50 MEMBER (M)=MEMBER(J)
      D(K)= P(I)
      MEMBER(K)=I
      GO TO 60
   55 CONTINUE
   60 CONTINUE
C CALCULATE DIFFERENCES
      DO 65 I=1,NCOMP
      L=I+1
   65 DP(I) =ABS(P(MEMBER(I))-P(MEMBER(L)))
C RANK DIFFERENCES
      DO 70 I=1,NCOMP2
   70 D(I)=-1.0
      RD(1)=1
      SUM=0.0
C CALCULATE RELATIVE ENTROPY
      DO 85 I=1,NCOMP
      IF (P(I).GT.0.001) SUM=SUM+P(I)*ALOG(P(I))
      DO 80 K=1,I
      IF (DP(I).LE.D(K))  GO TO 80
      L=I-K+1
      DO 75 JL=1,L
      J=I-JL+1
      M=J+1
      D(M)=D(J)
   75 RD(M)=RD(J)
      D(K)=DP(I)
      RD(K)=I
      GO TO 85
   80 CONTINUE
   85 CONTINUE
       ENTROP = 100.0*SUM/DENOM
      DF = 100.0*(1.0-(DP(RD(1))-DP(RD(2))))
      CLASS = RD(1)
```

POPULATION MAPS

Purpose: The program reads rectangular coordinates (x,y) and populations, and then draws population maps, with cities represented as circles.

Description: For cities it has been shown that the built up area is extremely well approximated by a circle of radius $r = aP^b$. For the United States a = 0.0219, b = 0.44 when the radius is in miles. These values are used as the default options in the program. If the population is zero, or if the radius is less than 0.01 inch at the plotting scale, the program produces a dot. Up to 9999 circles may be plotted on a single map.

Comments: Input/output--reads from unit 6, writes on 7, plots on 9. Output is planned for 30" CALCOMP plotter and the standard CALCOMP plotting library is assumed. X, Y, XMIN, YMIN, DV are assumed to be in miles; if in inches set SCAL = 63360. If ALPHA and BETA (see below) are defaulted, both must be defaulted; conversely if one is supplied both must be supplied.

Data Deck:
1) Control card
2) Title card
3) Variable format card
4) Observation deck

Data Card Composition: 1) Control card

columns

1-4	NOBS	Number of observations (≤ 9999) (I4)
5-9	XMIN	Minimum value of the x coordinate (F5.0)
10-14	YMIN	Minimum value of the y coordinate (F5.0)
15-24	SCAL	Denominator of the representative fraction (F10.0)
25-27	DV	Spacing of tick marks. (F3.0)
28-33	ALPHA	Coefficient of proportionality, a (F6.0) Default = 0.0219.
34-39	BETA	Exponent, b (F6.0) Default = 0.44

2) Title card

Any title (< 32 characters)

3) Variable format card

E or F FORTRAN type format for one observation.

4) Observation deck

Observations as described by the variable format card.

Programmer: Waldo R. Tobler

Reference: W.R. Tobler, "Satellite Confirmation of Settlement Size Coefficients," Area, Vol. 1, No. 3 (1969), pp. 30-34.

```
C       PROGRAM TO DRAW POPULATION MAPS
        INTEGER T
        DIMENSION TIT(8),FMT(18)
10      FORMAT (I4,2F5.0,F10.0,F3.0,2F6.0)
11      FORMAT (8A4)
12      FORMAT (7H1TIT = ,8A4)
14      FORMAT (8H0ALPH = ,F6.0/8H BETA = ,F6.0)
13      FORMAT (18A4)
15      FORMAT (10H0SUMPOP = ,F15.5)
        MILE = 63360.
        T=0
2       READ (6,10) NOBS,XMIN,YMIN,SCAL,DV,ALPH,BETA
        IF(ALPH.GT.0.0) GO TO 3
        ALPH = 0.0219
        BETA = 0.44
3       READ (6,11) TIT
        WRITE (7,12) TIT
        WRITE (7,14) ALPH,BETA
        PX = 1.0
        PY = 1.0
        SIZE = 0.1
        THO = 0.
        CALL PSYMB(PX,PY,SIZE,'GEOGRAPHY DEPARTMENT UNIVERSITY OF MICHIGAN
       1  1968 ',THO,50)
        PX = PX+6.0
        CALL PSYMB(PX,PY,SIZE,TIT,THO,32)
        SCAL = SCAL/MILE
        SUMPOP = 0.0
        CALL PLTOFS(XMIN,SCAL,YMIN,SCAL,1.0,3.0)
        XMAX = XMIN
        READ (6,13) FMT
        THF = 360.
        K=0
        DL = 0.
        SC = 10.
4       K = K+1
        READ (6,FMT) X,Y,POP
        RHO = ALPH*(POP**BETA)
        XMAX = AMAX1(X,XMAX)
        IF((RHO/SCAL).LT.0.01) RHO = 0.01*SCAL
        CALL PCIRCL(X,Y,THO,THF,RHO,RHO,DL,SC)
        SUMPOP = SUMPOP+POP
        IF(K.LT.NOBS) GO TO 4
        DV = DV/SCAL
        AXLTH = -(XMAX-XMIN)/SCAL
        CALL PAXIS(1.0,2.0,T,-0,AXLTH,THO,XMIN,SCAL,DV)
        WRITE (7,15) SUMPOP
        CALL PLTEND
        GO TO 2
        END
```

SPHERICAL DISTANCES

Purpose: To compute distances in kilometers assuming one degree equals 69.172 miles.

Description: The program reads a vector of latitude/longitude coordinates and produces the complete distance matrix; i.e., the distance between all pairs of points specified in the coordinate vector. Latitudes and longitudes are assumed given in degrees and minutes, with the usual convention that South latitudes and West longitudes are negative.

Deck Make-up: 1) Control and Format Card

2) Observations deck

Card Format: Control and format Card

columns

1-3	Number of observations ≤ 100
4-67	Fortran format for one observation

Observations deck (all E or F fields)

latitude in whole degrees

minutes of latitude

longitude in whole degrees

minutes of longitude

References: W.R. Tobler, "A Comparison of Spherical and Ellipsoidal Measures", Professional Geographer, XVI, 4(1964), pp. 9-12.

W.R. Tobler, Geographical Coordinate Computations", University of Michigan, Contract Nonr 1224 (48), Task 389-137; UM 05824-2,3-T, December, 1964, 107 pages.

COORDINATES FROM DISTANCES

Purpose: The program reads a matrix of distances, and then produces a vector of plane coordinates.

Deck Make-up: 1) Title Card

2) Control Card

3) Distance Matrix

Data Cards: <u>Title Card</u> (18A4)

any title

<u>Control Card</u> (2I2, I1, 1X, 16A4)

column

1-2	NRANDC	Number of rows and columns in the distance matrix.
3-4	NREIGN	Number of eigenvectors desired
5	IFPLOT	zero to plot the solutions.
7-70	FMT	Format for one row of the distance matrix.

<u>Distance Matrix</u>

The distance matrix as described by the format. Only the lower half is read.

Reference: W.S. Torgerson, <u>Theory and Methods of Scaling</u>, New York, J. Wiley, 1958, pp. 254-259.

J. Lingoes, "An IBM 7090 Program for Guttman- Lingoes Smallest Space Analysis", <u>Behavioral Science</u>, 10(1965), pp. 183-184.

W. Tobler, H. Mielke, T. Detwyler, "Geobotanical Distance Between New Zealand and Neighboring Islands", <u>Bioscience</u>, May 1970.

```
C
C
C   M A P   C O O R D I N A T E S   F R O M   G E O - D I S T A N C E S
C
C PROGRAM TO COMPUTE EUCLIDEAN COORDINATES FROM GIVEN DISTANCES
C WALDO R. TOBLER             1967
C DEPARTMENT OF GEOGRAPHY UNIVERSITY OF MICHIGAN
C HOUSEHOLDER AND YOUNG THEOREM - TORGERSONS METHOD
C FOR METRIC MULTI-DIMENSIONAL SCALING
C
C       INPUT
C READS 7 WRITES 5 AND 6
C READS TITLE, CONTROLS + FORMAT, THEN LOWER-HALF DISTANCE MATRIX
C       CARD1 TITLE (1-72)
C       CARD # 2
C
C       NRANDC (1-2) ROWS AND COLUMNS
C       NREIGN (3-4) COLS OF EIGEN (<11)
C       IFPLOT (5)   =0 FOR PLOT
C       FORMAT (7-72)
C       CARD # 3 AND ON = LOWER HALF OF DISTANCE MATRIX
        DIMENSION A(70,70), B(70,70), C(70,70), D(70,70), FMT(16), TITL(18
       1), DUMYA(70), EVALU(70), X1(70,2), RSUM(70), CSUM(70)
        EQUIVALENCE (DUMYA,EVALU), (NR,NRANDC), (NC,NREIGN)
        EQUIVALENCE (A,B)
        MD=70
C       READ TITLE
1       READ (7,37) (TITL(I),I=1,18)
        READ(7,38) NRANDC,NREIGN,IFPLOT,(FMT(I),I=1,16)
        IFSPM=0
C       OUTPUT OF INPUT
        WRITE (6,39) (TITL(I),I=1,18),(FMT(I),I=1,16)
        WRITE (6,40)
C       READ AND PRINT MATRIX
        IF ((IFSPM)) 2,3,2
2       WRITE (6,41)
        GO TO 4
3       WRITE (6,42)
4       DO 7 I=1,NRANDC
        IF (I-NRANDC) 5,6,6
5       IP1=I+1
        READ (7,FMT) (A(I,J),J=1,I),(DUMYA(J),J=IP1,NRANDC)
        GO TO 7
6       READ (7,FMT) (A(I,J),J=1,NRANDC)
 7      WRITE(5,43) I,(A(I,J),J=1,I)
        DO 8 I=2,NRANDC
        IM1=I-1
        DO 8 J=1,IM1
8       A(J,I)=A(I,J)
C       CHANGES TO DISTANCE MATRIX CAN BE PERFORMED HERE
C       (I.E., GET GEOBOTANICAL DISTANCES FROM SPECIES SIMILARITIES, ETC.)
        IF (IFSPM) 9,11,9
9       DO 10 I=1,NR
        DO 10 J=1,NR
10      C(I,J)=A(I,J)
```

```
25      CONTINUE
26      CONTINUE
        RMULT=F/SUM
C       PROGRAM NOW SCALES FIRST TWO EIGENVECTORS FOR GEO-MAPS
        DO 27 I=1,NR
        DO 27 J=1,2
27      D(I,J)=D(I,J)*RMULT
        WRITE (6,45) RMULT
C       PRINT OUT EIGENVECTORS
        WRITE (6,39) (TITL(I),I=1,18)
        WRITE (6,46) NRANCC,NREIGN
        DO 28 I=1,NRANCC
28      WRITE (6,47) I,(D(I,J),J=1,NREIGN)
C       PRINT RESULTING DISTANCES, USING FIRST TWO EIGENVECTORS
        WRITE (6,39) (TITL(I),I=1,18)
        WRITE (6,48)
        DO 31 I=1,NR
        IM1=I
        DO 30 J=1,IM1
        IM2=J
        IF (IM1-IM2) 30,29,30
29      B(I,J)=0.0
        GO TO 31
30      B(I,J)=SQRT((D(I,1)-D(J,1))**2+(D(I,2)-D(J,2))**2)
 31     WRITE(5,43) I,(B(I,J),J=1,IM1)
C       NORMALIZE FOR PLOTS
        IF (IFPLOT) 32,32,36
32      S=0
        DO 34 K=1,NC
        DO 34 I=1,NR
        IF (S-ABS (D(I,K))) 33,34,34
33      S=ABS (D(I,K))
34      CONTINUE
        DO 35 K=1,NC
        DO 35 I=1,NR
35      D(I,K)=D(I,K)/S
C       PLOT OUT THE VECTORS OF D
        CALL PLOT(D,X1(1,1),X1(1,2),NR,NC,MD)
36      GO TO 1
C
37      FORMAT (18A4)
38      FORMAT(2I2,I1,1X,16A4)
39      FORMAT('1EIGENVALUE AND EIGENVECTOR PROGRAM',/1H ,18A4/5X,18A4)
40      FORMAT('LOWER HALF OF DATA MATRIX')
41      FORMAT('DATA IS SCALER PRODUCT MATRIX')
42      FORMAT('DATA IS RAW DISTANCE MATRIX')
43      FORMAT (1H0,I3,10E12.5/(1H ,3X,10E12.5/))
44      FORMAT (1H0,I2,F17.2,2PF24.2)
45      FORMAT('OMULTIPLIER IS ', F11.5)
46      FORMAT(' EIGENVECTORS'/'ROWS=',I10,10X,'COLUMNS=',I10)
47      FORMAT (1H0,I2,2X,10F11.3)
48      FORMAT('RESULTING DISTANCES')
49      FORMAT('- E I G E N V A L U E S ' /)
50      FORMAT( ' TRACE= ',E12.5, 10X,'N= ',I10, /)
51      FORMAT(' ROWS',10X,'VALUE',10X,'CUM. PERCENT OF TOTAL',/1H0 )
        END
```

```
9       CONTINUE
        SCALAR=SCALAR/2.
        DO 10 J=I1,N
        QVEC(J)=PVEC(J)-SCALAR*WVEC(J)
        DO 10 K=I1,J
        A(K,J)=A(K,J)-(WVEC(K)*QVEC(J)+WVEC(J)*QVEC(K))
        A(J,K)=A(K,J)
10      CONTINUE
C       SAVE ROTATION FOR LATER APPIICATION TO CO-DIAGONAL VECTORS
        DO 12 K=2,N
        TEMP=0.
        DO 11 J=I1,N
11      TEMP=TEMP+WVEC(J)*B(J,K)
        DO 12 J=I1,N
        B(J,K)=B(J,K)-WVEC(J)*TEMP
12      CONTINUE
C       MOVE CO-DIAGONAL FORM ELEMENTS FOR ITERATIVE PROCEDURE
13      J=I
        DIAG(I)=A(J,I)
        SUPERD(I)=A(J+1,I)
14      CONTINUE
        DIAG(N)=A(N,N)
C       DETERMINE EIGENVALUES FROM STURM CHAIN OF CO-DIAGONAL MINORS
C       CALCULATE NORM OF MATRIX AND INITIALIZE EIGENVALUE BOUNDS
        ANORM2=DIAG(1)**2
        DO 15 L=2,N
        Q(L-1)=SUPERD(L-1)**2
        ANORM2=DIAG(L)**2+Q(L-1)+Q(L-1)+ANORM2
15      CONTINUE
C                                                    TERMINAL-F FLAG
        ANORM=SQRT(ANORM2)
        DO 16 L=1,M
        VALU(L)=ANORM
        VALL(L)=-ANORM
16      CONTINUE
        EPS1=ANORM*E1
        IF (EPS1) 17,73,17
C       CHOOSE NEW TRIAL VALUE WHILE TESTING BOUNDS FOR CONVERGENCE
17      DO 35 L=1,M
        ITER=0
        VTEMP=EPS1
18      TAU=(VALU(L)+VALL(L))/2.
        IF (ITER-10)  20,19,20
19      VTEMP=VTEMP*10.
        ITER=0
20      IF (2.*(TAU-VALL(L))-VTEMP) 35,35,21
C       DETERMINE SIGNS OF PRINCIPAL MINORS
21      MATCH=0
        ITER=ITER+1
        T2=0.
        T1=1.
        DO 30 L1=1,N
        P=DIAG(L1)-TAU
        IF (T2) 23,22,23
22      T1=SIGN (1.,T1)
23      IF (T1) 25,24,25
```

```
      T(J+1,K)=VTEMP
48    CONTINUE
49    VTEMP=T(J+1,1)/T(J,1)
      U(J)=VTEMP
      T(J+1,2)=T(J+1,2)-VTEMP*T(J.2)
      T(J+1,3)=T(J+1,3)-VTEMP*T(J.3)
50    CONTINUE
      ITER=1
      IF (I1) 58,51,58
C     BACK SUBSTITUTE TO OBTAIN EIGENVECTOR
51    DO 52 L1=1,N
      L=NP1-L1
      V(L)=(V(L)-T(L,2)*V(L+1)-T(L,3)*V(L+2))/T(L,1)
52    CONTINUE
      GO TO (53,58), ITER
C     PERFORM SECOND ITERATION
53    ITER=2
54    DO 57 L=2,N
      IF (INDEX(L-1))55,56,55
55    VTEMP=V(L-1)
      V(L-1)=V(L)
      V(L)=VTEMP
56    V(L)=V(L)-U(L-1)*V(L-1)
57    CONTINUE
      GO TO 51
C     ORTHOGONALIZE VECTOR TO OTHERS ASSOCIATED WITH REPEATED ROOT
58    IF (I1) 59,62,59
59    DO 61 L1=1,I1
      K=I-L1
      VTEMP=0.
      DO 60 J=1,N
60    VTEMP=VTEMP+A(J,K)*V(J)
      DO 61 J=1,N
61    V(J)=V(J)-A(J,K)*VTEMP
62    GO TO (54,63), ITER
C     NORMALIZE VECTOR TO UNIT LENGTH
63    VNORM2=0.
      SUM=0.
      DO 65 L=1,N
      IF (SUM-ABS (V(L))) 64,65,65
64    SUM=ABS (V(L))
65    CONTINUE
      DO 66 L=1,N
      V(L)=V(L)/SUM
66    VNORM2=VNORM2+V(L)**2
C                                                           TERMINAL-F FLAG
      VNORM=SQRT(VNORM2)
      DO 67 J=1,N
67    A(J,I)=V(J)/VNORM
68    CONTINUE
C     ROTATION OF CO-DIAGONAL VECTORS INTO MATRIX EIGENVECTORS
      DO 70 I=1,M
      DO 69 K=2,N
      U(K)=0.
      DO 69 J=2,N
69    U(K)=U(K)+B(J,K)*A(J,I)
```

```
      L=K+5
      IF (IW(K).EQ.BLANK) GO TO 12
      FMT(L)=WORD2
      GO TO 13
 12   FMT(L)=WORD1
 13   CONTINUE
      DO 15 K=26,50
      L=K+6
      IF (IW(K).EQ.BLANK) GO TO 14
      FMT(L)=WORD2
      GO TO 15
 14   FMT(L)=WORD1
 15   CONTINUE
      IF (KX-26) 17,16,17
 16   WRITE (JTAPE,20) I
 17   WRITE (JTAPE,FMT) L2(KX),(IW(L),L=1,50),L2(KX)
 18   WRITE (JTAPE,21) (L1(K),K=1,20)
      RETURN
C     *** FORMAT STATEMENTS ***
C
 19   FORMAT (1H1,12HVECTOR PLOTS/1H ,6HVECTOR,I5,1X22HPLOTTED AGAINST V
     1ECTOR,I5,21X6HVECTOR/1H ,60XI4/1H0,11X10I5,2X1H*,I4,9I5/1H ,13X50(
     22H* ),1H*)
 20   FORMAT (1H ,6HVECTOR,I3,4X50(2H* ),1H*)
 21   FORMAT (1H ,13X50(2H* ),1H*/1H ,11X10I5,2X1H*,I4,9I5)
      END
```

DATA PLOTTING

Purpose: Rapid data screening.

Description: The program produces the N*(N - 1)/2 printer plots, linear regressions, and simple correlations for all pairwise combinations of N variables. All operations are done according to single precision standard statistical procedures.

Comments: Current dimensioning allows 2000 observations for each of 30 variables. I/O reads controls and format from unit 7, reads observations from 4, and writes onto 6. Multiple observation decks may be run by repeating the data deck. The UMPLOT subroutines are available through SHARE. The end points of the linear regression are indicated by dots in the margin of the plots.

Data deck:
1) Title card
2) Control card
3) Format card
4) Observation Deck
5) End of observations card(s)

Data Card Composition:
1) Title Card

   Any title ($\leq$ 72 characters)

2) Control Card

   column

   1-2 NVAR The number of variables ($\leq$ 15) I2

3-4	INDEP	Independent variable indicator (i.e., 5 yields 5th variable on x axis). Default = last variable.
5-6	IDEP	Dependent variable indicator (y axis). Default = 1

3) Format Card

E or F type FORTRAN format for one observation of the N variables. ($\leq$72 characters)

4) Observation Deck

Observations as described by the FORMAT

5) "End of observations" card(s)

Put 9999 in columns 1-4; this avoids having to know the exact number of observations. If the observations are on M cards, follow this last card by M-1 blank cards.

Programmer: Program originally written by Krumbein and Benson of Northwestern University. Revised by Tobler of the University of Michigan, 1962.

References: Comparable programs are now available in the BIMED system.

```
C PROGRAM TO PLOT SCATTER DIAGRAMS
      DIMENSION TRANG(30), VMIN(30), VMAX(30), RANG(30)
      DIMENSION V(30,2000),             X(2000),Y(2000)
      DIMENSION TITL(30), FMT(21), GRAPH(1200), DELX(30), DELY(30), DELX
     1X(30), DELYY(30), SUMX(30), SUMXQ(30), SUMCOV(30,30), SIGX(30)
      DIMENSION XBAR(30), SSX(30), R(30,30), B(30,30), A(30,30), RSQ(30,
     130)
      DIMENSION W(30)
      DATA FMT/'(A4,T1, ',18*'     ',' )  '/,HALT/'9999'/
C     *****SET UP TABLE OF RANGES
      TRANG(1)=0.01
      DO 1 I=2,10
      L=I-1
    1 TRANG(I)=10.0*TRANG(L)
      NSCALE=0
      NHL=6
      NSBH=8
      NVL=8
      NSBV=10
C     *****READ TITLES AND MASTER
    2 READ (7,33) (TITL(I),I=1,18)
      WRITE (6,35)(TITL(I),I=1,18)
      READ (7,34)     NVAR,INDEP,IDEP
      WRITE (6,41) NVAR,INDEP,IDEP
      READ (7,33) (FMT(I),I=3,20)
      WRITE(6,33) (FMT(I),I=3,20)
      NVAR1 = NVAR
      IF (IDEP) 498,499,499
  498 IDEP = -IDEP
      NVAR1 = NVAR + NVAR
C     ** ZERO OUT SUMS
  499 DO 3 I =1,NVAR1
      SUMX(I)=0.0
      SUMXQ(I)=0.0
      DO 3 J =1,NVAR1
    3 SUMCOV(I,J)=0.0
      COUNT=0.0
C     *****READ AND STORE DATA MATRIX
      N=0
    4 N=N+1
      READ(4,FMT)        END,(V(I,N),I=1,NVAR)
      IF (NVAR-NVAR1) 501,505,505
  501 DO 500 K=1,NVAR
      JJ = K
      J = NVAR + JJ
  500 V(J,N)=V(JJ,N)**2
  505 IF(END.EQ.HALT) GO TO 11
    5 IF (INDEP) 7,7,6
    6 DUM1=V(NVAR,N)
      V(NVAR,N)=V(INDEP,N)
    7 V(INDEP,N)=DUM1
    8 IF (IDEP) 9,9,8
      DUM1=V(1,N)
      V(1,N)=V(IDEP,N)
      V(IDEP,N)=DUM1
```

```
9       CONTINUE
C       ** ACCUM SUMS
        DO 10 I =1,NVAR1
        SUMX(I)=SUMX(I)+V(I,N)
        SUMXQ(I)=SUMXQ(I)+V(I,N)**2
        DO 10 J =1,NVAR1
10      SUMCOV(I,J)=SUMCOV(I,J)+V(I,N)*V(J,N)
        COUNT=COUNT+1.0
        GO TO 4
C       *** END OF DATA
11      N=N-1
        NVAR=NVAR1
C       ** GET SAMPLE STATISTICS
        DO 12 I=1,NVAR
        XBAR(I)=SUMX(I)/COUNT
        SSX(I)=SUMXQ(I)-XBAR(I)*SUMX(I)
        VARX=SSX(I)/COUNT
12      SIGX(I)=SQRT(VARX)
        DO 13 I=1,NVAR
        DO 13 J=1,NVAR
        COVXX=SUMCOV(I,J)-XBAR(I)*SUMX(J)
        R(I,J)=(COVXX/(SIGX(I)*SIGX(J)))/COUNT
        B(I,J)=COVXX/SSX(J)
        A(I,J)=XBAR(I)-B(I,J)*XBAR(J)
13      RSQ(I,J)=R(I,J)**2*100.0
C       *****PRINT OUT DATA
        WRITE (6,35) (TITL(I),I=1,18)
        WRITE (6,36) NVAR
        DO 14 M=1,N
14      WRITE (6,37)       M ,(V(I,M),I=1,NVAR)
        WRITE (6,38)
        WRITE (6,37) N,(XBAR(I),I=1,NVAR)
        WRITE (6,37) N,(SIGX(I),I=1,NVAR)
C       ****FIND MIN AND MAX VALUES EACH VECTOR
        DO 24 I=1,NVAR
        VMAX(I)=V(I,1)
        VMIN(I)=V(I,1)
        DO 18 M=2,N
        IF (V(I,M)-VMAX(I)) 16,18,15
15      VMAX(I)=V(I,M)
        GO TO 18
16      IF (V(I,M)-VMIN(I)) 17,18,18
17      VMIN(I)=V(I,M)
18      CONTINUE
        RANG(I)=VMAX(I)-VMIN(I)
C       *****SELECT FROM TABLE A RANGE SLIGHTLY LARGER THAN ACTUAL RANGE,
C       *****FOR EACH VECTOR.
        DO 19 K=3,10
        M=K-2
        IF (RANG(I)-TRANG(K)) 20,19,19
19      CONTINUE
20      DO 21 K=1,N
21      V(I,K)=V(I,K)/TRANG(M)
        VMAX(I)=VMAX(I)/TRANG(M)
        VMIN(I)=VMIN(I)/TRANG(M)
        RANG(I)=RANG(I)/TRANG(M)
```

```
      W(I)=TRANG(M)
      DO 22 K=1,10
      VMARG=FLOAT(K)
      VMARG=VMARG*10.0
      XMIN=RANG(I)-VMARG
      IF (XMIN) 23,22,22
22    CONTINUE
C     *****SPREAD MAX + MIN(SPLIT DIFFERENCE) TO FIT NEW RANGE SELECTED
23    VMARG=-XMIN/2.0
      VMIN(I)=VMIN(I)-VMARG
      VMIN(I)=HFIX(VMIN(I))
      VMAX(I)=VMAX(I)+VMARG+0.5
      VMAX(I)=HFIX(VMAX(I))
      RANG(I)=VMAX(I)-VMIN(I)
C     *****CALC INCREMENTS EACH SPACE ALONG X AND Y AXIS
      DELX(I)=RANG(I)/77.0
      DELY(I)=RANG(I)/45.0
      DELXX(I)=RANG(I)/7.0
24    DELYY(I)=RANG(I)/5.0
C     ***SET UP GRAPHS EACH PAIR OF VECTORS
      DO 32 I=1,NVAR
C     ***ITH VECTOR BECOMES YVECTOR
      DO 25 M=1,N
25    Y(M)=V(I,M)
      YMIN=VMIN(I)
      YMAX=VMAX(I)
      DO 32 J=I,NVAR
      IF (I-J) 26,32,26
C     ***JTH VECTOR BECOMES X VECTOR
26    DO 27 M=1,N
27    X(M)=V(J,M)
      XMIN=VMIN(J)
      XMAX=VMAX(J)
C     ***SET UP GRID WITH AXES AND VALUES EACH XDIV + YDIV
      CALL PLOT1 (NSCALE,NHL,NSBH,NVL,NSBV)
      CALL PLOT2 (GRAPH,XMAX,XMIN,YMAX,YMIN)
C     ***DELETE GRID LINES FROM BODY OF GRAPH
      KX=4
      DO 28 K=1,KX
      YVAL=YMIN+FLOAT(K)*DELYY(I)
      DO 28 L=1,76
28    CALL PLOT3 (1H ,XMIN+FLOAT(L)*DELX(J),YVAL,1,4)
      KX=6
      DO 29 K=1,KX
      XVAL=XMIN+FLOAT(K)*DELXX(J)
      DO 29 L=1,44
29    CALL PLOT3 (1H ,XVAL,YMIN+FLOAT(L)*DELY(I),1,4)
C     ***PLOT END POINTS OF REGRESSION LINE IN MARGIN
      YVAL=(A(I,J)+B(I,J)*W(J)*XMIN)/W(I)
      CALL PLOT3 (1H.,XMIN,YVAL,1,4)
      YVAL=(A(I,J)+B(I,J)*W(J)*XMAX)/W(I)
      CALL PLOT3 (1H.,XMAX,YVAL,1,4)
      ABIJ=ABS(B(I,J))
      IF (ABIJ-0.000001) 31,31,30
30    XVAL=(YMIN*W(I)-A(I,J))/(B(I,J)*W(J))
      CALL PLOT3 (1H.,XVAL,YMIN,1,4)
```

```
      XVAL=(YMAX*W(I)-A(I,J))/(B(I,J)*W(J))
      CALL PLOT3 (1H.,XVAL,YMAX,1.4)
C     *****PLOT ALL DATA POINTS ITH AND JTH VECTORS
31    CALL PLOT3 (1H*,X(1),Y(1),N.4)
C     ***WRITE OUT GRAPH
      WRITE (6,39) I,J,A(I,J),B(I.J),R(I,J),PSQ(I,J),N
      CALL PLOT4 (25,25H                    ITH VECTOR)
      WRITE (6,40)
32    CONTINUE
      GO TO 2
C            W.R.TOBLER / UNIVERSITY OF MICHIGAN GEOGRAPHY / 1962
C
33    FORMAT (18A4)
   34 FORMAT(3I2)
35    FORMAT (1H1,18A4)
36    FORMAT (/10X,I2,16H VECTORS OF DATA/10X,18H------------------)
37    FORMAT (5X,I4,8F14.2/(9X,8F14.2))
38    FORMAT (//10X,44HN, MEANS, AND STANDARD DEVIATIONS ARE ..... /)
39    FORMAT (1H1,2HI=,I2,3X,2HJ=,I2,5X,5H( I =,F10.4,3H + ,F10.4,6H * J
     1 ),5X,7HR(I,J)=,F10.4,5X,10HPCT SSRED=,F10.4,1H,,5X,I6,1X,6HPOINTS
     2///)
40    FORMAT (/30X,10HJTH VECTOR)
   41 FORMAT('    NVAR=',I3,4X,'INDEP=',I3,4X,'IDEP=',I3)
      END
NORTON'S DATA FOR ITALY
 7 3 5
(7F5.1)
14.9 150.0145.5141.0157.0147.0147.2
.6   62.2 38.7 41.7 35.4 34.6 32.8
8.3  90.7 60.7 47.7 42.5 44.6 39.5
9.7  88.6 69.5 56.8 41.7 39.8 39.8
31.0 150.8142.0131.0133.0130.0139.2
2.6  41.1 44.6 25.7 26.3 28.7 25.2
21.4 104.088.0 79.0 75.0 81.7 78.5
3.9  74.7 54.1 44.2 38.1 34.3 33.7
0.5  72.8 46.1 41.1 41.7 44.6 35.4
2.0  53.2 47.2 35.8 29.4 27.9 28.8
17.0 98.0 93.0 98.3 79.5 83.0 81.1
14.8 72.5 64.6 58.9 48.7 51.9 50.8
10.9 74.3 49.2 39.8 42.2 39.4 37.3
2.2  53.1 42.5 40.3 34.3 34.0 33.2
0.5  56.0 61.0 67.5 51.6 46.1 47.1
9.2  122.8120.0116.0123.9144.6132.6
8.3  90.7 60.7 47.7 44.1 44.6 39.5
16.3 111.877.4 60.5 56.8 58.9 54.0
15.7 90.8 69.0 53.5 49.2 44.1 40.0
6.2  83.8 53.0 41.9 38.8 38.8 35.4
4.9  109.368.5 49.9 44.4 43.1 41.6
11.0 64.7 50.1 41.0 37.3 34.4 34.4
3.1  86.0 64.8 54.3 57.5 42.2 34.2
```

GEOGRAPHICAL INTERPOLATION

Purpose: Interpolation to a square lattice from measures given at scattered geographical (x,y) positions.

Description: The assumption is made that the data are a sample taken from a continuous scalar field, and that values at unobserved locations can be estimated from the observed values. The second assumption clearly requires information concerning the shape of the two-dimensional autocorrelation function, as is discussed by Heiskanen and Moritz; equivalently, the applicability of the sampling theorem in two-dimensions should be recognized.

More specifically, the program establishes a lattice and estimates a value at each matrix point by using a weighted average of the six nearest data points. The weights are the inverses of the squares of the distances of the data points from the lattice point (i.e., linear interpolation). This weighted average is then averaged with the value at the nearest observation point (implying an autocorrelation which has a large negative slope in the vicinity of the origin) to give the final estimate. Alternate procedures are alternate hypotheses about nature.

Deck Make up:
1) First Control Card
2) Second Control Card
3) Title Card
4) Observations
5) Boundary Card (optional)

Data Cards: <u>First Control Card</u>

column		
1-3	N	Number of observations
4-71	FMT	Format for the observations

Second Control Card

1-3	ROWS	Number of rows in the output matrix (I3)
4-6	COLS	Number of columns in the output matrix (I3)
7-10	GSIZE	size of lattice in coordinate units (F4.0)
11	BND	1 to read specific boundaries; default uses the observations to calculate the boundaries.
12	LIST	1 to list original observations 2 to list estimated values 3 to do both of the above.
13	IPUN	1 to punch matrix of resulting values on unit 7.
14-15	NMAP	Number of maps requested. This option requires that subroutine SCON (see program GEOFIT) be called at line 149 of the program. The current listing does not include this option.

Comment: GSIZE, COLS, ROWS (in that precedence order) establish the size of the matrix. All may be left blank and the program will use a value which yields approximately as many lattice points as there are data points.

Title Card

Any title in columns 1-72.

Observations

X, Y, Z in that order punched as described in the format in the first control card.

Boundary Card

Used iff BND = 1 on the second control card: Reads XMAX, XMIN, YMAX, YMIN as 4F10.0 fields. These define a subregion of the data for which the interpolation is desired.

Programmer: W.R. Tobler

References:

J.W. Goodman, Introduction to Fourier Optics, New York, McGraw-Hill, 1968, pp. 21-25.

W. Heiskanen, and H. Moritz, Physical Geodesy, San Francisco, Freeman, 1967 (Chapter 7).

H. Freeman, Discrete Time Systems, New York, J. Wiley, 1965, pp. 73-96.

B. Gilchrist, G. Cressman, "An Experiment in Objective Analysis", Tellus, 6, 4 (1954).

D. Shepard, " A Two Dimensional Interpolation Function for Irregularly-Spaced Data", Proceedings, 1968 ACM National Conference, pp. 517-524.

P. Switzer, "Reconstructing Patterns from Sample Data", Am. Math. Stat., 33(1967), pp. 138-154.

N. Wiener, Extrapolation, Interpolation, and Smoothing of Stationary Time Series, MIT Press, 1949.

```
C PROGRAM TO INTERPOLATE FROM SCATTERED VALUES TO A LATTICE
C READS NOBS+FORMAT, CONTROLS, TITLE, DATA (X, Y, Z)
C CONTROLS: (2I3,F4.0,3I1,I2) ARE (ROWS,COLS,GSIZE,BND,LIST,IPUN,NMAP)
C
C READS 5,WRITES 6, PUNCHES 7
      DIMENSION X(1000), Y(1000), W(1000), Z(100,100), D(20), FC
     1RM(10), TITL(12), A(6), B(6), C(6)
      DIMENSION FMT(17),NI(20)
      REAL MAXX,MINX,MAXY,MINY,MAXW,MINW,MAXZ,MINZ
      INTEGER TAPE,YES,BND,ROWS,COLS
       TAPE=7
  105 CONTINUE
      READ(5,99) N,(FMT(I),I=1,17)
 99   FORMAT(I3,17A4)
      READ(5,110) ROWS,COLS,GSIZE,BND,LIST,IPUN,NMAP
      READ(5,115) (TITL(I),I=1,12)
      DO 1 I=1,N
1     READ (5,FMT) X(I),Y(I),W(I)
      IF (LIST.LT.1.OR.LIST.GT.3) GO TO 5
      IF(LIST.NE.2) WRITE(6,120) (TITL(I),I=1,12)
    5 IF (IPUN.EQ.1) WRITE (7,125) (TITL(I),I=1,12)
      DO 10 I=1,N
   10 IF (LIST.EQ.1.OR.LIST.EQ.3) WRITE(6,130) I,X(I),Y(I),W(I)
      CALL STATS (N,1,1000,1,X,XBAR,DUM,XSIG,MAXX,MINX)
      CALL STATS (N,1,1000,1,Y,YBAR,DUM,YSIG,MAXY,MINY)
      CALL STATS (N,1,1000,1,W,WBAR,DUM,WSIG,MAXW,MINW)
      WRITE (6,135) (TITL(I),I=1,12),N
      IF (BND.NE.1) GO TO 15
      WRITE (6,140)
      READ (5,145) XMAX,XMIN,YMAX,YMIN
      GO TO 20
C
   15 WRITE (6,150)
      XMAX=MAXX
      XMIN=MINX
      YMAX=MAXY
      YMIN=MINY
   20 FN=MAXW-MINW
      WMAX=FN+MAXW
      WMIN=MINW-FN
      MAXZ=MAXW
      MINZ=MINW
      WRITE (6,155) NMAP,LIST,BND,TAPE
      M=0
      FN=FLOAT(N)
      DX=XMAX-XMIN
      DY=YMAX-YMIN
      AREA=DX*DY
      BEST=SQRT(AREA/FN)
      AVEDIS=BEST*1.07346
      WRITE (6,160) XBAR,XSIG,MAXX,MINX,YBAR,YSIG,MAXY,MINY,WBAR,WSIG,MA
     1XW,MINW,AREA,AVEDIS,BEST,GSIZE,ROWS,COLS
      IF (GSIZE.LE.0) GO TO 25
      COLS=(2.0+DX/GSIZE)
      ROWS=(2.0+DY/GSIZE)
```

```
      M=1
      GO TO 45
C
   25 IF (COLS.LE.1) GO TO 35
   30 GSIZE=DX/FLOAT(COLS-1)
      ROWS=(2.0+DY/GSIZE)
      M=1
      GO TO 45
C
35    IF (ROWS.LE.1) GO TO 40
      GSIZE=DY/FLOAT(ROWS-1)
      COLS=(2.0+DX/GSIZE)
      M=1
      GO TO 45
C
   40 GSIZE=BEST
      COLS=(2.0+DX/GSIZE)
      ROWS=(2.0+DY/GSIZE)
   45 CONTINUE
      IF (COLS.LE.100) GO TO 50
      COLS=100
      GO TO 30
C
   50 WRITE (6,165)
      IF (M.EQ.1) WRITE (6,170)
      IF (M.NE.1) WRITE (6,175)
      WRITE (6,180) XMAX,XMIN,YMAX,YMIN,GSIZE,ROWS,COLS
      IF (LIST.EQ.2.OR.LIST.EQ.3) WRITE (6,185) (TITL(I),I=1,12)
      DY=(DY-(ROWS-1.)*GSIZE)/2.0
      DX=(DX-(COLS-1.)*GSIZE)/2.0
      YMAX=YMAX-DY
      XMIN=XMIN+DX
      HGRID=GSIZE/25.0
      FN=(WMAX-MAXW)/2.0
      MAXX=MAXW+FN
      MINX=MINW-FN
       WRITE(6,199)
      DO 95 M=1,ROWS
         YP=YMAX-(FLOAT(M-1))*GSIZE
         DO 90 J=1,COLS
            XP=XMIN+(FLOAT(J-1))*GSIZE
            DO 55 I=1,9
   55       D(I)=1.E10
            N1(1)=1
            N1(2)=1
            DO 75 I=1,N
               DIST=((XP-X(I))**2)+((YP-Y(I))**2)
               IF (DIST.GE.HGRID) GO TO 60
               Z(M,J)=W(I)
               GO TO 85
C
60    K=I
      IF(K-8) 61,61,62
61    YES=K
      GO TO 63
62    YES=8
```

```
63      DO 70 K=1,YES
                  IF (DIST.GE.D(K)) GO TO 70
                  DO 65 JJ=K,YES
                     KK=YES+K-JJ
                     MM=KK+1
                     D(MM)=D(KK)
   65          N1(MM)=N1(KK)
               D(K)=DIST
               N1(K)=I
               GO TO 75
C
   70 CONTINUE
   75 CONTINUE
      DUM1=0.0
      PSUM=0.0
      DO 80 I=1,6
      YES=N1(I)
      MAXW=1.0/(SORT(D(I)))
      PSUM=PSUM+MAXW
   80 DUM1=DUM1+W(YES)*MAXW
      DUM6=DUM1/PSUM
      N11=N1(1)
      DUM2=W(N11)
      DUM=(DUM2+DUM6)/2.0
      IF (DUM.GT.MAXX.OR.DUM.LT.MINX) Z(M,J)=DUM2
      IF (DUM.LE.MAXX.AND.DUM.GE.MINX) Z(M,J)=DUM
   85 CONTINUE
   90 CONTINUE
      IF (LIST.EQ.2.OR.LIST.EQ.3) WRITE(6,190) (Z(M,I),I=1,COLS)
      IF (IPUN.EQ.1) WRITE(7,195) (Z(M,I), I=1,COLS)
   95 CONTINUE
      NEWTOT=ROWS*COLS
      CALL STATS (ROWS,COLS,100,100,Z,ZBAR,XP,ZSIG,ZMAX,ZMIN)
      YMIN=YP
      WRITE (6,200) NEWTOT,XMIN,YMIN,ZMAX,ZMIN,ZBAR,ZSIG
      WRITE (6,205)
      IF (NMAP.LT.1) GO TO 105
C
  110 FORMAT (2I3,F4.0,3I1,I2)
  115 FORMAT(12A4)
  120 FORMAT(1H1,12A4,//,3X,24H COORDINATE OBSERVATIONS
  125 FORMAT(6X,12A4)
  130 FORMAT (5X,4H I =,I3,7H X(I) =,F10.3,7H Y(I) =,F10.3,7H W(I) =,F10
     1.3)
  135 FORMAT(1H1,12A4,//,3X,I5,13H OBSERVATIONS   ,//)
  140 FORMAT (3X,25H BOUNDARIES PREDETERMINED)
  145 FORMAT (4F10.0)
  150 FORMAT (3X,32H BOUNDARIES CALCULATED FROM DATA)
  155 FORMAT (3X,6HNMAP =,I3,7H LIST =,I3,6H BND =,I2,7H TAPE =,I3)
  160 FORMAT (3X,15HOBSERVED VALUES,//,3X,6HXBAR =,F10.3,6HXSIG =,F10.3,
     16HMAXX =,F10.3,6HMINX =,F10.3,//3X,6HYBAR =,F10.3,6HYSIG =,F10.3,6
     2HMAXY =,F10.3,6HMINY =,F10.3,//3X,6HWBAR =,F10.3,6HWSIG =,F10.3,6H
     3MAXW =,F10.3,6HMINW =,F10.3.//3X,6HAREA =,F10.3,8HAVEDIS =,F10.3,6
     4HBEST =,F10.3,//3X,7HGSIZE =,F10.3,6HROWS =,I4,6HCOLS =,I4,/)
  165 FORMAT(3X, ' ROWS,COLUMNS, OR GRID SIZE ' )
  170 FORMAT(3X,' SPECIFIED A PRIORI ' )
```

Purpose The program produces isarithmic maps on the line printer from scattered observations by first interpolating to a lattice and then contouring these regular values. A location map of the observations, a list of the values at the grid intersections, and the contoured map are returned.

Description: For each lattice point the program examines every observation, deleting from consideration all but the nearest NIPT (<10), regardless of direction from the grid point. The only exception to this search procedure occurs when an observation within IMIN (real) of the point is encountered, at which time its value is accepted as the lattice value and no further observations are tested. In the general case the value assigned to the lattice point is

$$\frac{1}{2}\left[z_c + \frac{\sum_i z_i/d_i}{\sum_i 1/d_i}\right]$$

where z_c is the value of the nearest observation, d is distance, and the sums are carried from the nearest to the NIPTth nearest observation. The lattice values are then printed row by row in a table, after which it is possible to terminate the program.

The location map and contour map are constructed simultaneously. Values appropriate to the latter are stored in a scratch file for later printing, while the contents of the former are printed as the contouring algorithm is executed. This routine determines values along horizontal lines by linear interpolation of the values at the bounding intersections. The values of the printer elements within each row of cells are interpolated linearly from the appropriate elements of the two bounding horizontal lines.

The location map shows grid intersections by ticks and the observations by asterisks followed by the Z values. The contour map prints a different symbol for each contour interval and denotes observation locations by asterisks alone. Both have X and Y scales printed in the margins.

Comments: Within the restrictions that the observations be given in rectangular coordinates and that the map produced be isarithmic the user has great flexibility. The spacing of ticks on the map is fixed both horizontally and vertically at one inch, with the result that a large map is produced in sections of twelve inch width and unlimited length. The size of the map is determined by the user through the variable SCALE. This is the interval along both the X and Y axes which will correspond to one inch on the map. This value should be chosen with care, as it is in the interpolation of grid values that much time is consumed. The program examines each of the N observations, regardless of the section of the map into which it falls, to find the NIPT nearest observations to each grid point. SCALE may be chosen as zero, the default value, in which case the program assigns to it the value of

$$\left[(X_{max} - X_{min}) * (Y_{max} - Y_{min}) / N\right]^{1/2}$$

a rough approximation of the number of map cells to the number of observations.

The user may read in the maxima and minima of the X, Y, and Z values over which the map is to range, or he may default to the extrema of the observations. He may request a certain contour interval; default is one tenth the range of Z, whether read in or determined by the program. If the interval chosen yields more than 37 contours the Z increment is doubled. The user may state the maximum distance from a grid intersection at which an observation's Z value is acceptable for the grid value. It is advisable to make this distance IMIN small relative to SCALE; the default is SCALE/25. He may also define NIPT as the number of nearest neighboring observations from which the grid values are to be interpolated. Maximum is nine, default is six. There are also options to suppress printing of certain output features and to punch the values at the lattice points.

The isarithmic map comes in one of three styles. The user may choose to have the entire field filled with symbols, a completely contoured map with different symbols representing different Z intervals; contours are denoted by a change of symbol. Or alternate contour intervals may be suppressed, leaving white bands between the printed intervals. This will usually improve the visual impact. The third choice is an actual isarithmic map, although the program does not calculate points along the line but uses instead the lowest ring of characters in each interval as an approximation to the contour. All maps have the locating asterisks for the observations.

I/O units 8 and 9 are used for storage during execution and should be set to temporary files. Unit 7 is used for punch output; if no cards are desired set this to a temporary file. The first control card reads the units to be used for input and output. These are usually 5 and 6. 500 observations can be handled by current version of the program.

Deck Make-up: The main program reads two control cards and then the observations:

1) I/O Card

2) Variable Format Card

3) Observations

The subroutine Grid reads three more cards, Grid Cards I, II, and III.

Card Formats: <u>I/O Card</u>

1-4	N	Number of observations
8	INPUT	I/O unit to be used for input (reading)
12	OUTPUT	I/O unit to be used for output (writing)

<u>Variable Format Card</u>

This describes the N groups of X, Y, and Z values. As many of these triples may be placed on the same observation card as desired; the format card will include between parentheses a description of the arrangement.

<u>Observation deck</u>: X, Y, Z values

<u>Grid Card I</u>: This reads an alphanumeric title of fewer than 81 spaces.

<u>Grid Card II</u>:

1-10	SCALE	Data units per lattice cell (floating point); default: leave blank
11-22	UNITS	As many as 12 characters describing the units of the observation measurement, e.g., as "inches."

The following eight options are available; a 1 selects the option and a 0 or blank suppresses it:

23	PIN	Print a table of input values
24	PMXMN	Do not print input extrema
25	PGRID	Do not print the gridded values
26	PHGRID	Punch gridded values (0 is recommended)
27	PLMAP	Do not print the location map
28	PCMAP	Do not print the contour map
29	MTYPE	Map style: 0 is alternate bands, 1 is contour lines, and 2 is all contour bands.
30	PCNTRL	Print control values

The following four values are floating point; default is blanks:

31-40	MAXX	Right margin value
41-50	MINX	Left margin value
51-60	MAXY	Top margin value
61-70	MINY	Bottom margin value
72	IFLIM	If the margin values have been read in, set to 1; if blank and defaulted

to extrema of the observations, set to 0

Grid Card III: The first four values are floating point, while the last two are integer.

1-10	MAXZ	Highest contour altitude
11-20	MINZ	Lowest contour altitude
21-30	ZINC	Contour interval; default leave blank
31-40	DMIN	Maximum acceptance distance for grid point; default leave blank
42	NIPT	Number of interpolation points; default leave blank
44	IFZLIM	If the Z extrema have been read in, set to 1; if not and the default is chosen, set to 0

Programmer: Original program (GRID, SCON, STATS, in Mad) by Waldo R. Tobler, The University of Michigan, 1965. Modifications and translation to FORTRAN IV by D. Gill, The University of Michigan, May 1967. Further modifications by D. Bowman and D. Rhynsburger, 1969.

```
113   FORMAT (1H1,25X,20A4)
114   FORMAT('0',45X,'-',17X,'BELOW',6X,F12.4)
115   FORMAT(46X,A1,10X,F12.4,3X,'TO',F13.4)
116   FORMAT (46X,A1,10X,F12.4)
117   FORMAT(46X,'+',17X,'ABOVE',6X,F12.4)
118   FORMAT (1H0,59X,12HINPUT VALUES  )
119   FORMAT (1H0,3(15X,1HX,9X,1HY,9X,1HZ,4X))
120   FORMAT (1H0,6(3X'COL'7X'VALUE'3X))
121   FORMAT('0',25X,'LOCATION MAP..SCALE IS',E12.4,2X,3A4,'PER INCH')
122   FORMAT('0',25X,'CONTOUR MAP....SCALE IS',E12.4,2X,3A4,'PER INCH')
123   FORMAT('OMAXX = ',E12.5,4X,'MINX= ',E12.5,4X,'MAXY = 'E12.5,4X,'MI
     1NY = ', E12.5,4X,'MAXZ = ',F12.5,4X,'MINZ = ',E12.5)
124   FORMAT('OSCALE IS ',E12.5,1X,3A4,' PER INCH. ALTITUDE INCREMENT IS
     1 ',E12.5,'.  MAXIMUM ACCEPTANCE DISTANCE IS ',E12.5)
125   FORMAT('OOPTIONS ARE:  PRINT INPUT VALUES (',F2.0,')  SUPPRESS CHA
     1RACTERISTICS TABLE (',F2.0,')  SUPPRESS PRINTING OF GRIDDED VALUES
     2 (',F2.0,')')
126   FORMAT('OPUNCH GRIDDED VALUES (',F2.0,')  SUPPRESS LOCATION MAP  (
     1',F2.0,')  SUPPRESS CONTOUR MAP  (',F2.0,')  CONTOUR MAP TYPE  (',
     2F2.0,')')
127   FORMAT('ONUMBER OF INTERPOLATION POINTS  (',F2.0,')  PRINT CONTROL
     1 VALUES  (',F2.0,')   NUMBER OF DATA POINTS  (',I3,')')
128   FORMAT('0',25X,'SECTION ',I3,' OF ',I3)
129   FORMAT('0',57X,'INITIAL CONTROL INPUT')
130   FORMAT('0',57X,'FINAL CONTROL VALUES')
131   FORMAT('ONUMBER OF:  INTERPOLATION POINTS  (',I1,')   VERTICAL MAP
     1 CELLS  (',I2,')  HORIZONTAL MAP CELLS  (',I2,')   MAP VERTICAL S
     2ECTIONS  (',I2,')')
133   FORMAT (1H ,I2,'INTERPOLATION POINTS IS TOO MANY. NINE WILL BE USE
     1D')
134   FORMAT(' ',I1,' DATA POINTS ARE TOO FEW.  RETURN FROM GRID.')
      READ (INPUT,105) TITLE
      READ (INPUT,101) SCALE,UNITS,PIN,PMXMN,PGRID,PHGRID,PLMAP,PCMAP,
     1  MTYPE,PCNTRL,MAXX,MINX,MAXY,MINY, IFLIM
      READ (INPUT,110) MAXZ, MINZ,ZINC,IMIN,NIPT,IFZLIM
      IF (PCNTRL.EQ.0.) GO TO 76
      WRITE (OUTPUT,113) TITLE
      WRITE (OUTPUT,129)
      WRITE (OUTPUT,123) MAXX,MINX,MAXY,MINY,MAXZ,MINZ
      WRITE (OUTPUT,124) SCALE,UNITS,ZINC,IMIN
      WRITE (OUTPUT,125) PIN,PMXMN,PGRID
      WRITE (OUTPUT,126) PHGRID,PLMAP,PCMAP,MTYPE
      WRITE (OUTPUT,127) NIPT,PCNTRL,N
   76 IF (PIN.EQ.0.) GO TO 4
      WRITE (OUTPUT,113) TITLE
      WRITE (OUTPUT,118)
      WRITE (OUTPUT,119)
      WRITE (OUTPUT,103) (X(I),Y(I),Z(I),I=1,N)
C                             FIND MAXIMA AND MINIMA
    4 IF (IFLIM .EQ. 1)  GO TO 501
      XMAX=X(1)
      XMIN=XMAX
      YMAX=Y(1)
      YMIN=YMAX
      DO 3 I=2,N
      IF (XMAX.LT.X(I)) XMAX=X(I)
```

```
      IF (YMAX.LT.Y(I)) YMAX=Y(I)
      IF (XMIN.GT.X(I)) XMIN=X(I)
      IF (YMIN.GT.Y(I)) YMIN=Y(I)
    3 CONTINUE
      GC TO 502
  501 XMAX = MAXX
      XMIN = MINX
      YMAX = MAXY
      YMIN = MINY
  502 IF (IFZLIM .EQ. 1)  GO TO 504
      ZMAX=Z(1)
      ZMIN=ZMAX
      DO 503 I = 2,N
      IF (ZMAX.LT.Z(I)) ZMAX=Z(I)
      IF (ZMIN.GT.Z(I)) ZMIN=Z(I)
  503 CONTINUE
      GC TO 505
  504 ZMAX = MAXZ
      ZMIN = MINZ
  505 IF (PMXMN .NE. 0)  GO TO 7
      WRITE (OUTPUT,113) TITLE
      WRITE (OUTPUT,102) XMAX,XMIN,YMAX,YMIN,ZMAX,ZMIN
C                             CALCULATE CONTOUR INTERVALS
    7 DZ=ZMAX-ZMIN
      IF (ZINC.EQ.0.) ZINC=DZ/10.
   30 CON=DZ/ZINC
      IF (CON.LE.36) GO TO 29
      WRITE (OUTPUT,111) ZINC
      ZINC=2.*ZINC
      GC TO 30
C                             CALCULATE CONSTRUCTION PARAMETERS
C                             NCOL IS TOTAL NUMBER OF HORIZONTAL CELLS
C                             NSECT IS TOTAL NUMBER OF VERTICAL MAP
C                               SECTIONS PRODUCED
C                             NSECTC DETERMINES WHICH SECTION WE'RE IN
C                             ROWS IS VERTICAL NUMBER OF CELL BOUNDARIES
C                             ROWSM1 IS VERTICAL NUMBER OF CELLS
   29 IF (SCALE.EQ.0.) SCALE=SQRT((XMAX-XMIN)*(YMAX-YMIN)/N)
      NCOL=(XMAX-XMIN)/SCALE+.99
      NSECT=(NCOL-1)/12+1
      NSECTC=0
      ROWS=(YMAX-YMIN)/SCALE +1.99
      ROWSM1=ROWS-1
      MAXX=XMIN
      IF (IMIN.EQ.0.) IMIN=SCALE/25.
      IMIN = IMIN * IMIN
      IF (NIPT.EQ.0) NIPT=6
      IF (NIPT.LT.10) GO TO 12
      WRITE (OUTPUT,133) NIPT
      NIPT=9
   12 IF (N.GE.NIPT) GO TO 80
      WRITE (OUTPUT,134) N
      RETURN
C                             CYCLIC SETUP
C                             INDEX IS AN ARRAY WHICH CORRESPONDS TO
C                             THE INPUT DATA ARRAY.  VALUES MEAN:
```

```
C                              1 - DATA POINT IS IN CURRENT SECTION
C                              0 - DATA POINT IS IN A FUTURE SECTION
C                             -1 - DATA POINT IS IN A PREVIOUS SECTION
   80 DO 77 I=1,N
   77 INDEX(I)=0
      XINC=3.
      YINC=5.
      PS=30.
      IF (PCNTRL.EQ.0.) GO TO 79
      WRITE (OUTPUT,113) TITLE
      WRITE (OUTPUT,130)
      WRITE (OUTPUT,123) XMAX,XMIN,YMAX,YMIN,ZMAX,ZMIN
      WRITE (OUTPUT,124) SCALE,UNITS,ZINC,IMIN
      WRITE (OUTPUT,131) NIPT,ROWSM1,NCOL,NSECT
C                              HERE BEGINS THE CYCLIC PROCEDURE WHICH
C                              PRODUCES VERTICAL MAP SECTIONS
   79 CALL REWIND(8)
      CALL REWIND(9)
C                              DECIDE MAP WIDTH
C                              COLS IS THE HORIZONTAL NUMBER OF CELL
C                              BOUNDARIES IN THE CURRENT SECTION
C                              COLSM1 IS THE HORIZONTAL NUMBER OF CELLS
C                              IN THE CURRENT SECTION
C                              NUM IS WHERE MARGIN VALUE PRINTING STARTS
      IF (NCOL.GT.12) GO TO 13
      COLS=NCOL+1
      NCOL=0
      GO TO 14
   13 COLS=13
      NCOL=NCOL-12
   14 COLSM1=COLS-1
      NUM=10*COLSM1+4
      IF (PHGRID.EQ.0.) GO TO 1
      WRITE(7,105) TITLE
      WRITE(7,106) SCALE,ROWS,COLS
C                              FIND HORIZONTAL MAX AND MIN FOR
C                              THIS SECTION
    1 MINX=MAXX
      MAXX=MINX+(COLS-1)*SCALE
C                              ELIMINATE OUT-OF-SECTION POINTS
      DO 24 I=1,N
      IF (INDEX(I).LT.0) GO TO 24
      IF (X(I).LE.MAXX.AND.X(I).GE.MINX.AND.Y(I).LE.YMAX.AND.Y(I).GE.
     1  YMIN) INDEX(I)=1
   24 CONTINUE
      IF (PGRID.NE.0.) GO TO 78
      WRITE (OUTPUT,113) TITLE
      WRITE (OUTPUT,120)
C                              SET UP ITERATION WHICH DOES GRIDDING.
C                              FOR EACH GRID ROW:
   78 YP=YMAX
      NIP = NIPT + 1
      DO 15 M=1,ROWS
C                              FOR EACH GRID COLUMN WITHIN THE ROW
C                              XP AND YP ARE COORDINATES OF GRID POINT
      XP=MINX
```

```
      DC 16 J=1,COLS
      DXM = 1.E10
      DYM = 1.E10
      DO 20 I = 1,NIP
   20 DELV(I) = 1.E10
C                                 FIND 'NIPT' DATA POINTS NEAREST GRID POINT
      DC 19 I=1,N
      TEMP=XP-X(I)
      TEMP1=YP-Y(I)
      IF (TEMP .GT. DXM .AND. TEMP1 .GT. DYM)  GO TO 19
      DIST=TEMP*TEMP+TEMP1*TEMP1
      K = NIP
      IF (DIST .GE. IMIN)  GO TO 21
      ZROW1(J)=Z(I)
      GO TO 25
   23 K=K-1
      DELV(K+1)=DELV(K)
      IPREV(K+1)=IPREV(K)
   21 IF (K.EQ.1) GO TO 22
      IF (DIST.LT.DELV(K-1) ) GO TO 23
      IF (K .NE. NIPT)  GO TO 22
      DXM = TEMP
      DYM = TEMP1
   22 DELV(K)=DIST
      IPREV(K)=I
19    CONTINUE
C                                 INTERPOLATE WITH 'NIPT' NEAREST POINTS
      TEMP=0.0
      TEMP1=0.0
      DO 17 I=1,NIPT
      TEMP2=1./SQRT(DELV(I))
      TEMP1=TEMP1+TEMP2
   17 TEMP=TEMP+Z(IPREV(I))*TEMP2
      ZROW1(J)=(Z(IPREV(1))+TEMP/TEMP1)/2.
   25 XP=XP+SCALE
16    CONTINUE
      YP=YP-SCALE
      IF (PHGRID .EQ. 1)  WRITE (7,109) (ZROW1(I),I=1,COLS)
      IF (PGRID .EQ. 0)  WRITE (OUTPUT,107) M, (I, ZROW1(I),I=1,COLS)
      IF (PLMAP.NE.0..AND.PCMAP.NE.0.) GO TO 15
      WRITE(8,109) (ZROW1(I),I=1,COLS)
15    CONTINUE
C
C                                 DONE WITH GRIDDING
      IF (PLMAP.NE.0..AND.PCMAP.NE.0.0) GO TO 72.
      NSECTC=NSECTC+1
      IF (PLMAP.NE.0.) GO TO 74
      WRITE (OUTPUT,113) TITLE
      WRITE (OUTPUT,121) SCALE,UNITS
      WRITE (OUTPUT,128) NSECTC,NSECT
   74 IF (PCMAP.NE.0.) GO TO 75
      WRITE(9,113) TITLE
      WRITE(9,122) SCALE, UNITS
      WRITE(9,128) NSECTC, NSECT
C                                 SET UP THE TOP ROW OF NUMBERS
   75 DO 32 I=1,131
```

```
      PSYM(I)=BLANK
   32 PSYMA(I)=BLANK
      TEMP=MINX
      DO 31 I=1,COLS
      K=10*(I-1)+1
      CALL BNBCD(TEMP,PSYM(K) )
   31 TEMP=TEMP+SCALE
      IF (PLMAP .EQ. 0)  WRITE (OUTPUT,112) MINUS, PSYM
      IF (PCMAP .EQ. 0)  WRITE (9,112) MINUS, PSYM
C                                SET UP THE TOP MARGIN
      PSYM(1)=BLANK
      DO 38 I=1,NUM,10
      PSYM(I+1)=VBAR
      DO 38 J=2,10
   38 PSYM(I+J)=BLANK
      IF (PLMAP .EQ. 0)  WRITE (OUTPUT,112) NINE, PSYM
      IF (PCMAP .EQ. 0)  WRITE (9,112) NINE, PSYM
C                                SET UP THE FIRST ROW
      CALL REWIND(8)
      READ(8,109) (ZROW1(I),I=1,COLS)
      COL=1
      DO 36 J=1,COLSM1
      DELH=ZROW1(J+1)-ZROW1(J)
      VNOW(J)=ZROW1(J)
      XP=-XINC
   37 XP=XP+XINC
      IF (XP.GE.PS) GO TO 36
      COL=COL+1
      INEXT(COL) = (ZROW1(J) + XP * DELH / PS - ZMIN) / ZINC + 1.
      PSYM(COL)=BLANK
      PSYMA(COL)=BLANK
      GO TO 37
36    CONTINUE
      COL=COL+1
      VNOW(COLS)=ZROW1(COLS)
      INEXT(COL) = (ZROW1(COLS) - ZMIN) / ZINC + 1.
      PSYM(COL)=BLANK
      PSYMA(COL)=BLANK
      YVAL=YMAX
C                                DETERMINATION OF Z VALUES FOR INTERIOR
      DO 42 M=1,ROWSM1
      READ(8,109) (ZROW1(I),I=1,COLS)
C                                CALCULATE ALTITUDE DIFFERENCES
C                                BETWEEN ROWS
      DO 40 I=1,COLS
   40 DELV(I)=ZROW1(I)-VNOW(I)
      PSYM(1)=MINUS
      PSYMA(1)=MINUS
      K=NUM-1
      DO 57 I=2,K,10
   57 PSYMA(I)=PLUS
      PSYMA(NUM-1)=MINUS
      PSYM(NUM-1)=MINUS
      CALL BNBCD(YVAL,PSYM(NUM) )
      CALL BNBCD(YVAL,PSYMA(NUM) )
C                                SET UP INTERMEDIATE ROW INCREMENTS
```

```
      YP=-YINC
   41 YP=YP+YINC
      IF (YP.GE.PS) GO TO 42
      COL=1
      VNOW(1)=YINC*DELV(1)/PS+VNOW(1)
C                                FOR EACH GRID COLUMN
      DO 43 J=1,COLSM1
      VNOW(J+1)=YINC*DELV(J+1)/PS+VNOW(J+1)
      DELH=VNOW(J+1)-VNOW(J)
C                                SET UP INTERMEDIATE COLUMN INCREMENTS
      XP=-XINC
   45 XP=XP+XINC
      IF (XP.GE.PS) GO TO 43
      IPREV(COL)=I
      COL=COL+1
      I=INEXT(COL)
      INEXT(COL) = (VNOW(J) + XP * DELH / PS - ZMIN) / ZINC + 1.
C                                CHOOSE THE PRINT SYMBOL
      IF (I.GE.1) GO TO 46
      PSYM(COL) =MINUS
      GO TO 45
   46 IF (I.LE.CON) GO TO 47
      PSYM(COL)=PLUS
      GO TO 45
   47 PSYM(COL)=SYMTAB(I)
      IF ((M.EQ.1.AND.YP.EQ.0.).OR.COL.EQ.2) GO TO 45
      IF (MTYPE-1.) 49,48,45
   49 IF (I/2*2.EQ.I) PSYM(COL)=BLANK
      GO TO 45
   48 IF (I.GT.IPREV(COL).OR.I.GT.IPREV(COL-1).OR.I.GT.INEXT(COL+1).OR.
     1  I.GT.INEXT(COL)) GO TO 45
      PSYM(COL)=BLANK
      GO TO 45
43    CONTINUE
      IPREV(COL)=I
      COL=COL+1
      I=INEXT(COL)
      INEXT(COL) = (VNOW(COLS) - ZMIN) / ZINC + 1.
      IF (I.GE.1) GO TO 60
      PSYM(COL)=MINUS
      GO TO 61
   60 IF (I.LE.CON)  GO TO 50
      PSYM(COL)=PLUS
      GO TO 61
   50 PSYM(COL)=SYMTAB(I)
   61 COL=1
   62 DO 51 I=1,N
      IF (INDEX(I).LE.0) GO TO 51
      IF (ABS(Y(I)-YVAL).GT.SCALE/12.) GO TO 51
      INDEX(I)=-1
      J=10.*(X(I)-MINX)/SCALE+2.5
      PSYM(J)=ASTRK
      PSYMA(J) = ASTRK
      CALL BNBCD(Z(I),PSYMA(J+1))
51    CONTINUE
      YVAL=YVAL-SCALE/6.
```

```
      GO TO 79
      END
      SUBROUTINE BNBCD(VAL,ST)
      DIMENSION ST(1), TS(8)
      REAL X/'X'/, POINT/'.'/, BLANK/' '/, QMINU/'-'/
      REAL SYM(10)/'0','1','2','3','4','5','6','7','8','9'/
      MAX = 5
      L = 1
      IF (VAL .GE. 0.)  GO TO 1
      ST(L) = QMINU
      L = 2
    1 V = ABS(VAL)
      I = MAX
      ST(L) = SYM(1)
      IF (V .LT. 10.**MAX)  GO TO 3
      DO 2 L = L,MAX
    2 ST(L) = X
      RETURN
C          INTEGRAL PART
    3 IF (V .LT. 1.)  GO TO 7
    4 IF (V .LT. 1.)  GO TO 5
      TS(I) = SYM(V - 10. * AINT(V / 10.) + 1.)
      V = V / 10.
      I = I - 1
      IF (I .GT. 0)  GO TO 4
    5 I = I + 1
      DO 6 I = I,MAX
      ST(L) = TS(I)
    6 L = L + 1
      L = L - 1
      IF (L .GE. MAX)  RETURN
C          FRACTIONAL PART
    7 L = L + 1
      V = ABS(VAL - AINT(VAL))
      ST(L) = POINT
      IF (V .EQ. 0.)  RETURN
      L = L + 1
    8 V = 10. * V
      ST(L) = SYM(V + 1.)
      L = L + 1
      IF (L .GT. MAX)  RETURN
      V = V - AINT(V)
      IF (V .NE. 0.)  GO TO 8
      RETURN
      END
```

GEOGRAPHICAL NEIGHBORS

Purpose: Given NC points identified by rectangular coordinates X_i, Y_i, $i = 1 \ldots NC$ the program produces the adjacency matrix of neighbors of order $NN \leq 9$. These can then be used to find the neighbors to points of a regular lattice.

Description: Subroutine THIESS calculates the adjacency matrix. Neighbors are defined on the basis of adjacent Thiessen polygons: Let A_i be the area closer to point i than to any other point, and similarly for A_j with respect to point j. If A_i and A_j touch (contact>1) then i and j are first order neighbors. Higher order neighbors to a point are defined by deleting neighbors of one lower order. The search algorithm is from a program by Gambini, and requires examination of the lines of equilibrium between all pairs of observations. Specifically, the search proceeds by incrementing along the perpendicular bisector of the line connecting each pair until a position is reached for which the "attraction" is greatest. If the equilibrium point is not attained by the time the search reaches the boundary, another pair is considered. The algorithm is slow and tedious, but no better procedure is known.

A rectangular lattice can now be superimposed on the original point distribution, and the points of the original set which are the neighbors to the lattice points are calculated. The variable SCALE determines the number of coordinate units between the lattice points. If the adjacency matrix of the original set of points is known in advance, this may be entered as data.

Deck Make-up: 1) Control Card

2) Limits Card

3) Format Card

4) Observations

5) Adjacency Matrix (optional)

Data Cards: <u>Control Card</u> (2I5, 4X, I1, 4X, I1)

columns		
1-5	NC	number of observations
6-10	NN	maximum number of neighbors desired.
15	INNBR	0 if neighbors are to be computed 1 if adjacency matrix is to be read as data
20	IFPCH	1 if adjacency matrix is to be punched.

<u>Limits Card</u>

1-10	XO	minimum X coordinate
11-20	X1	maximum X coordinate
21-30	YO	minimum Y coordinate
31-40	Y1	maximum Y coordinate The above limits all apply to the lattice region.
41-50	DMAX	Search increment; blank yields valid default of SCALE/100. Small value of DMAX results in excessive computation

		time; large value increases the probability of missing a short boundary between adjacent polygons (i.e., missing a neighbor).
51-60	RLIM	maximum search radius; blank yields valid default.
61-70	SCALE	unit lattice size; blank yeilds default of $[(XMAX - XMIN)*(YMAX - YMIN)/N]^{\frac{1}{2}}$

Format Card (20A4)

Format for the X, Y coordinates of the observations.

Observations

X, Y coordinates punched as described on the format card.

Adjacency Matrix

Read iff INNBR = 1. Punched as 80 I 1; the number is the order of the adjacency, zero or blank implies non-adjacent.

Programmer: Dierk Rhynsburger, University of Michigan.

Reference: R. Gambini, "A Computer Program for Calculating Lines of Equilibrium between Multiple Center of Attraction," Center for Regional Studies, University of Kansas, Lawrence, no date.

```
C        THIESSEN POLYGONS OF GRID INTERSECTIONS
      DIMENSION XC(100),YC(100), UC(100), VC(100), IZ(100,100), IFMT(20)
      COMMON XC, YC, NC
      COMMON /LIM/ XO, X1, YO, Y1
      COMMON /MGT/NN, IZ, DMAX, RLIM
   10 FORMAT (16I5)
   11 FORMAT (80I1)
   12 FORMAT (1H1,15HLIMITS OF DATA:/1H0,10X,6HXMAX =,F10.3,10X,6HYMAX =
     &   F10.3/1H0,10X,6HXMIN =,F10.3,10X,6HYMIN =,F10.3)
   13 FORMAT (1H-,8HSCALE IS,F8.3,25H DATA UNITS PER GRID CELL)
   14 FORMAT (8E10.3)
   15 FORMAT (20A4)
   16 FORMAT (1H-,23HINCREMENT FOR SEARCH IS,E18.6)
   17 FORMAT (1H-,16HSEARCH RADIUS IS,E18.6)
   18 FORMAT (1H-,39HNEIGHBORS OF OBSERVATIONS READ AS INPUT)
      READ 10, NC, NN, INNBR, IFPCH
      READ 14, XO, X1, YO, Y1, DMAX, RLIM, SCALE
      READ 15, (IFMT(I),I=1,20)
      READ (5,IFMT) (XC(I), YC(I), I=1,NC)
      IF (RLIM .EQ. 0.)  RLIM = SQRT((X1 - XO)**2 + (Y1 - YO)**2)
      IF (SCALE .EQ. 0.)  SCALE = SQRT((X1 - XO) * (Y1 - YO) / NC)
      IF (DMAX .EQ. 0.)  DMAX = SCALE / 100.
      PRINT 12, X1, Y1, XO, YO
      PRINT 13, SCALE
      PRINT 16, DMAX
      PRINT 17, RLIM
      IF (INNBR .EQ. 0)  GO TO 98
      READ 11, ((IZ(I,J),J=1,NC),I=1,NC)
      PRINT 18
      GO TO 99
   98 CALL THIESS(IFPCH)
   99 NB = (X1 - XO) / SCALE + 1.99
      NA = (Y1 - YO) / SCALE + 1.99
      YP = Y1 + SCALE
      DO 100 I = 1,NA
      YP = YP - SCALE
      XP = XO - SCALE
      DO 100 J = 1,NB
      XP = XP + SCALE
  100 CALL GNBR(XP,YP)
      CALL EXIT
      END
      SUBROUTINE THIESS(IFP)
      DIMENSION XC(500), YC(500), IX(500), IS(500), IY(500,15)
      DIMENSION IZ(50,50), IR(500), IN(500,10)
      COMMON XC, YC, NC
      COMMON /LIM/ XO, X1, YO, Y1
      COMMON /MGT/NN, IZ, DMAX, RLIM
   25 FORMAT (1H1)
   26 FORMAT (1H ,50I2)
   27 FORMAT (1H1,35HFIRST NEIGHBORS OF THE OBSERVATIONS)
   28 FORMAT (1H1,9HNEIGHBORS,I4)
   29 FORMAT (1H0,11X,1HX,9X,1HY)
   30 FORMAT (1H-,I4,2F10.2,I8,4X,20I4)
   31 FORMAT (1H-,I4,2F10.2)
```

```
      DO 1010 I = 1,NC
 1010 PRINT 26, (IZ(I,J),J=1,NC)
      IF (NN .EQ. 1)  RETURN
      DO 2000 I = 1,NL
      DO 2000 J = 1,NS
      JJ = J + 1
      DO 1995 K = JJ,NC
      IF (IZ(J,K) .NE. I)  GO TO 1995
      DO 1990 L = 1,NC
      IF (IZ(K,L) .NE. 1)  GO TO 1990
      IF (IZ(J,L) .NE. 0)  GO TO 1990
      IF (J .EQ. L)  GO TO 1990
      IZ(J,L) = I + 1
      IZ(L,J) = I + 1
 1990 CONTINUE
 1995 CONTINUE
 2000 CONTINUE
      PRINT 33
      DO 2050 I = 1,NC
 2050 PRINT 51, (IZ(I,J),J=1,NC)
      PRINT 25
      IF (IFP .EQ. 1)  WRITE (7,52) ((IZ(I,J),J=1,NC),I=1,NC)
      DO 2100 I = 1,NC
      PRINT 31, I, XC(I), YC(I)
      DO 2100 J = 1,NN
      IW = 0
      DO 2090 K = 1,NC
      IF (IZ(I,K) .NE. J)  GO TO 2090
      IW = IW  +1
      IS(IW) = K
 2090 CONTINUE
      IF (IW .EQ. 0)  GO TO 2100
      PRINT 32, IW, J, (IS(K),K=1,IW)
 2100 CONTINUE
      PRINT 25
 9999 RETURN
      END
      SUBROUTINE GNBR(UC,VC)
      DIMENSION XC(500), YC(500), AA(500), BB(500)
      DIMENSION IZ(50,50), IR(100), IX(20), IY(20)
      COMMON XC, YC, NC
      COMMON /LIM/ X0, X1, Y0, Y1
      COMMON /MGT/NN, IZ, DMAX, RLIM
   12 FORMAT (1H1,12HNEIGHBORS TO,F9.2,1H,,F7.2)
   13 FORMAT (1H-,9HNEIGHBORS,I4,I15)
   14 FORMAT (1H ,15X,I4,2F8.2)
   23 FORMAT (1H ,21HRLIM HAS BEEN DOUBLED)
   24 FORMAT (1H-,30HRLIM DOUBLED: CHANGE GRID SIZE)
      II = 0
      IO = 1
      IRL = 1
      PRINT 12, UC, VC
      DO 100 J = 1,NC
      A = UC - XC(J)
      B = VC - YC(J)
      IF (A .EQ. 0. .AND. B .EQ. 0.)  GO TO 920
```

```
      AA(J) = A
  100 BB(J) = B
  105 DO 900 J = 1,NC
      DX = AA(J)
      DY = BB(J)
      DEN = SQRT(DX * DX + DY * DY)
      IF (DEN .GT. RLIM)  GO TO 900
      IRL = 0
      AX = (UC + XC(J)) / 2.
      AY = (VC + YC(J)) / 2.
      CALL LMT(DX,DY,AX,AY,DMAX,U,NSTEP)
      ITEST = 2
      DO 600 K = 1,NSTEP
      U = U + DMAX
      X = AX + U * DY
      Y = AY - U * DX
      IP = ITEST
      ITEST = IFUN(X,Y,UC,VC,J)
      IF (ITEST .NE. IP)  GO TO 700
  600 CONTINUE
      GO TO 900
  700 II = II + 1
      IX(II) = J
  900 CONTINUE
      IF (IRL .EQ. 0)  GO TO 940
      PRINT 23, UC, VC
      RLIM = RLIM + RLIM
      IF (IRL .EQ. 3)  GO TO 9998
      IRL = IRL + 1
      GO TO 105
  920 II = 1
      IX(1) = J
  940 PRINT 13, IO, II
      PRINT 14, (IX(I), XC(IX(I)), YC(IX(I)), I=1,II)
      IF (NN .EQ. 1)  RETURN
      IT = II
      DO 1050 I = 1,II
      IY(I) = IX(I)
 1050 IR(I) = IX(I)
      DO 1500 I = 2,NN
      IQ = 0
      DO 1070 J = 1,II
 1070 IX(J) = IY(J)
      DO 1300 J = 1,II
      DO 1200 K = 1,NC
      IF (IZ(IX(J),K) .NE. 1)  GO TO 1200
      DO 1100 L = 1,IT
      IF (IR(L) .EQ. K)  GO TO 1200
 1100 CONTINUE
      IQ = IQ + 1
      IT = IT + 1
      IY(IQ) = K
      IR(IT) = K
 1200 CONTINUE
 1300 CONTINUE
      IF (IQ .EQ. 0)  RETURN
```

```
      II = 10
      PRINT 13, I, II
 1500 PRINT 14, (IY(J), XC(IY(J)), YC(IY(J)), J=1,II)
      RETURN
 9998 PRINT 24
 9999 RETURN
      END
      SUBROUTINE LMT(DX,DY,AX,AY,USTEP,U,NSTEP)
      COMMON /LIM/ X0, X1, Y0, Y1
      IF (DY) 111,112,113
  111 IF (DX) 114,115,116
  112 IF (DX) 117,9999,118
  113 IF (DX) 119,120,121
  114 U = -AMIN1((AX - X1) / DY, (Y0 - AY) / DX)
      UL = AMIN1((X0 - AX) / DY, (AY - Y1) / DX)
      GO TO 124
  115 U = (X1 - AX) / DY
      UL = (X0 - AX) / DY
      GO TO 124
  116 U = -AMIN1((AX - X1) / DY, (Y1 - AY) / DX)
      UL = AMIN1((X0 - AX) / DY, (AY - Y0) / DX)
      GO TO 124
  117 U = (AY - Y0) / DX
      UL = (AY - Y1) / DX
      GO TO 124
  118 U = (AY - Y1) / DX
      UL = (AY - Y0) / DX
      GO TO 124
  119 U = -AMIN1((AX - X0) / DY, (Y0 - AY) / DX)
      UL = AMIN1((X1 - AX) / DY, (AY - Y1) / DX)
      GO TO 124
  120 U = (X0 - AX) / DY
      UL = (X1 - AX) / DY
      GO TO 124
  121 U = -AMIN1((AX - X0) / DY, (Y1 - AY) / DX)
      UL = AMIN1((X1 - AX) / DY, (AY - Y0) / DX)
  124 NSTEP = (UL - U) / USTEP + 1.
      U = U + USTEP
 9999 RETURN
      END
      FUNCTION IFUN(X,Y,UC,VC,J)
      DIMENSION XC(500), YC(500)
      COMMON XC, YC, NC
      COMMON /LIM/ X0, X1, Y0, Y1
      IFUN = 2
      ATT = 1. / SQRT((X - UC)**2 + (Y - VC)**2)
      DO 100 L = 1,NC
      IF (J .EQ. L)  GO TO 100
      IF (X .EQ. XC(L) .AND. Y .EQ. YC(L))  RETURN
      AT = 1. / SQRT((X - XC(L))**2 + (Y - YC(L))**2)
      IF (ATT .LT. AT)  RETURN
  100 CONTINUE
      IFUN = 3
      RETURN
      END
      FUNCTION JFUN(X,Y,I,J)
```

```
      DIMENSION XC(500), YC(500), A(500), ATTR(500)
      COMMON XC, YC, NC
      COMMON /LIM/ X0, X1, Y0, Y1
      JFUN = 2
      DC 100 L=1,NC
      IF (X .EQ. XC(L) .AND. Y .EQ. YC(L))  RETURN
      ATTR(L) = 1. / SQRT((X-XC(L))**2 + (Y-YC(L))**2)
  100 CONTINUE
      DO 200 M=1,NC
      IF((M.EQ.I).OR.(M.EQ.J)) GO TO 200
      IF (ATTR(I) .LT. ATTR(M))  RETURN
  200 CONTINUE
      JFUN = 3
  400 RETURN
      END
```

CONTOUR PLOTTING

Purpose: The program uses the 30 inch Calcomp plotter to draw contour maps from data given in the form of geographical matrices. Stereograms and perspective contours can also be obtained. An option allows conversion of the contours to a map projection before plotting.

Description: The contouring algorithm is that described by Dayhoff. The perspective plotting is based on that of Puckett.

Comments: File 2 must be used as a scratch tape. Input is from unit 7, comments are written onto unit 6, plotting is done via unit 9, and execution requires concatenation with *PLØTSYS.

Deck make-up: The controls are specified by an integer code punched in columns 1 and 2. The code also specifies whether further data are required on the same card, or whether additional cards are to be read. Reading of controls continues until plot code (20) is read. The number of control cards may vary from run to run. A basic sequence has been indicated by asterisks. The plot control card may be called repeatedly.

Map Projection Control Card:

Punch 05 in columns 1-2. This calls a map projection subroutine PROJ (X, Y, XMAX, YMAX, SCALE) which may read additional parameters. In the present instance the first call on PROJ is activated immediately following the first plot (code 20) which follows an 05 code, and the program then expects to read one card as follows:

Center latitude (decimal degrees) of the map in columns 1-7.

Center longitude of the map in 8-14.

Size of quadrilateral if different from one degree in columns 15-21.

Center latitude of an oblique stereographic projection in columns 22-28.

Center longitude of the oblique stereographic projection in columns 29-35.

Subsequent plots must repeat the 05 code if conversion to map projection coordinates is desired, and, as currently written, must use the same projection parameters. The width of the map using the 05 code is defined to be the width to scale at the center latitude on the square projection. Other map projections may be used by changing the subroutine PROJ.

*<u>Title Control Card</u>

Punch 12 in columns 1-2.

Follow this with any single card title (which will be plotted at the top of the drawing). Repeat as desired for all runs.

*<u>Format Control Card</u>

Punch 19 in columns 1-2.

Follow this with a single card (18A4) describing the E or F Fortran format for one row of the observation matrix.

*<u>Constant Contour Interval Card</u>

Punch 15 in columns 1-2

Punch lowest contour (with decimal point) in 3-12

Punch contour increment in 13-22

Punch highest contour in 23-32.

*<u>Data Definition Card</u>

Punch 01 in columns 1-2

Punch number of rows (LE.100) in data matrix as an E or F number in columns 3-12.

Punch number of columns (LE.100) in data matrix in columns 13-22.

Punch desired width (inches LE.28) of contour map in columns 23-32.

Punch denominator of vertical scale transformation in columns 33-42.

Follow this card by the data matrix as described by the earlier format control card.

*<u>Plot Control Card</u>

Punch 20 in columns 1-2

Plotting with the current controls is initiated. The remaining control cards (below) could also have been read before the Plot Control Card, with obvious exceptions.

<u>Width Change Card</u>

Punch 21 in columns 1-2

Punch desired width in columns 3-12.

Change Constant Contour Interval Card

Punch 16 in columns 1-2

Columns 3-12, 13-22, 23-32 are punched as for the Constant Contour Interval Card.

Variable Contour Interval Card

Punch 17 in columns 1-2

Punch the number of contour intervals desired in columns 3-12.

Follow this by one card which gives the format for the variable contours.

Follow this by the definition of the variable contours, from lowest to highest, on the appropriate number of of cards (as specified by the format).

Change Variable Contour Interval Card

Punch 18 in columns 1-2

Columns 3-12 and remaining cards punched as described under the Variable Contour Interval Card.

Non-Standard Plot File Card

Punch 11 in columns 1-2.

Punnh the file unit onto which plots are to be written in columns 3-12.

Extra Labels Card

Punch 14 in columns 1-2.

Punch starting x-coordinate of the label in columns 3-12.

Punch starting y-coordinate of the label in columns 13-22.

Punch starting z-coordinate of the label in columns 23-32.

Height of the lettering in inches goes into columns 33-42.

Follow this by one card containing any label in columns 1-36.

Repeat the entire sequence up to 10 times.

<u>Translation Card</u>

Punch 10 in columns 1-2.

Punch DX (inches)in columns 3-12.

Punch DY in columns 13-22.

Punch DZ in columns 23-32.

<u>Perspective Card</u>

Punch 08 in columns 1-2

Punch distance to viewing plane in columns 3-12. This should be larger than max (Z).

Punch distance to the object in columns 13-22.

<u>X-Rotation Card</u>

Punch 02 in columns 1-2

Punch degrees rotation relative to the X-axis in columns 3-12.

<u>Y-Rotation Card</u>

Punch 03 in columns 1-2

Punch degrees rotation relative to the Y-axis in columns 3-12.

<u>Z-Rotation Card</u>

Punch 04 in columns 1-2

Punch degrees rotation relative to the Z-axis in columns 3-12.

*<u>Termination Card</u>

Punch 13 in columns 1-2.

This terminates the entire program.

Programmer: F.J. Rens.

References: M.O. Dayhoff, "A Contour-Map Program for X-ray Crystallography," Communications, ACM, 6, 10 (October, 1963), pp. 620-622.

H.R. Puckett, "Computer Method for Perspective Drawing," Journal of Spacecraft and Rockets, 1,1(January 1964), pp. 44-48.

R. Osborn, " An Automated Procedure for Producing Contour Charts", IM No. 67-4, Oceanographic Office, Washington D.C.

B. Kubert, "Two Computer Programs for Contour Plotting", Aerospace Corp., 1969. AD 682307.

G. McCue, H. DuPrie, " Improved Fortran IV Function Contouring Program", North Am. Aviation, SID-65-672, 1965.

D.L. Light, "Ranger Mapping by Analytics", Photogrammetric Engineering, XXXII, (1966), pp. 792-800.

R.F. Walters, "Contouring by Machine", Bulletin Am. Assn. Petroleum Geologists, V. 53, No. 11 (1969) pp. 2343-2340

B. Bengtsson, S. Nordbeck, "Construction of Isarithms and Isarithmic Maps by Computers", BIT, 4 (1964), pp. 87-105.

D. McIntyre, D. Pollard, R. Smith, "Computer Programs for Automatic Contouring", Kansas Geological Survey Computer Contribution #22, 1968.

W. Tobler, "l'automation dans la préparation des cartes thématiques", International Yearbook of Cartography, 1966, pp

A. Noll, "Sterographic Projections by Computer", Computers and Automation, May 1965, pp. 32-34.

```
C      CONTOUR-PERSPECTIVE PLOTTING / FRANK J RENS / GEOG..COMP.CNTR. / 1
C     2 IS USED AS A SCRATCH FILE
C
       COMMON MT,NT,NI,IX,IY,IDX,IDY,ISS,IT,IV,NP,N,JT,PY,REC,CV,X,Y,IPT,
      1INX,INY,DL,AM,IOP,DP,DO,ILAB,XL,YL,ZL,LAB,DX,DY,DZ,XMAX,HT,SCALE,Y
      2MAX,CTHETA,STHETA,FM2,FNN,NN,M,CL,D,NLINES,NCURV,Z
       COMMON TIT,IZX
       DIMENSION AM(100,100), REC(800), X(1500), Y(1500), IPT(3,3), INX(8
      1), INY(8), IOP(23), XL(10), YL(10), ZL(10), LAB(10,6), DX(2), DY(2
      2), DZ(2), HT(10), CTHETA(4), STHETA(4)
       DIMENSION TEMP(6), FMT(18)
       DIMENSION Z(1500)
       DIMENSION TIT(18)
       DIMENSION CFMT(18), VCON(20)
       CALL TIME(0)
       CALL PLTXMX(30.0)
       NPLOT=1
       ISTAP=0
       IOP(1)=0
       IOP(23)=0
       D=0.
1      ILAB=0
       CALL REWIND(2)
       DO 2 I=2,20
2      IOP(I)=0
       NBOUND=0
C      READ CONTROL INSTRUCTIONS ACCORDING TO WRITE UP
3      READ (7,78) NCON,(TEMP(I),I=1,6)
       WRITE (6,79) NCON,(TEMP(I),I=1,6)
       IF(NCON.EQ.1) NLINES=0
       IF(NCON.GT.14.AND.NCON.LT.19) NLINES=0
C*    CALL CPER(NCON)
       IF (NCON) 74,74,4
4      IF (NCON-23) 5,5,74
5      GO TO (6,21,21,21,22,23,24,25,25,28,29,30,31,32,36,37,38,39,40,41,
      161,62,63), NCON
C      NCON=1
 6     CONTINUE
       CALL TIME(3,1)
C     M= # OF ROWS        NN= # OF COLS
       M=TEMP(1)
       NN=TEMP(2)
       TEMP(2)=TEMP(2)-1.
       SCALE=TEMP(2)/TEMP(3)
       IF (TEMP(4)) 8,7,8
7      TEMP(4)=1.
8      ZSCALE=TEMP(4)
       FM2=FLOAT(M)/2.
       FNN=FLOAT(NN)/2.
       XMAX=(FNN/SCALE)+2.5
       YMAX=(FM2/SCALE)+2.5
       IF (XMAX-14.) 9,9,70
9      IF (YMAX-14.) 10,70,70
10     IF (IOP(1)) 11,12,11
11     IOP(15)=1
```

```
      ICPV=1
      READ (7,85)  CFMT
      READ (7,CFMT) (VCON(I),I=1,NVCON)
      WRITE (6,84) (CFMT(I),I=1,18),(VCON(I),I=1,NVCON)
      GO TO 3
C     NCON=18
39    NVCON=TEMP(1)
      IOPV=1
      READ (7,85)  CFMT
      READ (7,CFMT) (VCON(I),I=1,NVCON)
      WRITE (6,84) (CFMT(I),I=1,18),(VCON(I),I=1,NVCON)
      GO TO 16
C     NCON=19
C     READ FORMAT FOR MATRIX
40    READ (7,85)  FMT
      WRITE (6,86)  FMT
      IOP(19)=1
      GO TO 3
C     NCON=20
C     BEGIN PLOTTING
 41   CONTINUE
      CALL TIME(3,1)
      IF (IOP(23)) 64,42,64
42    CONTINUE
43    IF (IOP(8)) 45,44,45
44    IOP(9)=1
45    IF (IOP(1)) 95,73,95
95    CONTINUE
C     SAFETY VALVE
      ASSIGN 46 TO NNL
      GO TO NNL, (53,46)
46    ASSIGN 53 TO NNL
C*  CALL SETEFL (NNL)
C     CALL DRAFT TO PLOT PERSPECTIVELY
      CBGN=-53139.E15
      NCURV=0
      NCOUNT=0
      NCMAX=0
      IF (NLINES) 52,52,47
47    READ(2,1000) N,CV
1000  FORMAT(I5,E12.4)
      IF (CBGN+53139.E15) 49,48,49
48    CBGN=CV
49    READ(2,1001)(X(I),Y(I),I=1,N)
1001  FORMAT(6E12.4)
      NCURV=NCURV+1
      NCOUNT=NCOUNT+N
      IF (N-NCMAX) 51,50,50
50    NCMAX=N
      CLMAX=CV
51    CV=CV/ZSCALE
      CALL DRAFT
      IF (NCURV-NLINES) 47,53,53
52    CBGN=0.
53    WRITE (6,88) NPLOT,NLINES,NCOUNT,NCMAX,CLMAX,CBGN
C     MODIFY BOUNDARY FOR MAP PROJECTION
```

```
      CALL TIME(3,1)
      GO TO 1
C     NCON = 2?
C     RED AND GREEN BUSINESS
63    CONTINUE
      FLABGT=PSYMLN(.2,99)
C     MTS SYSTEM DOES NOT INCORPORATE COLORED PENS AT THIS TIME - SUMMER
C     CALL PSYMB (.2,FLABGT+2.,-.2,99HPLEASE INSERT RED PEN NOW - AT NEX
C    1T 999 STOP INSERT THE GREEN PEN - AT 3RD 999 STOP INSERT BPBLK    ,
C    2270.,99)
C     CALL PLTSTP
      CALL PLTEND
      IOP(23)=3
      GO TO 3
64    IF (IOP(23)-3) 65,66,66
65    CONTINUE
66    IOP(23)=IOP(23)-1
      GO TO 43
67    IF (IOP(23)-1) 69,68,69
68    CONTINUE
      GO TO 62
69    GO TO 1
70    XMAX=XMAX*2.
      YMAX=YMAX*2.
      WRITE (6,90) YMAX,XMAX
      GO TO 75
71    WRITE (6,91)
      GO TO 75
72    WRITE (6,92)
      GO TO 75
73    WRITE (6,93)
      GO TO 75
74    WRITE (6,94) NCON
75    IF (NPLOT-1) 76,76,77
76    CALL PLTTRM (0)
77    CALL ERROR (0)
C
78    FORMAT (I2,6F10.0)
   79 FORMAT(1H0,I2,6F10.2)
80    FORMAT (1H0,(10F12.6))
   81 FORMAT(12A4)
   83 FORMAT(9A4)
   84 FORMAT(1H0,18A4//(10X,10E12.6))
   85 FORMAT(18A4)
86    FORMAT(70X,'THE FORMAT FOR THE DATA MATRIX IS : ',18A4)
88    FORMAT(' PLOT NUMBER ',I3,'  NUMBER OF LINES = ',I6,',NUMBER OF PO
     1INTS = ',I10/' MAXIMUM NUMBER OF POINTS = ',I10,' WHICH FORMED A C
     2URVE WITH CONTOUR LEVEL OF ',F10.2,', BOUNDARY AT ',F10.2)
90    FORMAT (14H4****A MAP OF ,F10.1,4H BY ,F10.1,14H WAS REQUESTED)
91    FORMAT (32H4****FORMAT NOT DEFINED FOR DATA)
92    FORMAT (31H4****CONTOUR LEVELS NOT DEFINED)
93    FORMAT (12H4****NO DATA)
94    FORMAT (27H4****NCON NOT LEGAL, NCON =,I5)
      END
C     CONTOUR-PERSPECTIVE PLOTTING / FRANK J RENS / GEOG.,COMP.CNTR. / 1
      SUBROUTINE SCAN
```

```
      IF (AM(MT,I)-CV) 7,9,9
7     IF (AM(MT,I+1)-CV) 9,8,8
8     IX=MT
      IY=I+1
      IDX=0
      IDY=-1
      CALL TRACE
9     CONTINUE
      DO 12 I=1,MT1
      MT2=MT+1-I
      IF (AM(MT2,NT)-CV) 10,12,12
10    IF (AM(MT2-1,NT)-CV) 12,11,11
11    IX=MT2-1
      IY=NT
      IDX=1
      IDY=0
      CALL TRACE
12    CONTINUE
      DO 15 I=1,NT1
      NT2=NT+1-I
      IF (AM(1,NT2)-CV) 13,15,15
13    IF (AM(1,NT2-1)-CV) 15,14,14
14    IX=1
      IY=NT2-1
      IDX=0
      IDY=1
      CALL TRACE
15    CONTINUE
      ISS=1
      NT1=NT-1
      MT1=MT-1
      DO 21 J=2,NT1
      DO 21 I=1,MT1
      IF (AM(I,J)-CV) 16,21,21
16    IF (AM(I+1,J)-CV) 21,17,17
17    COM=100*(I+1)+J
      IF (NP) 18,20,18
18    DO 19 ID=1,NP
      IF (REC(ID)-COM) 19,21,19
19    CONTINUE
20    IX=I+1
      IY=J
      IDX=-1
      IDY=0
      CALL TRACE
21    CONTINUE
      RETURN
      END
C     CONTOUR-PERSPECTIVE PLOTTING / FRANK J RENS / GEOG.,COMP.CNTR. /
      SUBROUTINE TRACE
C     ADAPTED FOR U OF M SYSTEM BY FRANK J RENS
      COMMON MT,NT,NI,IX,IY,IDX,IDY,ISS,IT,IV,NP,N,JT,PY,REC,CV,X,Y,IPT,
     1INX,INY,DL,AM,IOP,DP,DO,ILAB,XL,YL,ZL,LAB,DX,DY,DZ,XMAX,HT,SCALE,Y
     2MAX,CTHETA,STHETA,FM2,FNN,NN,M,CL,D,NLINES,NCURV,Z
      COMMON TIT
      DIMENSION TIT(18)
```

```
      PY=2.0
      CALL CALC
      IX=IX+IDX
      GC TO 25
24    IY=IY+IDY
      IDY=-IDY
      PY=2.0
      CALL CALC
      IY=IY+IDY
25    IF (AM(IX-1,IY)-CV) 26,27,27
26    NP=NP+1
      REC(NP)=100*IX+IY
27    IS=IS+5
      IX=IX2
      IY=IY2
      GO TO 6
28    XT=MT
      IF (AM(IX-1,IY)-CV) 29,30,30
29    NP=NP+1
      REC(NP)=100*IX+IY
30    NLINES=NLINES+1
      DO 31 K=1,N
      X(K)=X(K)-FNN
      Y(K)=Y(K)-FM2
31    CONTINUE
C     STORE CURVE ON TAPE 2  (B2)
      IF (N) 33,33,32
32    WRITE(2,1000) N,CL
1000  FORMAT(I5,E12.4)
      WRITE(2,1001)(X(I),Y(I),I=1,N)
1001  FORMAT(6E12.4)
33    N=-1
      RETURN
C
   34 FORMAT('0ACONTOURATLEVEL',E11.5,'WASTERMINATEDATX=',F5.1,'Y=',F5.1
     1,'BECAUSEITCONTAINEDMORETHAN1500PLOTPOINTS')
      END
C     CONTOUR-PERSPECTIVE PLOTTING / FRANK J RENS / GEOG.,COMP.CNTR. / 1
      SUBROUTINE CALC
C     ADAPTED FOR U OF M SYSTEM BY FRANK J RENS
      COMMON MT,NT,NI,IX,IY,IDX,IDY,ISS,IT,IV,NP,N,JT,PY,REC,CV,X,Y,IPT,
     1INX,INY,DL,AM,IOP,DP,DO,ILAB,XL,YL,ZL,LAB,DX,DY,DZ,XMAX,HT,SCALE,Y
     2MAX,CTHETA,STHETA,FM2,FNN,NN,M,CL,D,NLINES,NCURV,Z
      COMMON TIT
      DIMENSION TIT(18)
      DIMENSION AM(100,100), REC(800), X(1500), Y(1500), IPT(3,3), INX(8
     1), INY(8), IOP(23), XL(10), YL(10), ZL(10), LAB(10,6), DX(2), DY(2
     2), DZ(2), HT(10), CTHETA(4), STHETA(4)
      DIMENSION Z(1500)
      IT=0
      N=N+1
      IF (IDX**2+IDY**2-1) 4,1,4
1     IF (IDX) 3,2,3
    2 X(N)=IX
      Z(N)=IY
      IY2=IY+IDY
```

```
      DY(2)=IDY
      Y(N)=((AM(IX,IY)-CV)/(AM(IX,IY)-AM(IX,IY2)))*DY(2)+Z(N)
      RETURN
3     Y(N)=IY
      W=IX
      DX(2)=IDX
      IX2=IX+IDX
      X(N)=((AM(IX,IY)-CV)/(AM(IX,IY)-AM(IX2,IY)))*DX(2)+W
      RETURN
4     IX2=IX+IDX
      IY2=IY+IDY
      W=IX
      Z(N)=IY
      DX(2)=IDX
      DY(2)=IDY
      DCP=(AM(IX,IY)+AM(IX2,IY)+AM(IX,IY2)+AM(IX2,IY2))/4.0
      IF (PY-2.0) 5,6,5
5     IF (DCP-CV) 6,6,7
6     AL=AM(IX,IY)-DCP
      V=.5*(AL+DCP-CV)/AL
      X(N)=V*DX(2)+W
      Y(N)=V*DY(2)+Z(N)
      PY=0.0
      RETURN
7     IT=1
      AL=AM(IX2,IY2)-DCP
      V=.5*(AL+DCP-CV)/AL
      X(N)=-V*DX(2)+W+DX(2)
      Y(N)=-V*DY(2)+Z(N)+DY(2)
      RETURN
      END
C     CONTOUR-PERSPECTIVE PLOTTING / FRANK J RENS / GEOG..COMP.CNTR. / 1
      SUBROUTINE CRAFT
C     WRITTEN BY FRANK J RENS/ GEOGRAPHY / 1966
      COMMON MT,NT,NI,IX,IY,IDX,IDY,ISS,IT,IV,NP,N,JT,PY,REC,CV,X,Y,IPT,
     1INX,INY,DL,AM,IOP,DP,DO,ILAB,XL,YL,ZL,LAB,DX,DY,DZ,XMAX,HT,SCALE,Y
     2MAX,CTHETA,STHETA,FM2,FNN,NN,M,CL,D,NLINES,NCURV,Z
      COMMON TIT
      DIMENSION TIT(18)
      DIMENSION AM(100,100), REC(800), X(1500), Y(1500), IPT(3,3), INX(8
     1), INY(8), IOP(23), XL(10), YL(10), ZL(10), LAB(10,6), DX(2), DY(2
     2), DZ(2), HT(10), CTHETA(4), STHETA(4)
      DIMENSION Z(1500)
      DIMENSION IRES(3)
1     KTYPE=3
      DO 21 I=2,14
      IF (IOP(I)) 21,21,2
2     GO TO (3,3,4,5,21,21,21,8,8,19,21,21,21,14), I
C     ROTATION ABOUT X
3     D1=1.
      D2=0.
      D3=0.
      D4=0.
      D5=CTHETA(2)
      D8=STHETA(2)
      D7=0.
```

```
      GO TO (16,15), NN1
15    E=(DO-ZL(J))/DP*SCALE
      F=(DO/DP)*SCALE
      C1=(XL(J)-FNN)/E+XMAX
      C2=(YL(J)-FM2)/E+YMAX
      GO TO 17
16    C1=XL(J)/SCALE+XMAX
      C2=YL(J)/SCALE+YMAX
      F=SCALE
17    HT(J)=HT(J)/F
18    CALL PSYMB (C2,C1,HT(J),LAB(1,J),90.0,36)
      IOP(14)=0
      GO TO 21
C     TRANSLATE DATA
19    DO 20 J=1,N
      X(J)=X(J)+DX(1)
      Y(J)=Y(J)+DY(1)
20    Z(J)=Z(J)+DZ(1)
C     PLOT PLANE OR PERSPECTIVE VIEW OF SURFACE
21    CONTINUE
      IOPC= IOP(2)+IOP(3)+IOP(4)
      DO 29 J=1,N
      IF (IOPC) 22,22,23
22    Z(J)=CV
23    GO TO (24,25), NN1
24    C1=X(J)/SCALE+XMAX
      C2=Y(J)/SCALE+YMAX
      GO TO 26
25    E=(DO-Z(J))/DP*SCALE
      C1=X(J)/E+XMAX
      C2=Y(J)/E+YMAX
26    IF (IOP(5)) 28,28,27
27    CALL PROJ (C1,C2,XMAX,YMAX,SCALE)
28    CALL PLOTCC (C2,C1,KTYPE)
      KTYPE=2
29    CONTINUE
      RETURN
   30 FORMAT('0      DISTANCE TO PLANE =',F9.2,', DISTANCE TO OBJECT = '
     1,F9.2)
      END
      SUBROUTINE PROJ (X,Y,XMAX,YMAX,SCALE)
      DATA K/'1'/
C     MAP PROJECTION SUBROUTINE
C     OBLIQUE STEREOGRAPHIC PROJECTION
C     INSERT CARD CONTAINING PROJECTION CONSTANTS IMMEDIATELY AFTER
C     "20" CARD WHICH FOLLOWS THE FIRST USE OF A "05" CARD.
C     PROJECTION CONSTANTS MUST INCLUDE THE LAT/LONG OF THE CENTER OF
C     THE MATRIX TO BE CONTOURED AND THE GRID SIZE IF DIFFERENT FROM
C     1 DEGREE. OTHER PROJECTION CONTROLS DEPEND ON THE SPECIFIC
C     PROJECTION SUBROUTINE. ALL DATA USING "05" IN ONE RUN SHOULD
C     HAVE THE SAME CENTER LAT/LON AND GRID SIZE, AND WILL PLOT ON THE
C     SAME PROJECTION.
C     WIDTH OF MAP WHEN USING "05" SHOULD REFER TO THE WIDTH AT THE
C     CENTER LATITUDE ON THE SQUARE PROJECTION (ASSUMES NORTH
C     ORIENTATION).
      IF (K) 1,4,1
```

```
15.115.1
12.011.611.913.114.214.713.813.313.413.213.111.510.310.808.808.210.215.9
15.717.2
13.715.115.814.814.915.616.516.616.413.915.613.514.714.216.113.211.814.2
10.711.7
15.416.415.917.615.717.217.917.717.916.817.114.316.814.613.914.216.416.3
17.916.8
18.217.117.917.416.916.719.318.018.416.814.616.417.417.717.115.409.416.9
15.013.7
18.818.717.818.217.518.419.018.418.617.917.818.118.718.418.917.817.318.3
17.818.4
18.718.218.018.017.118.318.218.018.418.117.917.918.018.217.617.617.818.2
18.218.1
12
TEMP AT 200M*SAME/FOLIO2/PLATE2/MP6
20
  +40.0  -54.0    1.0  +54.0  -38.0
13
```

BLOCK DIAGRAM PLOTTING

Purpose: This program is used for rapid plotting of isometric profiles from a rectangular matrix of data.

Description: The block diagrams are drawn from a ROWS x COLS matrix of floating point Z values. Data is read in row by row from the top down (i.e., the origin is (1,1) in the upper left, or Northwest). By simply changing the ROWS parameter to a minus number, the block diagram will be drawn as viewed from the North. Controls and format are read from unit 5; the title and observation deck from unit 4, and plotting is done via unit 9. The program is easy to use since it has many default options, the only essential control variables being ROWS and COLS. This version of the program does not delete hidden lines.

Comments: The program is designed for use with 30 inch CALCOMP plotter. Current dimensioning allows an input matrix of 100 x 100; this may be adjusted by changing the dimensioning of X, Y, XL, and YL. Multiple sets of data may be run using the same vertical exaggeration by setting SAME parameter. Plotting calls are from the standard CALCOMP library.

Data Deck:
1) Control card
2) Variable Format card
3) Title card
4) Observation deck

Data Card
Composition: 1) Control card

columns		
1-3	ROWS	Number of rows in input matrix (100 or less). If negative, -ROWS is the number of rows, and a view from the North is plotted. (I3)
4-6	COLS	Number of columns in input matrix (100 or less) (I3)
7-9	RINC	Spacing increment of row profiles Default = 1. (I3)
10-12	CINC	Spacing increment of column profiles. Default = Cols -1, i.e., only the edges are drawn (I3)
13-15	SAME	Number of matrices of input to be run with same controls and format, and same vertical exaggeration. Default = 0 (I3)
16-20	MAXHGT	Floating point height of diagram in inches. Default = 0.5 inches (F5.0)
21-25	MAX	
26-30	MIN	Control vertical exaggeration, in floating point (Default = MAX and MIN of data) (F5.0)
31-35	WIDTH	
36-40	FRONT	Specify horizontal size of diagram in floating point. Default uses EPS to calculate these val-

		ues. (F5.0)
41-45	EPS	Inches between columns of diagram. floating point. Default = 0.1 inches. (F5.0)
46-50	ALF	Isometric viewing angle in degrees floating point. Default = 35. (F5.0)

2) Variable format card

E or F type FORTRAN format for one row of data

3) Title card

Any title (≤ 72 characters)

4) Observation Deck

Observations by rows as described by variable format card.

References: B. Kubert, J. Szabo, and S. Giulieri, "The Perspective Representation of Functions of Two Variables", Journal, ACM, 15,2 (April 1968), pp. 193-204.

R.L. Mitchell, "A Computerized 3-D Plotting Program", Los Angeles, 1967, AD 658857.

```
C .PLOCK DIAGRAM PLOTTING PROGRAM
C C FORTRAN TRANSLATION OF A 1965 MAD PROGRAM BY W.R. TOBLER
            DIMENSION FMT(18),TITL(18),X(100,100),Y(100,100),XL(5000),YL
     1(5000)
      INTEGER ROWS,COLS,RINC,CINC.SAME
      INTEGER ROWS1
      REAL MAXHGT,MAX,MIN
999   FORMAT(5I3,7F5.0)
1000  FORMAT(18A4)
1001  FORMAT (2X,18A4)
1003  FORMAT (2X,5I4,2X,6(F10.3,2X))
1004  FORMAT(2X,F10.3)
1005  FORMAT(1H ,I5)
      RAD=174532925E-10
5     READ(5,999) ROWS,COLS,RINC,CINC,SAME,MAXHGT,MAX,MIN,WIDTH,FRONT,
     1EPS,ALF
C READS FORMAT FOR ONE ROW OF DATA MATRIX
      READ(5,1000) FMT
      WRITE(6,1001) (FMT(I),I=1,18)
      IVU =0
      IF (ROWS) 500,500,501
 500  IVU=1
      ROWS=-ROWS
      ICOLS=COLS+1
 501  GRID=ROWS*COLS
      IF(RINC) 50,50,52
50    IF(CINC) 51,51,52
51    RINC=1
52    KK=0
      IF (ALF) 56,55,56
55    ALF=35.0
56    IF(ALF-90) 57,57,55
57    ALF = ALF*RAD
      SALF = SIN(ALF)
      CALF = COS(ALF)
10    CONTINUE
      READ(4,1000)(TITL(I),I=1,18)
      WRITE(6,1001) (TITL(I),I=1,18)
      WRITE(6,1005)SAME
      CALL PLTXMX(30.0)
      ROWS1 = ROWS+1
      DO 60 II=1,ROWS
      I = ROWS1-II
C READS DATA MATRIX
      IF (IVU ) 502,502,503
 502  READ (4,FMT) (X(I,J),J=1,COLS)
      GO TO 505
 503   READ(4,FMT) (Y(II,J),J=1,COLS)
      DO 504 J=1,COLS
      JJ=ICOLS-J
 504  X(II,JJ)=Y(II,J)
 505  IF (I-1) 3001,3000,3001
3000  WRITE(6,1003)ROWS,I,II,ROWS1,SAME,X(I,1),X(I,2),
     1X(I,3),X(I,4),X(I,5),X(I,6)
3001  IF(I-ROWS)3003,3002,3003
```

```
3002   WRITE(6,1003)ROWS,I,II,ROWS1,SAME,X(I,1),X(I,2),
      1X(I,3),X(I,4),X(I,5),X(I,6)
3003   CONTINUE
60     CONTINUE
       IF(RINC) 61,61,62
61     RINC = ROWS-1
62     IF(CINC) 63,63,64
63     CINC = COLS-1
64     IF(SAME) 67,67,65
65     IF(KK) 66,66,20
66     KK = KK+1
67     CALL STATS(ROWS,COLS,X,ZBAR,ZVAR,ZSIG,ZMAX,ZMIN,ROWS1)
       IF(MAX-MIN) 68,69,68
68     ZMAX = MAX
       ZMIN = MIN
C      HORIZONTAL SCALES
69     IF(WIDTH) 70,71,70
70     EPS = (WIDTH-1)/(COLS-1+(ROWS-1)*CALF)
       GO TO 75
71     IF (EPS) 75, 72, 75
72     IF(FRONT) 73,74,73
73     EPS = FRONT/(COLS-1)
       GO TO 75
74     EPS = 0.1
75     WIDTH = EPS*((COLS-1)+(ROWS-1)*CALF)+1.0
       FRONT = EPS*(COLS-1)
C      VERTICAL SCALES
       RANGE = ZMAX-ZMIN
       IF(MAXHGT) 76,76,77
76     MAXHGT=0.5
77     DZ = MAXHGT/RANGE
       XCON=0.5
       YCON=1.5
       SUM=YCON-0.1
       SF = SALF*EPS
20     CONTINUE
       WRITE (6,1003) ROWS,COLS,RINC,CINC,SAME,FRONT,EPS,SALF,MAXHGT,
      1 ZMAX,ZMIN
       CALLPSYMB(0.5,0.1,.1,' GEOGRAPHY DEPARTMENT, UNIVERSITY OF MICHIGA
      1N    ',0.0,45)
       CALL PSYMB(0.5,0.5,.1,TITL,0.,72)
C      BEGIN PLOTTING
C      COMPUTE THE DIAGRAM LATTICE POINT COORDINATES
       DO 30 I=1,ROWS
       M = I-1
       RC1 = SF*M+YCON
       RC2 = M*CALF
       DO 30 J=1,COLS
       Y(I,J)=(X(I,J)-ZMIN)*DZ+RC1
       X(I,J) = EPS*((J-1)+RC2)+XCON
30     CONTINUE
       BOT=0.0
        WRITE (6,1004) BOT
C      DRAW PROFILES ACROSS THE ROWS
       DO 40 I=1,ROWS,RINC
         DO 2000 J=1,COLS
```

```
      XL(J)=X(I,J)
2000  YL(J)=Y(I,J)
      CALL PLINE (XL(1),YL(1),COLS,1,0,0,0)
40    CONTINUE
        BCT=1.0
       WRITE (6,1004) BOT
C     DRAW PROFILES DOWN THE COLUMNS
      DO 42 J=1,COLS,CINC
      DO 41 I=1,ROWS
      XL(I) = X(I,J)
      YL(I) = Y(I,J)
41    CONTINUE
      CALL PLINE(XL(1),YL(1),ROWS,1,0,0,0)
42    CONTINUE
      BOT=BOT+2.0
       WRITE (6,1004) BOT
C     PUT BOTTOM ON THE DIAGRAM
      XL(1)=XCON
      YL(1)=Y(1,1)
      XL(2)=XCON
      YL(2)=SUM
      XL(3)=X(1,COLS)
      YL(3)=SUM
      XL(4)=XL(3)
      YL(4)=Y(1,COLS)
      CALL PLINE (XL(1),YL(1),4,1,0,0,0)
      XL(1)=X(1,COLS)
      XL(2)=X(ROWS,COLS)
      YL(3)=Y(ROWS,COLS)
      YL(1)=SUM
       YL(2)=RC1-0.1
      XL(3)=XL(2)
      CALL PLINE(XL(1),YL(1),3,1,0,0,0)
      BCT=BOT+4.0
       WRITE (6,1004) BOT
      SAME = SAME-1
      CALL PLTEND
      WRITE(6,114)(TITL(I),I=1,18)
114    FORMAT(1H1,18A4)
      IF(SAME) 5,5,10
      END
      SUBROUTINESTATS(NR,NC,X,XBAR,XVAR,XSIG,XMAX,XMIN,PRINT)
      DIMENSION X(100,100)
      INTEGER PRINT
200   FORMAT(1H ,2(I5,1X),5(F10.3,1X))
      FN=NR*NC
      XMAX=X(1,1)
      XMIN=XMAX
      XSUM=0.
      XSQR=0.
      DO10I=1,NR
      DO10J=1,NC
      XX=X(I,J)
      XMAX=AMAX1(XMAX,XX)
      XMIN=AMIN1(XMIN,XX)
      XSUM=XSUM+XX
```

```
10  XSQR=XSQR+XX*XX
    XBAR=XSUM/FN
    XVAR=(XSQR-2.*XBAR*XSUM+FN*XBAR*XBAR)/FN
    XSIG=SQRT(XVAR)
    IF(PRINT)30,30,20
20  WRITE(6,200)NR,NC,XBAR,XVAR,XSIG,XMAX,XMIN
30  RETURN
    END
```

SPATIAL DERIVATIVE PROGRAM

Purpose: This program calculates the absolute value of the gradient ("slope") for data given as a geographical matrix.

Description: The spatial derivative is calculated by the equation

$$\nabla(x,y) = \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}$$

Where $\nabla(x,y)$ is the magnitude of the gradient obtained from a pair of orthogonal partial derivatives of Z, approximated by finite differences. That is, ∇_{ij} is calculated from its four neighbors ($Z_{i\pm1,\ j\pm1}$). Edge effects appear at the boundaries, where the three neighboring points are used, and in the corners where the two neighboring points are used in the calculations. The program will calculate as high as nine derivatives. The results may also be punched out.

Comments: Currently the program will accomodate an input matrix of Z values of 100 x 100. This may be adjusted by changing the dimensioning of values Z and W. This program may be used at a remote terminal which has access to shared files with no difficulty. Punched output may be used directly (with proper control cards) in the block profile and contour mapping programs.

Data Deck:
1) Control Card
2) Title Card
3) Variable Format Card
4) Observation Deck

Data Card Composition: 1) Control Card

columns

1-3	ROWS	Number of rows of input matrix
4-6	COLS	Number of columns of input matrix
7	G	One for first derivative, two for second derivative, etc.
8	PUN	One if any derivatives are to be punched out, zero otherwise
9-14	SCALE	Distance between matrix points (F6.0)
15	SPUN	If zero punches out all derivatives (PUN =1), if 1-9 will only punch out derivative number indicated (convenient for remote terminal usage).

2) Title Card

Any title (≤ 80 characters)

3) Variable Format Card

E or F type FORTRAN format for one row of data (≤ 80 characters)

4) Observation Deck

Observations by rows as described by variable format card.

Reference: C.M. Davis, "A Study of the Land Type", University of Michigan, Contract DA-31-124-ARO-D-456, UM 08055-2-F, March 1969, pp. 59-85.

```
C PROGRAM TO COMPUTE THE K'TH(MAX 9) DERIVATIVE AT GRID POINTS
C W.R. TOBLER / UNIVERSITY OF MICHIGAN / GEOGRAPHY
C ADAPTED TO FORTRAN BY H. MOELLERING
      INTEGER ROWS,COLS,G,PUN,SPUN
      DIMENSION TITL(20),FMT(20),Z(100,100),W(100,100),DIV(9)
      DATA DIV/'1ST','2ND','3RD','4TH','5TH','6TH','7TH','8TH','9TH'/
  100 FORMAT(2I3,2I1,F6.0,I1)
  101 FORMAT(20A4)
  102 FORMAT(1H1,20A4)
  103 FORMAT('0','N=',I5,5X,'ROWS=',I3,5X,'COLS=',I3,5X,'K=',I2,5X,
     1'SCALE=',F6.0,5X,'G=',I2,2X,'SPUN=',I2///55X,'THE',1X,A3,1X,
     2'DERIVATIVE IS')
  104 FORMAT(8F10.3)
  105 FORMAT(10X,'END OF DATA SET')
  106 FORMAT(1H0,' END OF DERIVATIVES')
  108 FORMAT('0',12F10.3/(' ',12F10.3))
  109 FORMAT('1',20A4)
  110 FORMAT ('0  PROGRAM IS PUNCHING OUT RESULTS OF ALL DERIVATIVES REQ
     1UESTED')
  111 FORMAT('0  PROGRAM IS PUNCHING OUT RESULTS OF ONLY DERIVATIVE #',
     1I2)
    5 READ(5,100) ROWS,COLS,G,PUN,SCALE,SPUN
      READ(5,101)  TITL
      WRITE(6,109)  TITL
      READ(5,101)  FMT
      WRITE (6,101)  FMT
      IF(SCALE.LE.0.0) XH2=2.0
      IF(SCALE.GT.0.0) XH2=SCALE*2.0
      IF(G.LE.0) G=1
      DO 70 I=1,ROWS
   70 READ (5,FMT)  (Z(I,K),K=1,COLS)
      DO 60 I=1,ROWS
   60 WRITE(6,108) (Z(I,K),K=1,COLS)
      CALL STATS(ROWS,COLS,Z,ZBAR,DUM,ZSIG,ZMAX,ZMIN,K)
      IF(PUN.GE.1 .AND. SPUN.EQ.0) WRITE(6,110)
      IF(PUN.GE.1 .AND. SPUN.GT.0) WRITE(6,111)  SPUN
      DO 50 K=1,G
      N=ROWS*COLS
      WRITE(6,102)  TITL
      WRITE(6,103) N,ROWS,COLS,K,SCALE,G,SPUN,DIV(K)
      DO 30 M=1,ROWS
      MM=M-1
      M1=M+1
      DO 20 J=1,COLS
      JJ=J-1
      J1=J+1
      IF(JJ.LT.1) DX=((Z(M,J1)-Z(M,J))*2.0)/XH2
      IF(J1.GT.COLS) DX=((Z(M,J)-Z(M,JJ))*2.0)/XH2
      IF(JJ.GE.1 .AND. J1.LE.COLS) DX=(Z(M,J1)-Z( M,JJ))/XH2
      IF(MM.LT.1) DY=((Z(M,J)-Z(M1,J))*2.0)/XH2
      IF(M1.GT.ROWS) DY=((Z(MM,J)-Z(M,J))*2.0)/XH2
      IF(MM.GE.1 .AND. M1.LE.ROWS) DY=(Z(MM,J)-Z(M1,J))/XH2
      W(M,J)=SQRT(DX*DX+DY*DY)
   20 CONTINUE
      IF(PUN.GE.1 .AND. SPUN.EQ.0) WRITE(7,104) (W(M,J),J=1,COLS)
```

TREND SURFACES BY EIGENVECTOR DYADS

Purpose: This program produces a series of trend surfaces each consecutive member of which better approximates an original input distribution.

Description: It is assumed that the geographical data are given in the form of an m by n geographical matrix $Z = (z_{ij})$. This matrix is decomposed by the program into the sum of k matrices of order m by n. Each of these trend matrices is in turn separated into the product of an m by one vector f_k multiplied by a one by n vector g_k, i.e., the model is

$$Z = f_1(x)g_1(y) + \dots + f_k(x)g_k(y)$$

where x can be interpreted as the column coordinate j, and y the row coordinate. The column vector f_k is the k^{th} eigenvector of ZZ^t, and the row vector g_k is obtained as $g_k = Z^t f_k / f_k^t f_k$. The program orders the dyads (f, g) from the largest eigenvalue to the smallest, each dyad accounting for the maximum amount of variance which can be extracted by any product of the form f(x)g(y). The number of eigenvector dyads required for a complete representation of Z is at most the lesser of m,n. The current dimensioning allows up to m = n = 70. The program reads from 7, writes on 6, and punches on 5.

Data Deck: 1) Control Card.

2) Format Card.

3) Observation deck.

Repeat 1 through 3 as many times as desired.

Data Card Composition: 1) Control Card

column

1-2	NR	number of rows
3-4	NC	number of columns
5-6	MAXE	Maximum number of dyads to be extracted
7	IFPCH	2 if for each dyad the X matrix and the cumulative Y matrix are to be punched. 1 if for each dyad only the cumulative Y matrix is to be punched. 0 if no punching.

2) Format Card

The variable format, including left and right parentheses, is punched in the first 72 columns of this card in E or F type FORTRAN notation.

References: Waldo R. Tobler, "Geographical Filters and their Inverses," Geographical Analysis, Vol 1, (1969), pp. 243-253.

Peter R. Gould, "On the Geographical Interpretation of Eigenvalues," Transactions, Institute of British Geographers, (1967), pp. 53-86.

```
C     TREND SURFACES VIA EIGENVECTOR DYADS
      DIMENSION Z(70,70), ZT(70,70), ZZT(70,70), FMT(18), EVAL(70), A(70
     1), B(70), EVEC(70,70)
31    FORMAT (3I2,I1)
32    FORMAT (18A4)
33    FORMAT (10H1CONTROLS ,4(2X,I2)/1H 18A4)
34    FORMAT ('ORIGINAL MATRIX')
35    FORMAT (1H 10F10.2/)
36    FORMAT ('-EIGENVALUES'/'TRACE=',E15.6,10X,'N=',I5/'VECTOR',10X,
     1'VALUE',10X,'CUMULATIVE PERCENT'/1H0)
37    FORMAT (1H0,I2,E15.6,2PF24.2)
38    FORMAT ('1 DYAD NUMBER',I5)
   39 FORMAT (1H0,'EIGENVECTOR TRANSPOSED')
   40 FORMAT (1H0,'THE OTHER VECTOR')
   41 FORMAT (1H0,'THE PRODUCT MATRIX')
42    FORMAT ('1 CUMULATIVE EXPANSION, THRU DYAD',I5)
   43 FORMAT (1H-,'EXPANSION COMPLETED')
   44 FORMAT (1H0,34HMATRIX DIMENSIONS EXCEED ALLOTMENT)
1     READ (7,31) NR,NC,MAXE,IFPCH
      READ (7,32) (FMT(J),J=1,18)
      WRITE (6,33) NR,NC,MAXE,IFPCH,(FMT(K),K=1,18)
      WRITE (6,34)
      IF (MAXO(NR,NC) .GT. 70)  GO TO 51
C     READ AND PRINT DATA MATRIX
      DO 2 I=1,NR
      READ (7,FMT) (Z(I,J),J=1,NC)
2     WRITE (6,35)  (Z(I,J),J=1,NC)
C     FORM TRANSPOSE
      DO 3 I=1,NR
      DO 3 J=1,NC
3     ZT(J,I)=Z(I,J)
      DO 4 I=1,NR
      DO 4 J=1,NR
      ZZT(I,J)=0.0
4     EVEC(I,J)=0.0
C     FORM PRODUCT
      DO 5 K=1,NC
      DO 5 I=1,NR
      DO 5 J=1,NR
5     ZZT(I,J)=ZZT(I,J)+Z(I,K)*ZT(K,J)
      MAXE = MINO(NR,NC)
11    TRACE=0.0
      DO 12 K=1,NR
12    TRACE=TRACE+ZZT(K,K)
C     PRINT EIGENVALUES
      WRITE (6,36) TRACE,MAXE
      PERC=0
      CALL EIGEN (ZZT,EVEC,NR,EVAL,MAXE,70)
      DO 15 K=1,MAXE
      IF (TRACE) 13,14,13
13    PERC=PERC+EVAL(K)/TRACE
      WRITE (6,37) K,EVAL(K),PERC
      GO TO 15
14    WRITE (6,37) K,EVAL(K)
15    CONTINUE
```

```
      DIMENSION A(MD,MC), B(MD,MD), VALU(MD), T(70,3), DIAG(70), SUPERD(1
     170), WVEC(70), PVEC(70),QVEC(70), VALL(70), C(70), U(70), INDEX(70
     2), V(70)
C
      EQUIVALENCE (WVEC,VALL,U),                (PVEC,QVEC,C,V), (I1,T1),
     1(I2,T2),        (TEMP,T0), (SUM,MATCH), (I,P), (DIV,SCALAR,TAU), (AN
     2ORM2,ANORM), (VTEMP,VNORM2,VNORM)
C
C
C     INITIALIZATION
      N=NSUP
      M=MSUB
      NP1=N+1
      NM1=N-1
       E1=1.E-8
C     GENERATE IDENTITY MATRIX
      DO 3 I=1,N
      DO 3 J=1,N
      IF (I-J) 2,1,2
1     B(I,J)=1.
      GO TO 3
2     B(I,J)=0.
3     CONTINUE
C     HOUSEHOLDER SIMILARITY TRANSFORMATION TO CO-DIAGONAL FORM
C     REDUCE COLUMN OF MATRIX
      DO 14 I=1,NM1
      IF (I-NM1) 4,13,4
4     I1=I+1
      I2=I1+1
      SUM=0.
      DO 5 J=I2,N
5     SUM=SUM+A(J,I)**2
      IF (SUM) 6,13,6
6     J=I1
      TEMP=A(J,I)
C                                                             TERMINAL-F FLAG
      SUM=SQRT(SUM+TEMP**2)
      A(J,I)=-SIGN (SUM,TEMP)
C                                                             TERMINAL-F FLAG
      WVEC(J)=SQRT(1.+ABS (TEMP)/SUM)
      DIV=SIGN (WVEC(J)*SUM,TEMP)
      DO 7 J=I2,N
7     WVEC(J)=A(J,I)/DIV
      SCALAR=0.
      DO 9 J=I1,N
      PVEC(J)=0.
      DO 8 K=I1,N
8     PVEC(J)=PVEC(J)+A(K,J)*WVEC(K)
      SCALAR=SCALAR+PVEC(J)*WVEC(J)
9     CONTINUE
      SCALAR=SCALAR/2.
      DO 10 J=I1,N
      QVEC(J)=PVEC(J)-SCALAR*WVEC(J)
      DO 10 K=I1,J
      A(K,J)=A(K,J)-(WVEC(K)*QVEC(J)+WVEC(J)*QVEC(K))
      A(J,K)=A(K,J)
```

```
10    CONTINUE
C     SAVE ROTATION FOR LATER APPLICATION TO CO-DIAGONAL VECTORS
      DO 12 K=2,N
      TEMP=0.
      DO 11 J=I1,N
11    TEMP=TEMP+WVEC(J)*B(J,K)
      DO 12 J=I1,N
      B(J,K)=B(J,K)-WVEC(J)*TEMP
12    CONTINUE
C     MOVE CO-DIAGONAL FORM ELEMENTS FOR ITERATIVE PROCEDURE
13    J=I
      DIAG(I)=A(J,I)
      SUPERD(I)=A(J+1,I)
14    CONTINUE
      DIAG(N)=A(N,N)
C     DETERMINE EIGENVALUES FROM STURM CHAIN OF CO-DIAGONAL MINORS
C     CALCULATE NORM OF MATRIX AND INITIALIZE EIGENVALUE BOUNDS
      ANORM2=DIAG(1)**2
      DO 15 L=2,N
      Q(L-1)=SUPERD(L-1)**2
      ANORM2=DIAG(L)**2+Q(L-1)+Q(L-1)+ANORM2
15    CONTINUE
C                                                      TERMINAL-F FLAG
      ANORM=SQRT(ANORM2)
      DO 16 L=1,M
      VALU(L)=ANORM
      VALL(L)=-ANORM
16    CONTINUE
      EPS1=ANORM*E1
      IF (EPS1) 17,73,17
C     CHOOSE NEW TRIAL VALUE WHILE TESTING BOUNDS FOR CONVERGENCE
17    DO 35 L=1,M
      ITER=0
      VTEMP=EPS1
18    TAU=(VALU(L)+VALL(L))/2.
      IF (ITER-10)  20,19,20
19    VTEMP=VTEMP*10.
      ITER=0
20    IF (2.*(TAU-VALL(L))-VTEMP) 35,35,21
C     DETERMINE SIGNS OF PRINCIPAL MINORS
21    MATCH=0
      ITER=ITER+1
      T2=0.
      T1=1.
      DO 30 L1=1,N
      P=DIAG(L1)-TAU
      IF (T2) 23,22,23
22    T1=SIGN (1.,T1)
23    IF (T1) 25,24,25
24    T0=-SIGN (1.,T2)
      T2=0.
      IF (Q(L1-1)) 26,22,26
25    T0=P-Q(L1-1)*T2/T1
      T2=1.
C     COUNT AGREEMENTS IN SIGN (ZERO CONSIDERED POSITIVE)
26    IF (T0) 29,27,28
```

```
      ITER=1
      IF (I1) 58,51,58
C     BACK SUBSTITUTE TO OBTAIN EIGENVECTOR
51    DO 52 L1=1,N
      L=NP1-L1
      V(L)=(V(L)-T(L,2)*V(L+1)-T(L,3)*V(L+2))/T(L,1)
52    CONTINUE
      GO TO (53,58), ITER
C     PERFORM SECOND ITERATION
53    ITER=2
54    DO 57 L=2,N
      IF (INDEX(L-1))55,56,55
55    VTEMP=V(L-1)
      V(L-1)=V(L)
      V(L)=VTEMP
56    V(L)=V(L)-U(L-1)*V(L-1)
57    CONTINUE
      GO TO 51
C     ORTHOGONALIZE VECTOR TO OTHERS ASSOCIATED WITH REPEATED ROOT
58    IF (I1) 59,62,59
59    DO 61 L1=1,I1
      K=I-L1
      VTEMP=0.
      DO 60 J=1,N
60    VTEMP=VTEMP+A(J,K)*V(J)
      DO 61 J=1,N
61    V(J)=V(J)-A(J,K)*VTEMP
62    GO TO (54,63), ITER
C     NORMALIZE VECTOR TO UNIT LENGTH
63    VNORM2=0.
      SUM=0.
      DO 65 L=1,N
      IF (SUM-ABS (V(L))) 64,65,65
64    SUM=ABS (V(L))
65    CONTINUE
      DO 66 L=1,N
      V(L)=V(L)/SUM
66    VNORM2=VNORM2+V(L)**2
C                                                   TERMINAL-F FLAG
      VNORM=SQRT(VNORM2)
      DO 67 J=1,N
67    A(J,I)=V(J)/VNORM
68    CONTINUE
C     ROTATION OF CO-DIAGONAL VECTORS INTO MATRIX EIGENVECTORS
      DO 70 I=1,M
      DO 69 K=2,N
      U(K)=0.
      DO 69 J=2,N
69    U(K)=U(K)+B(J,K)*A(J,I)
      DO 70 J=2,N
70    A(J,I)=U(J)
C     NORMALIZE LENGTH OF VECTORS TO EIGENVALUES AND STORE IN B(I,J)
      DO 72 J=1,M
      IF (VALU(J)) 73,73,71
C                                                   TERMINAL-F FLAG
71    VTEMP=SQRT(VALU(J))
```

GEOGRAPHICAL GROUPING

Purpose: The program groups observations using the criterion of Euclidean proximity in a p dimensional vector space; if the data are given in the form of geographical matrices $Z = (z_{ij})$ the program can automatically impose a geographical contiguity constraint on the grouping.

Description: The program is basically a modification of an earlier program by Neely and Mazukelli (see references below). The major modification consists of (1) an option to read data in the form of geographical matrices, and (2) restriction to only one of the possible grouping algorithms. The use of data not in geographic matrix form is still permitted, and contiguity matrices can be prespecified in these cases. Each of the n observations consists of p variables.

The original n observations are initially considered to consist of n p-dimensional groups each containing one element. The grouping procedure examines all of the $n (n - 1) / 2$ squared distances among these n observations and joins the two groups separated by the minimum distance. Each step of the grouping procedure decreases by one the number of groups still to be consolidated. Each grouping replaces two joined groups by a new group located at the center of mass of the pair, and this group then contains all the elements of the pair. Before a new step is initiated the distances from each of the remaining groups to the new group are calculated; these replace the distances to both of the component groups, with a consequent reduction by one of the distances to be examined at the next step. This aggregation requires n - 1 steps, after which all observations are in the final group. The minimum distance is added to the value of SUM at each step, thus providing a measure of the efficacy of each grouping. If the contiguity option is specified only distances between contiguous groups will be considered in the minimum proximity comparison.

Limits: The current version of the program will accept as many as 400 observations for each of the p variables, where p is defined by

$$n \times p \leq 80{,}200$$

Although subroutine GRPING is called only once, it is here that, for larger problems, the computer time is used. The program searches through the upper half of the distance squared triangle at each step of the grouping algorithm, an operation which requires examination of a number of distances varying from one to the triangle of n. The number of searches for the entire problem of n observations is given by

$$N = (n - 1)\, n\, (n + 1) / 6$$

The trivial case of two observations produces one search, while 85 observations require over 100,000, and 183 require more than one million. The maximum requires 10,666-600 searches to produce the 399 groupings.

Although 400 observations may seem a reasonably large number, it unfortunately yields geographical matrices of only 20 rows by 20 columns, a rather trivially small size. A nine inch by nine inch aerial photograph could be digitalized to a resolution of only circa one half inch, and a 20 inch topographical map to only a one inch resolution. It appears that other algorithms will be required for geographical grouping; i.e., capable of operating on the circa 1,000,000 observations contained in a single still coarsely digitalized geographical map or aerial photograph. The severity of the problem does not grow as rapidly with increases in p, the number of variables (e.g., for multispectral imagery or map overlays) and the contiguity constraint, essentially a local neighborhood operator, should allow greater efficiencies than are given in the present program.

There is an option not to print (IDSP = 1) the original distance matrix; it requires n/10 pages to print this. Forty pages, each containing up to five hundred E-format numbers, are best left in memory, although smaller problems and illustrative examples may utilize this form of output. Option KODE is also best left at 0; the distance matrices at the subsequent steps have similar paper requirements.

Deck Make-Up:

1) Problem card
2) Format card(s)
3) Data cards
4) Contiguity cards (optional)
5) Finish card.

As many sets of cards 1 through 4 as desired may be placed before the Finish card. If IDS (defined below) is 2, repetition of cards 2 and 3 is required for each of the variables specified on the Problem card.

Card Format: Problem Card:

Columns

1-6	PROBLM	
8-10	NOOB	Number of observations
11-14	MVAR	Number of variables
18	NFMT	Number of format cards
22	IDS	Input code: 2 if the variables are in geographical matrix form. In this case each of the p matrices, of n observations each, is read separately, always preceeded by a format card. This option may also be used to read one variable at a time. 1 To read the upper triangular matrix of distances squared, rather than raw data. 0 To read one observation of p variables.
24-26	NR	Number of rows (required only if both IDS and NCLUS are 2)
28-30	NC	Number of columns (required only if both IDS and NCLUS are 2)

34 NCLUS	Contiguities requested	
	2	to be computed from matrix data
	1	to be read in from data cards (see below)
	0	to be ignored
38 IDSP	Initial printing of the distances squared (upper triangular matrix)	
	1	to be suppressed
	0	to be effected
42 KODE	Stepwise printing of the distances squared.	
	1	to be effected
	0	to be suppressed
48 IP	Page control; set equal to one for best results.	
49-80 TITLE	An alphanumeric title of fewer than 33 characters.	

Format Cards:

There may be as many as nine E or F format cards, including initial and final parentheses. These will describe one complete observation of p variables or n observations on one variable. The word FORMAT must be punched in the first six columns of each card. If IDS is 2 there must be an identical number (NFMT) of format cards placed before the data for each of the p variables, although the data need not be punched in discrete rows if in matrix form and the content of the format cards may vary from variable to variable.

<u>Contiguity Cards:</u>

The contiguity matrix is symmetric, having n rows and n columns. This optional input requires the contiguity matrix in its entirety, except for the last row, although the contiguity states are read only from the upper triangle. Cards are punched according to format 80I1 for each row, with each row beginning on a new card. There will be an integral multiple of (n - 1) cards in the set. A 1 denotes contiguous observations, a blank the converse.

If contiguities are requested for data in geographical matrix form, the neighbors to an observation are defined as those observations which have a side in common with the cell in question: diagonally adjacent cells are not considered neighbors.

Output: In addition to the optional distances, at each step of the algorithm the program provides a statement of which groups were joined and the increment to the sum of squared distances. At each step the groups are labelled by the smaller identification number of the two in the pair. A dendrogram of the grouping is provided, and this depicts the groups present at a particular step by printing a plus sign. Each line of this graph includes one plus sign which is approximately halfway between two plus signs of the previous line, and this represents the two groups which have been joined at this step. The two groups will always be adjacent on the graph. This output permits division of the observations into any number of exclusive groups (partitions), although it is impossible to specify the total number of observations contained in any group at any step other than the first and last.

Programmer: Dierk Rhynsburger; subroutines TRI, ERROR, FORMAT, and PAGE adapted from a similar program at the University of Chicago by Neely and Mazukelli.

Reference: D.F. Marble, Some Computer Programs for Geographic Research, Department of Geography, Northwestern University, Evanston, 1967; program Congroup, pp. 23-34.

```
      DIMENSION A(80200), NO(400), FMT(162), TITLE(8), IFMT(162)
      DIMENSION IB(2), IPA(7), IPB(6), IPC(6), IPE(6)
      COMMON A
      INTEGER WORD, TITLE
      EQUIVALENCE (FMT,IFMT)
      DATA IB/'PROB','FINI'/, IEA/'MAIN'/, IEB/' N*P'/
      DATA IPA/'STAR','T OF',' PRO','BLEM',' - G','ROUP','ING '/
      DATA IPB/'FORM','AT F','OR I','NPUT',' DAT','A   '/
      DATA IPC/'DIST','ANCE',' SQU','ARED',' MAT','RIX '/
      DATA IPE/'INPU','T FO','RMAT',', VA','RIAB','LE  '/
 1001 FORMAT (A4,2X,9I4,5X,I1,8A4)
 2001 FORMAT(23H0 NUMBER OF VARIABLES = I9, 70X,17HSUBROUTINE MAIN    /
     1 26H0 NUMBER OF OBSERVATIONS =  I6,70X,19HDATED 30 JUNE 1969 ///)
 2002 FORMAT (1H0,3HTHE,I6,40H  ELEMENTS OF THE D SQUARED MATRIX HAVE ,
     1    17HBEEN READ AS DATA)
 2003 FORMAT (1H0,3HTHE,I3,40H MATRICES OF VARIABLES HAVE BEEN READ AS,
     1    5H DATA)
 2004 FORMAT (1H0,3HTHE,I3,37H MATRIX OF VARIABLES HAS BEEN READ AS,
     1    5H DATA)
 2005 FORMAT(1H0/1H0 20X A6)
      NPAGE = 1
      NSYZ = 80200
    1 READ 1001, WORD, NOOB, MVAR, NFMT, IDS, NR, NC, NCLUS, IDSP,
     1   KODE, IP, (TITLE(I),I=1,8)
      IF (WORD .EQ. IB(1))  GO TO 4
      IF (WORD .EQ. IB(2))  GO TO 3
    2 CALL ERROR(IEA,2,WORD)
    3 WRITE (6,2005) WORD
      GO TO 101
    4 IF (MVAR * NOOB)  6,6,5
    5 NTR = NOOB * (NOOB + 1) / 2
      IF (NTR - NSYZ)  7,7,6
    6 CALL ERROR(IEA,6,IEB)
    7 CALL PAGE(NPAGE,IP,TITLE,7,IPA)
      WRITE (6,2001) MVAR, NOOB
      DO 8 I = 1,NOOB
    8 NO(I) = I
C         THREE WAY BRANCH FOR INPUT
      IF (IDS - 1)  9,25,11
    9 I1 = NSYZ - NOOB * MVAR + 1 - MVAR
      I2 = I1 + MVAR - 1
      CALL FORMAT(NFMT,IFMT,162,6,IPB,0)
      DO 10 I = 1,NOOB
      I1 = I1 + MVAR
      I2 = I2 + MVAR
   10 READ (5,IFMT) (A(J),J=I1,I2)
      GO TO 15
   11 DO 12 I = 1,MVAR
      II = I
      CALL FORMAT(NFMT,IFMT,162,6,IPE,II)
      K = NSYZ - NOOB * MVAR + I
   12 READ (5,IFMT) (A(L),L=K,NSYZ,MVAR)
      IF (MVAR - 1)  15,13,14
   13 PRINT 2004, MVAR
      GO TO 15
```

```
   14 PRINT 2003, MVAR
   15 K = 0
      I1 = NSYZ - NOOB * MVAR + 1
      DO 24 I = 1,NOOB
      J1 = I1
      DO 22 J = I,NOOB
      K = K + 1
      A(K) = 0.
      I2 = I1
      J2 = J1
      DO 20 L = 1,MVAR
      A(K) = A(K) + (A(I2) - A(J2)) **2
      I2 = I2 + 1
   20 J2 = J2 + 1
   22 J1 = J1 + MVAR
   24 I1 = I1 + MVAR
      GO TO 26
   25 CALL FORMAT(NFMT,IFMT,162,6,IPR,0)
      READ (5,IFMT) (A(I),I=1,NTR)
      PRINT 2002, NTR
   26 CONTINUE
C          END OF INPUT AND DISTANCE CALCULATIONS
      IF (NCLUS)  101,29,28
   28 CALL CNTGTY(NOOB,NCLUS,NR,NC)
   29 IF (IOSP .EQ. 0)  CALL TRI(A,NO,NOOB,NPAGE,IP,6,IPC,0,TITLE)
   60 CALL GRPING(NOOB,MVAR,KODE,NCLUS,NPAGE,TITLE)
      GO TO 1
  101 STOP
      END
      SUBROUTINE GRPING(NOOB,MVAR,KODE,NCLUS,NPAGE,TITLE)
      DIMENSION TITLE(8), IB(4), IA(2), IPC(6), IPD(4), IPF(6)
      DIMENSION A(80200), IS(400,2), NZ(799), NX(799), ATR(400)
      DIMENSION NY(400), NC(400), NO(400), WG(400), TR(400), PCT(400)
      COMMON A
      INTEGER BLANK,DATA, TITLE
      DATA IB/' ','+','I','-'/, IFC/'GRUP'/, IED/'KODE'/
      DATA IPC/'DIST','ANCE',' SQU','ARED',' MAT','RIX '/
      DATA IPD/'GROU','PING',' SUM','MARY'/
      DATA IPF/'STEP',' GRA','PH O','F GR','OUPI','NG  '/
 2000 FORMAT(23H1NUMBER OF OBSERVATIONS I8,50X,39HSUBROUTINE GRPING
     1 DATED 21 JUNE 69 //)
 2001 FORMAT (I1,4HSTEP,I4,3X,5HSUM =,E15.6,13H   MIN DIST =,E15.6,
     1 11H  FOR ITEMS I5,5H  AND I5,6H  WITH 2I5,9H  MEMBERS)
 2006 FORMAT(1H0)
 2008 FORMAT(19H0      ITEMS GROUPED 5X,12I8)
 2009 FORMAT(1X,I4,2I5,E14.5,2H I,101A1)
 2010 FORMAT (1H ,29X,1H-,101A1)
 2012 FORMAT (1H ,25X,12I8)
 2014 FORMAT (1H ,29X,12I8)
 2015 FORMAT(25H STEP     I     J     VALUE   3X,12I8)
 5006 FORMAT (1H1,37HSTEP     I     J          %     WITHIN DSQ,7X,
     1    8HBETW DSQ/1H ,3X,1H0,16X,4H0.00,4X,11H0.00000E 00,E15.5)
 5007 FORMAT (1H ,I4,2I5,F10.2,2E15.5)
 5070 FORMAT (1H ,3X,1H0,13X,13H0.00000E 00 I,101A1)
      AVOG = 6.02257 * 10. ** 23
      NBB = NOOB - 1
```

```
      DO 502 I = 1,NOOB
      NZ(I) = 0
      NC(I) = I
  502 NO(I) = 1
      IC = 0
      TRACE = 0.
      IF (NOOB .GT. 16 .AND. KODE .EQ. 1)  IQ = 1
      IF (KODE)  15,17,17
   15 CALL ERROR(IFC,15,IED)
   17 WRITE (6,2000) NOOB
      CALL PAGE(NPAGE,0,TITLE,4,IPD)
      WRITE (6,2006)
      DO 550 NCYC = 1,NBB
      AMIN = AVOG
      N = NOOB - NCYC + 1
      NB = N - 1
      NTR = N * (N + 1) / 2
C          SEARCH WITH DIAGONAL DELETION
      K = 0
      DO 514 I = 1,NB
      K = K + 1
      II = I + 1
      DO 514 J = II,N
      K = K + 1
      IF (A(K))  514,516,512
  512 IF (A(K) .GE. AMIN)  GO TO 514
      AMIN = A(K)
      MR = I
      MC = J
  514 CONTINUE
      GO TO 518
  516 MR = I
      MC = J
      AMIN = 0.
C          THE SEARCH HAVING BEEN COMPLETED WITH THE LOCATION OF AMIN, BO
C          KEEPING PROCEEDS
  518 IR = NC(MR)
      IC = NC(MC)
      IS(NCYC,1) = IR
      IS(NCYC,2) = IC
      WG(NCYC) = AMIN
      TRACE = TRACE + AMIN
      TR(NCYC) = TRACE
      WR = NO(IR)
      WC = NO(IC)
      PRINT 2001, IQ, NCYC, TRACE, AMIN, IR, IC, NO(IR), NO(IC)
      IF (KODE .EQ. 1)  CALL TRI(A,NC,N,NPAGE,0,6,IPC,1,TITLE)
      IF (AMIN .EQ. 0.)  GO TO 707
C          COMBINE ELEMENTS OF THE CLOSEST PAIR
      MRR = MR - 1
      MPR = MR + 1
      MCC = MC + 1
      IF (MR .EQ. 1)  GO TO 703
      DO 702 I = 1,MRR
      II = (I - 1) * (2 * N - I) / 2 + MR
      JJ = (I - 1) * (2 * N - I) / 2 + MC
```

```
  702 CALL AA(NCLUS,II,JJ,WR,WC,AMIN)
  703 A(MR + MRR * N - MR * MRR / 2) = 0.
      DO 704 I = MRR,MC
      II = MRR * (2 * N - MR) / 2 + I
      JJ = (I - 1) * (2 * N - I) / 2 + MC
  704 CALL AA(NCLUS,II,JJ,WR,WC,AMIN)
      IF (MC .EQ. N)  GO TO 707
      DO 706 I = MCC,N
      II = MRR * (2 * N - MR) / 2 + I
      JJ = (MC - 1) * (2 * N - MC) / 2 + I
  706 CALL AA(NCLUS,II,JJ,WR,WC,AMIN)
  707 DO 708 I = MC,NP
  708 NC(I) = NC(I+1)
      NC(IR) = WR + WC
      NC(N) = 0
      NO(IC) = 0
C          CONDENSE THE VECTOR
  710 KK = MC - 1
      K = KK
      DO 714 I = 1,KK
      NN = N - I - 1
      IF (NN .EQ. 0)  GO TO 550
      DO 712 J = 1,NN
  712 A(J + K) = A(I + J + K)
  714 K = K + NN
      K = K + 1
      JJ = N * (N - 1) / 2
      DO 716 I = K,JJ
  716 A(I) = A(I + N)
  550 CONTINUE
C          PRINT THE GROUPING GRAPH
      NZ(1) = IS(NBB,1)
      NZ(2) = IS(NBB,2)
      DO 566 K = 2,NBB
      KK = NBB - K + 1
      DO 564 I = 1,K
      IF (IS(KK,1) .NE. NZ(I))  GO TO 564
      II = I + 1
      DO 562 J = II,K
      JJ = K - J + II
  562 NZ(JJ+1) = NZ(JJ)
      NZ(II) = IS(KK,2)
      GO TO 566
  564 CONTINUE
  566 CONTINUE
      JX = IB(1)
      JY = IB(4)
      I5 = 0
      LIM = NOOB / 49 + 1
      DO 587 L = 1,LIM
      I1 = I5 + 1
      I2 = I5 + 2
      I3 = I5 + 3
      I4 = I5 + 4
      II5 = I5 + 48
      I5 = MINO(II5,NOOB)
```

```
      N1 = 2 * I1 - 1
      N5 = 2 * I5
      IF (I5 .EQ. NOOB)  N5 = N5 - 1
      IF (I5 .EQ. NOOB)  JX = IB(3)
      NOB = 2 * NOOB - 1
      DO 572 I = 1,NOB
      NY(I) = 2 * I - 1
  572 NX(I) = IB(2)
      DO 574 I = 2,NOB,2
  574 NX(I) = IB(1)
      CALL PAGE(NPAGE, 1,TITLE,6,IPF)
      PRINT 2008, (NZ(I),I=I1,I5,4)
      PRINT 2012, (NZ(I),I=I2,I5,4)
      PRINT 2015, (NZ(I),I=I3,I5,4)
      PRINT 2014, (NZ(I),I=I4,I5,4)
      PRINT 5070, (NX(I),I=N1,N5), JX
      DO 586 K = 1,NBR
      DO 584 J = 1,2
      DO 582 I = 1,NOOB
      IF (NZ(I) .NE. IS(K,J))  GO TO 582
      IF (J .EQ. 1)  IX = I
      II = NY(I)
      NX(II) = IB(1)
      IA(J) = II
      GO TO 584
  582 CONTINUE
  584 CONTINUE
      II = (IA(1) + IA(2)) / 2
      NY(IX) = II
      NX(II) = IB(2)
      PRINT 2009, K, IS(K,1), IS(K,2), TR(K), (NX(I),I=N1,N5), JX
  586 CONTINUE
      PRINT 2010, (JY,I=N1,N5)
  587 CONTINUE
      DO 588 K = 1,NBR
      PCT(K) = TR(K) / TR(NBR) * 100.
  588 ATR(K) = TR(NBR) - TR(K)
      PRINT 5006, TR(NBR)
      PRINT 5007, (I, IS(I,1), IS(I,2), PCT(I), TR(I), ATR(I), I=1,NBR)
      RETURN
      END
      SUBROUTINE AA(NCLUS,I,J,WR,WC,D)
      DIMENSION A(80200)
      COMMON A
      WS = WR + WC
      WP = WR * WC
      IF (NCLUS - 1)  11,10,10
   10 IF (SIGN(1.,A(I)) + SIGN(1.,A(J)))  13,12,11
   11 A(I) = (WS * (WR * A(I) + WC * A(J)) - WP * D) / WS / WS
      RETURN
   12 A(I) = (WS * ABS(WR * A(I) - WC * A(J)) - WP * D) / WS / WS
      RETURN
   13 A(I) = (WS * (WR * A(I) + WC * A(J)) + WP * D) / WS / WS
   20 RETURN
      END
      SUBROUTINE CNTGTY(NOOB,NCLUS,NR,NC)
```

```
      DIMENSION TITLE(8), IR(10)
      INTEGER TITLE
 2002 FORMAT (1H0,8HOBS. NO.,I9,9I12)
 2004 FORMAT (1H ,I4,4X,10E12.4)
C          PRINTS A TRIANGLE OF A VALUES
      DO 9 I = 1,NOOB
      LL = NOOB - I + 1
      NB(LL) = NC(I)
    9 NA(LL) = I
      DO 10 I = 1,NOOB
   10 ND(I) = NB(I)
      NN = ((NOOB - 1) / 10) * 10 + 1
      DO 16 I = 1,NN,10
      LL = MIN0(I + 9,NOOB)
      MM = NOOB - I + 1
      DO 15 M = 1,MM
      IF (IS .EQ. 1)  GO TO 12
      IF (MOD(M - 1,50))  13,11,13
   11 CALL PAGE(NPAGE,1,TITLE,IQ,IR)
      PRINT 2002, (NC(K),K=I,LL)
      GO TO 13
   12 IF (M .EQ. 1)  PRINT 2002, (ND(K),K=I,LL)
   13 LL = MIN0(NOOB-M+1,LL)
      L = NA(I) + ((2 * NOOB - M) * (M - 1)) / 2
      DO 14 K = I,LL
      B(K) = A(L)
   14 L = L - 1
   15 PRINT 2004, NC(M), (B(K),K=I,LL)
   16 CONTINUE
      RETURN
      END
      SUBROUTINE FORMAT(N,F,L,M,W,II)
      DIMENSION F(162), H(6), W(6), IB(2)
      INTEGER F, H, W, WORD
      DATA IB/'    ','FORM'/, IEF/' LIM'/
 1001 FORMAT (A4,2X,18A4)
 2001 FORMAT (1H0,6A4,31X,18A4/(46X,18A4))
 2002 FORMAT (1H0,6A4,I4,27X,18A4/(46X,18A4))
      IF (N)  11,11,1
    1 DO 2 I = 1,L
    2 F(I) = IB(1)
      I2 = 0
      IF (9 - N)  3,4,4
    3 CALL ERROR(IB(2),3,IEF)
    4 DO 6 J = 1,N
      I1 = I2 + 1
      I2 = I2 + 18
      READ 1001, WORD, (F(I),I=I1,I2)
      IF (WORD .EQ. IB(2))  GO TO 6
    5 CALL ERROR(IB(2),5,WORD)
    6 CONTINUE
      DO 7 I = 1,6
    7 H(I) = IB(1)
      DO 8 I = 1,M
    8 H(I) = W(I)
      K = 18 * N
```

```
      IF (II)  9,9,10
    9 PRINT 2001, (H(I),I=1,6), (F(I),I=1,K)
      GO TO 11
   10 PRINT 2002, (H(I),I=1,6), II, (F(I),I=1,K)
   11 RETURN
      END
      SUBROUTINE PAGE(NP,IP,II,NW,WD)
      DIMENSION HEAD(10), WD(10), II(8)
      INTEGER HEAD,II,WD
      DATA IB/'    '/
 2001 FORMAT (I1,8A4,15X,10A4,32X,4HPAGE,I4)
      DO 2 I = 1,10
    2 HEAD(I) = IB
      IF (NW)  5,5,3
    3 DO 4 I = 1,NW
    4 HEAD(I) = WD(I)
    5 PRINT 2001, IP, II, HEAD, NP
      NP = NP + 1
      RETURN
      END
      SUBROUTINE ERROR(WORD,NUMBER,WD)
      INTEGER WORD,WD
 2001 FORMAT (1H0,23HERROR STOP,    ROUTINE    ,A4,15H    AT STATEMENT,I6,
     1    18X,10HTROUBLE IS,10X,A4)
      WRITE (6,2001) WORD, NUMBER, WD
      CALL EXIT
      END
```

GEOGRAPHICAL MATRIX DISAGGREGATION

Purpose: To prepare data from pairs of geographical matrices for input to a multiple regression program.

Description: Given two geographical matrices $A = (a_{ij})$ and $B = (b_{ij})$ both n x m, the assumption is made that $B = f(A)$. This program disaggregates A and B so that b_{ij}, and $a_{i+p,\ j+q}$ are repunched onto the same card, ready to be inserted into a multiple regression program (e.g., BMD 02R) with b_{ij} as the dependent variable. This provides a quick procedure for empirically estimating two dimensional weighting functions. If $b_{ij} = 0 = a_{ij}$, the observation is skipped. The program reads data from unit, punches on 7, and writes on 6, read from unit 5.

Deck Make-up: 1) Control Card

2) Format Card

3) Title Card

4) A matrix

5) Title Card

6) B matrix

Card Format: 1) Control Card

columns	
1-2	Number of rows in the matrices
3-4	Number of columns in the matrices
5	Number of rows in the $\pm p$ neighborhood
6	Number of columns in the $\pm q$ neighborhood. The result is $(2p + 1)\cdot(2q + 1)$ independent variables.

2) Format Card

1-80	FORTRAN format for one row of the input matrices

3) Title Card

1-80	title of the A matrix

4) A matrix

Data Cards

5) Title Card

1-80	title of the B matrix

6) B matrix

Data Cards

Repeating of cards (5) and (6) causes the B matrix to replace the A matrix, and a new "B" matrix is read in, ad infinitum.

Reference: W.R. Tobler, "Geographical Filters and their Inverses," Geographical Analysis, Vol. I, No. 3 (1969), pp. 234-253; and the references therein. Also see program EXTRAP.

```
C PROGRAM PREPARING DATA FROM GEOGRAPHICAL MATRICES
C FOR INPUT TO A MULTIPLE REGRESSION PROGRAM
C TO ESTIMATE NEIGHBORHOOD WEIGHTING FUNCTIONS
C (EMPIRICAL TWO-DIMENSIONAL CONVOLUTION WEIGHTS)
C B(I,J) = F(  A(I+P,J+Q)  )
C     WHERE THE INPUT MATRICES A AND B ARE OF
C     SIZE NROWS BY NCOLS, AND THE MATRIX OF NEIGHBORS IS
C     NR BY NC.
C     READ FROM UNIT FIVE.
C READS: CONTROLS;FORMAT OF THE MATRICES; THEN
C A'S TITLE, A MATRIX, B'S TITLE, B MATRIX,
C DOES COMPUTATIONS AND PUNCHING, THEN
C REPLCES A BY B, READS C'S TITLE, AND LOADS B WITH C MATRIX, ETC.
C
C     WRITE ON UNIT SIX.
C     PUNCH ON UNIT SEVEN.
C     ORDER OF OUTPUT FROM UNIT SEVEN.
C       (1) THE INDEX NUMBER OF THE POINT.
C       (2) (NR*NC-1) NEIGHBORS FROM THE A MATRIX BEGINNING
C           AT THE POSITION CORRESPONDING TO (1,1) IN THE MATRIX
C           OF NEIGHBORS AND SKIPPING THE CENTRAL POINT.
C       (3) THE CENTRAL POINT FROM THE A MATRIX.
C       (4) THE CENTRAL POINT FROM THE B MATRIX.
C     OUTPUT FORMAT FROM UNIT SEVEN IS 19F4.0
C     COLUMNS 77-80 HAVE BEEN LEFT BLANK FOR PURPOSES OF NUMBERING.
C
      DIMENSION A(100,100),B(100,100),FMT(18), T(18),W(100)
C
      READ(5,1000) NROWS,NCOLS,NR,NC
      READ(5,1010) (FMT(I),I=1,18)
      READ(5,1010) (T(I),I=1,18)
      WRITE(6,1020) (FMT(I),I=1,18),(T(J),J=1,18)
      DO 10 I=1,NROWS
10    READ(5,FMT)(A(I,J),J=1,NCOLS)
20    READ(5,1010,END=100) (T(I),I=1,18)
      WRITE(6,1030) (T(I),I=1,18)
      DO 30 I=1,NROWS
30    READ(5,FMT) (B(I,J),J=1,NCOLS)
      MINR=NR/2+1
      MINC=NC/2 +1
      MAXR=NROWS-MINR+1
      MAXC=NCOLS-MINC+1
      K=0
      M=NR*NC
      N=M+1
      DO 80 I=MINR,MAXR
      DO 80 J=MINC,MAXC
      IF((A(I,J).LT..00001).AND. (B(I,J).LT..00001)) GO TO 80
      K=K+1
      KK=0
      DO 60 IP=1,NR
      IPI=I+IP-MINR
      DO 50 JQ=1,NC
      JQJ=J+JQ-MINC
      IF((IP .EQ.MINR).AND.(JQ .EQ.MINC))GO TO50
```

UNIVARIATE GEOGRAPHICAL FORECASTING

Purpose: Extrapolation of geographical matrices in the time domain using a positionally invariant, time varying, linear, local operator.

Description: Data from an observation matrix, D, and matrices of weights are used to extrapolate the data with respect to time. Three options are available:

Option I: One weight matrix is used, and the results are time invariant.

$$Z_{i,j} = \sum_{p=1}^{K}\sum_{q=1}^{L} A_{p,q} D_{i+r,\ j+s}$$

where

$$r = \begin{cases} p - \frac{(k+1)}{2} & \text{if k is odd.} \\ p - \frac{(k+2)}{2} & \text{if k is even.} \end{cases}$$

and

$$s = \begin{cases} q - (l+1) & \text{if l is odd.} \\ q - \frac{(l+2)}{2} & \text{if l is even.} \end{cases}$$

Option II: Two weight matrices are used, and a time variant linear equation is employed.

$$Z_{i,j}^{t+\Delta t} = \sum_{p=1}^{K}\sum_{q=1}^{L} (A_{p,q} + B_{p,q}\Delta t) D_{i+r,\ j+s}$$

Option III: Three weight matrices are used, and a time variant quadratic equation is employed.

$$Z_{i,j}^{t+\Delta t} = \sum_{p=1}^{K}\sum_{q=1}^{L}(A_{p,q} + B_{p,q}\Delta t + C_{p,q}(\Delta t)^2)D_{i+r,\ j+s}$$

If there are an even number of rows or columns in the weight matrices, an asymmetrical calculation will result. Calculation of the weight matrices may be performed using the DECOMP program. Predicted negative values are set to zero.

Deck makeup: 1) Card 1. FORTRAN format (2I3, 2I2, I1)

columns

1-3 The number of rows in the observation matrix. (<100)

4-6 The number of columns in the observation matrix. (<100)

7-8 The number of rows in the weight matrices. (≤15)

9-10 The number of columns in the weight matrices. (≤15)

11 NUM, an integer which denotes the option desired.

If NUM = 1, only the A weight matrix is read and option I is chosen.

If NUM = 2, the A and B matrices are read, and the linear option is chosen.

If NUM = 3, the A, B, and C matrices are read, and option III is chosen.

If NUM = 4, the A, B, and C matrices are read, and results are printed for both option II and and option III.

It is assumed that all of the weight matrices are of the same size. Neither the weight matrices nor the observation matrix need be square arrays.

2) Card 2. FORTRAN format (4F8.3)

columns

1-8 The earliest time for which the results are desired.

9-16 The time assigned to the observation matrix.

17-24 The final time for which the results are desired.

25-32 The time increment.

Any unit of time may be used, as long as all times are expressed in a common unit.

3) The weight matrices, FORTRAN format (5F8.6)

If NUM = 1, one weight matrix must be provided.

If NUM = 2, two matrices must be provided and the A matrix should precede the B matrix.

If NUM = 3, or NUM = 4, three weight matrices must be provided, A first, B second, and C third.

4) The observation matrix, FORTRAN format (20F4.0)

Output: 1) The weights used, five numbers per line.

2) The time and the corresponding results for that time, fifteen numbers per line.

Reading and printing are done on logical devices 5 and 6, respectively. Writing on tape or on files is accomplished through logical device 7.

Programmer: L. Trevillyan

References:

R.G. Brown, Smoothing, Forecasting and Prediction of Discrete Time Series, Prentice Hall, 1962.

W. Tobler, "Spectral Analysis of Spatial Series", Proceedings, Fourth Annual Conference on Urban Planning Information Systems, University of California, Berkeley, 1966, pp. 179-186.

W. Tobler, "A Computer Movie Simulating Urban Growth in the Detroit Region", Economic Geography, IGU Proceedings, June 1970.

```
C PROGRAM FOR THE TIME DOMAIN EXTRAPOLATION
C OF GEOGRAPHICAL MATRICES, Z(I,J) = F (  D(I,J), T  )
C WALDO R. TOBLER, GEOGRAPHY, U. OF MICHIGAN, 1969
C
C READS FROM 5, WRITES ON 6 AND 7
C NROWS, NCOLS = # OF ROWS, COLUMNS (<101)
C IN THE OBSERVATION MATRIX D AND THE PREDICTED MATRIX Z
C IROWS, ICOLS = # OF ROWS, COLUMNS (<16)
C IN THE WEIGHTING MATRICES A, B, C.
C THE EXTRAPOLATION EQUATION IS
C Z(I,J) = THE SUM OVER P AND Q OF W(P,Q)*D(I+P,J+Q)
C WITH W(P,Q) = A(P,Q) + B(P,Q)*T + C(P,Q)*T*T
C  INPUT:
C     CARD 1.  NROWS,NCOLS,IROWS,ICOLS,NUM, FORMAT2I3,2I2,I1
C     NUM DENOTES THE EQUATION TO BE USED.
C     CARD 2.  TO,TI,TF,DT , FORMAT 4F8.3
C     TO IS THE EARLIST TIME FOR WHICH EXTRAPOLATION IS DESIRED
C     TI IS THE TIME FOR WHICH DATA IS PROVIDED
C     TF IS THE LATEST TIME FOR WHICH EXTRAPOLATION IS DESIRED
C     DT IS THE TIME INCREMENT
C     (TIMES MAY BE GIVEN IN ANY CONSISTENT UNITS)
C
C     THE A WEIGHT MATRIX, FORMATTED 5F8.6
C     THE B WEIGHT MATRIX, 5F8.6
C     THE C WEIGHT MATRIX, FORMATTED 5F8.6
C     THE D MATRIX, FORMAT 20F4.0
C     A,B,C ARE WEIGHT MATRICES OF SIZE (IROWS,ICOLS)
C     D IS THE OBSERVATION MATRIX . SIZE NROWS X NCOLS
C
        DIMENSION Z(100,100),D(100,100),A(15,15),B(15,15),AZ(100,100),
     1 BZ(100,100),C(15,15),CZ(100,100)
1      READ(5,10,END=200) NROWS,NCOLS,IROWS,ICOLS,NUM
10      FORMAT(2I3,2I2,I1)
        READ(5,20) TO,TI,TF,DT
20      FORMAT(4F8.3)
        READ(5,40)  ((A(I,J),J=1,ICOLS),I=1,IROWS)
        WRITE(6,55)  ((A(I,J),J=1,ICOLS),I=1,IROWS)
        IF (NUM.EQ.1) GO TO 30
        READ(5,40)  ((B(I,J),J=1,ICOLS),I=1,IROWS)
        WRITE(6,56)  ((B(I,J),J=1,ICOLS),I=1,IROWS)
        IF (NUM.EQ.2) GO TO 30
        READ(5,40) (( C(I,J),J=1,ICOLS),I=1,IROWS)
        WRITE(6,57)  ((C(I,J),J=1,ICOLS),I=1,IROWS)
30      READ(5,50)  ((D(I,J),J=1,NCOLS),I=1,NROWS)
40      FORMAT(5F8.6)
50      FORMAT((20F4.0))
55      FORMAT('1 THE A WEIGHT MATRIX'  //5(' ',F8.6,3X))
56      FORMAT('0 THE B WEIGHT MATRIX'//5(' ',F8.6,3X))
57      FORMAT('0 THE C WEIGHT MATRIX'//5(' ',F8.6,3X))
        JROW=IROWS/2
        JCOL=ICOLS/2
        LROW=NROWS-2*JROW
        LCOL=NCOLS-2*JCOL
        DO 60 I=1,NROWS
        DO 60 J=1,NCOLS
```

```
      Z(I,J)=0.
60    AZ(I,J)=0.
      IF(NUM.EQ.1) GO TO 65
      DO 61 I=1,NROWS
      DO 61 J=1,NCOLS
61    BZ(I,J)=0.
      IF(NUM.EQ.2) GO TO 65
      DO 62 I=1,NROWS
      DO 62 J=1,NCOLS
62    CZ(I,J)=0.
65    CONTINUE
C
C    CALCULATION OF AZ,BZ,CZ
C
      DO 73 I=1,LROW
      IR=I+JROW
      DO 73 J=1,LCOL
      JR=J+JCOL
      DO 73 K=1,IROWS
      DO 73 L=1,ICOLS
      M=IR+K-JROW-1
      N=JR+L-JCOL-1
      GO TO (73,70,69),NUM
69    CZ(IR,JR)=CZ(IR,JR)+C(K,L)*D(M,N)
70    BZ(IR,JR)=BZ(IR,JR)+B(K,L)*D(M,N)
73    AZ(IR,JR)=AZ(IR,JR)+A(K,L)*D(M,N)
C
C   BEGIN EXTRAPOLATION
C
      IF(NUM.NE.1) GO TO 72
      DO 71 I=1,NROWS
      DO 71 J=1,NCOLS
      IF(AZ(I,J).LT.0. .OR.D(I,J).EQ.0.) AZ(I,J)=0.
71    Z(I,J)=AZ(I,J)
      GO TO 200
72    T=TO-TI
      DO 130 JKL=1,5001
      XM=T+(JKL-1)*DT
      TK=TI+XM
      IF(TK .GT. TF) GO TO 135
      DO 100 I=1,LROW
      IR=I+JROW
      DO 100 J=1,LCOL
      JR=J+JCOL
      IF (D(IR,JR)) 90,90,80
80    CONTINUE
      IF(NUM.EQ.3) GO TO 81
      Z(IR,JR)=AZ(IR,JR)+XM*BZ(IR,JR)
      GO TO 82
81    Z(IR,JR)=AZ(IR,JR)+XM*BZ(IR,JR)+XM*XM*CZ(IR,JR)
82    CONTINUE
      IF(Z(IR,JR)) 90,100,100
90    Z(IR,JR)=0.
100   CONTINUE
      WRITE(6,110) TK
110   FORMAT('1THE COMPUTED VALUES IN ',F8.3)
```

```
      WRITE(6,120) ((Z(I,J),J=1,NCOLS),I=1,NROWS)
120   FORMAT(15(' ',F4.0,3X))
      WRITE(7) NROWS,NCOLS,TK
      DO 125 I=1,NROWS
125   WRITE(7) (Z(I,J),J=1,NCOLS)
130   CONTINUE
135   CONTINUE
      IF (NUM .NE.4) GO TO 1
      NUM=3
      WRITE(6,140)
140   FORMAT('0 RECALCULATION USING THE QUADRATIC EQUATION')
      GO TO 72
200   CONTINUE
C
C     SET NROWS AND NCOLS EQUAL TO NEGATIVE NUMBERS AS A SWITCH FOR *PIB
C
      NROWS=-3
      NCOLS=-3
      WRITE(7) NROWS,NCOLS,TK
      END
```

BINOMIALLY WEIGHTED SMOOTHING

Purpose: The program removes high frequency spatial components from a matrix of geographical data by use of a nine point binomially weighted local smoothing operation.

Description: The local smoothing operator is passed over each A value (input) to produce a matrix of B values (output). This operation may be repeated many times. The fundamental equation is:

$$B_{ij} = \sum_{m=-1}^{1} \sum_{n=-1}^{1} A_{i+m\ j+n} W_{mn}$$

where W is a set of weights (3 x 3) and A_{ij} is the original data. Boundary effects are noted as edges and corners are reached. The weight fields are as follows:

in interior			at boundary		at corner	
0.06	0.125	0.06	0.167	0.08	0.22	0.11
0.125	0.25	0.125	0.33	0.167	0.44	0.22
0.06	0.125	0.06	0.167	0.08		

Although there is no mathematical limit to the number of times the observations may be smoothed, edge effects will creep inward towards the center of the matrix with each consecutive smoothing operation (i.e., observations smoothed 3 times, boundary effects felt three rows in). Currently the program will handle an input matrix of 48 x 48. This may be adjusted by changing the dimensioning of A and B to (ROWS + 2 by COLS + 2). In the current version of the program the results will be punched out after all smoothings. The observation deck is read by rows beginning in the NW (upper left).

Data Deck: 1) Control Card

2) Title Card

3) Variable Format Card

4) Observation Deck

Data Card Composition: 1) Control Card

columns

1-2	MB	Number of Rows
3-4	NB	Number of Columns
5-6	NO	Number of smoothings requested

2) Title Card

Any title (≤ 72 characters)

3) Variable Format Card

Any E or F type FORTRAN type format for one row of data (≤ 72 characters)

4) Observation Deck

Observations as described by the format card

Programmer: D. Rhynsburger and H. Moellering, after a more general MAD program by W.R. Tobler

References:
W.R. Tobler, "Numerical Map Generalization," MICMOG Paper No. 7, University Microfilms, OP - 33067.

W. Tobler, "Of Maps and Matrices", Journal of Regional Science, 7 (1967), pp. 275-280.

```
C BINOMIALLY WEIGHTED SMOOTHING PROGRAM
C READ=5  WRITE=6   PUNCH=7
C WRITTEN BY D. RHYNSBURGER, MODIFIED BY H. MOELLERING
      DIMENSION A(50,50), B(50,50), C(18), D(18)
   81 FORMAT (18A4)
   82 FORMAT (40I2)
  100 FORMAT(1X,'INPUT DATA')
  101 FORMAT('1OUTPUT SMOOTHED ',I2,' TIMES')
  102 FORMAT('0',13F10.4/' '13F10.4)
  104 FORMAT(8F10.2)
  105 FORMAT('1ROWS=',I5,'COLS=',I5,'WITH ',I5,' ITERATIONS')
      READ (5,82)  MB, NB, NO
      WRITE (6,105) MB,NB,NO
      READ (5,81) D
      WRITE (6,81) D
      READ(5,81) C
      MA = MB + 2
      MC = MB + 1
      MD = MB - 1
      NA = NB + 2
      NC = NB + 1
      ND = NB - 1
      DO 4 I = 2,MC
    4 READ (5,C)  (A(I,J),J=2,NC)
      WRITE(6,100)
      DO 3 I=2,MC
    3 WRITE (6,102)  (A(I,J),J=2,NC)
      DO 15 K = 1,NO
C          EVALUATION OF THE INTERIOR
      DO 11 I = 3,MB
      DO 11 J = 3,NB
   11 B(I,J) = (A(I-1,J-1) + A(I-1,J+1) + A(I+1,J+1) + A(I+1,J-1))/16.
     1     + (A(I,J-1) + A(I,J+1) + A(I-1,J) + A(I+1,J))/8. + A(I,J)/4.
C          EVALUATION OF THE HORIZONTAL EDGES
      DO 12 J = 3,NB
      B(2,J) = (A(2,J-1) + A(2,J+1) + A(3,J))/6. +
     1    (A(3,J-1) + A(3,J+1))/12. + A(2,J)/3.
   12 B(MC,J) = (A(MC,J-1) + A(MC,J+1) + A(MB,J))/6. +
     1    (A(MB,J-1) + A(MB,J+1))/12. + A(MC,J)/3.
C          EVALUATION OF THE VERTICAL EDGES
      DO 13 I = 3,MB
      B(I,2) = (A(I-1,2) + A(I+1,2) + A(I,3))/6.  +
     1    (A(I-1,3) + A(I+1,3))/12. + A(I,2)/3.
   13 B(I,NC) = (A(I-1,NC) + A(I+1,NC) + A(I,NB))/6.  +
     1    (A(I-1,NB) + A(I+1,NB))/12. + A(I,NC)/3.
C          EVALUATION OF THE CORNERS
      B(2,2) = (4. * A(2,2) + 2. * (A(2,3) + A(3,2)) + A(3,3)) / 9.
      B(2,NC) = (4. * A(2,NC) + 2. * (A(2,NB) + A(3,NC)) + A(3,NB)) / 9.
      B(MC,2) = (4. * A(MC,2) + 2. * (A(MC,3) + A(MB,2)) + A(MB,3)) / 9.
      B(MC,NC) = (4. * A(MC,NC) + 2. * (A(MB,NC) + A(MC,NB)) +
     1    A(MB,NB)) / 9.
      WRITE(6,101) K
      WRITE (6,81) D
      DO 15 I=2,MC
      DO 14 J = 2,NC
```

Michigan Geographical Publications

University of Michigan
Department of Geography
Ann Arbor, Michigan 48104

Papers are issued several times a year. Priced variously; standing subscriptions granted twenty per cent reduction (payment must accompany single title orders).

1. Tobler, Waldo R., editor, Selected Computer Programs, 1970, 162 pp., $5.50. Second printing 1973.

2. Yuill, Robert S., General Model for Urban Growth: A Spatial Simulation, 1970, 221pp., $4.00.

3. Shannon, Gary W., Spatial Diffusion of an Innovative Health Care Plan, 1970, 166 pp., $4.00.

4. Tilmann, Sister Jean Paul, O.P., An Appraisal of the Geographical Works of Albertus Magnus, 1971, 190 pp., $4.00.

5. Deskins, Donald R., Jr., Residential Mobility of Negroes in Detroit, 1837-1965, 1972, 298 pp., $4.00.

6. Ma, Laurence J.C., Commercial Development and Urban Change in Sung China (960-1279), 1971, 196 pp., $4.00.

7. Jacoby, Louis R., Perception of Noise, Air, and Water Pollution in Detroit, 1972, 286 pp., $4.00.

8. Tobler, Waldo R., Lambert's Notes on Maps (1772), a translation with introduction of J.H. Lambert's "Notes and Comments on the Composition of Terrestrial and Celestial Maps," 1972, 125 pp., $4.00.

Manuscripts to be considered by the Department for publication may be submitted to Professor Thomas R. Detwyler or Professor John F. Kolars, editors.

Cartographic Data Structures

Thomas K. Peucker and Nicholas Chrisman

ABSTRACT. Efficient and flexible data structures are important to the development of computer mapping. Most current data banks are characterized by 1) structures which are convenient at the input stage rather than at the stages of use within computer programs, 2) separate and uncoordinated files for different types of geographic features, and 3) a lack of information about neighboring entities. The term "neighborhood function" may be used to indicate the relative location of a geographic entity and is a concept which is involved in all three of these characteristics. Ongoing research on data structures had led to work on the GEOGRAF system for encoding planar data and the GDS ("Geographic Data Structure") for encoding three-dimensional surfaces. Both involve data manipulation between the digitizing stage and the actual use of the data within computer mapping programs.

INTRODUCTION

A series of ongoing research projects are concerned with efficient and flexible data structures for geographic and cartographic analysis. The three main points of concern in the research can be summarized as follows:

1) 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 performed after the data have been input into the system from maps.

2) 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 combine them.

3) The data structure is usually very simple and lacks one facet in particular which is essential for much geographic and cartographic analysis—an indication of the relative location of a geographic entity, i.e., the position of a geographic entity with respect to its neighboring entities.

These three points may be abbreviated with the terms flexibility, comparability, and topology. This paper will characterize types of existing geographic and cartographic data systems for planar and three-dimensional surfaces, especially with respect to these three points. The paper will also describe attempts which have been made by the authors to produce data systems which eliminate some of the problems of existing ones. The term "neighborhood function" will play a major role throughout the paper and will therefore be explained in more detail in the following section.

NEIGHBORHOOD FUNCTION

When asked for the location of a city, we will give the location with respect to a river, a seacoast, a pass, a neighboring larger city, or other feature. Rarely will we use the geographic coordinates of longitude or latitude, nor will we use map coordinates. 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 de-

Thomas K. Peucker is associate professor, Simon Fraser University, Burnaby, Canada. Nicholas Chrisman is senior programmer, Laboratory for Computer Graphics and Spatial Analysis, Harvard University, Cambridge, Massachusetts. 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.

The American Cartographer, Vol. 2, No. 1, 1975, pp. 55–69

scribe 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) as orientation characteristics.

In contrast, when a geographic data bank of any kind is created, it utilizes some kind of absolute coordinate system. Usually neither the geographic evaluation at the time of digitizing nor the mapping system used with the data allow the inclusion of such cartographic features as streets, rivers, and roads which would give us an indication of relative location.

While the human user can be aided in his orientation on a map through overlays or map comparison, the computer has difficulty in determining relative location If the relative location in terms of the closest points for each of, say, 5,000 points is to be calculated, the program literally has to compute every point's distance to every other point. Indeed, some widely-used programs do this computation several times within one program run.

Some indication of the relative location of a geographic feature can be very useful. This neighborhood relationship will be referred to as a "neighborhood function." It can be expressed in different ways: as an *explicit* or *implicit function,* or as a discrete function in the form of a *table.*

The *explicit function* can be a polynomial or trigonometric equation 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, et al. (1973). Two-dimensional spline functions also fall into 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 closest neighbors. This is done in various interpolation algorithms to produce a regular grid of points (Shepard, 1968; Heiskanen and Moritz, 1967). The computations increase close to the square of the increase in the number of points, since the search has to be repeated for every point and all points, or at least a large number of them, must be processed each time.

This search procedure also a... the case of planar surfaces where neighboring polygons must be found. For example, when contiguity constraints are imposed in problems of factorial ecology and other regional correlations, all polygon points must be searched to find those which are in common for a pair of polygons. Again, the problem increases in complexity according to the square of the number of the items being searched.

The *implicit function* expressing neighborhood relationships is usually a function that describes the coding structure of the geographic entities. One very good case is described in Rosenfeld (1969) for different types of neighborhood relationships within a regular grid in which the point P_{ij} has the four neighbors $(i+1, j)$, $(i, j+1)$, $(i-l, j)$, $(i, j-1)$ and it has 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" indicating 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 accompany the record of each point The most widely-known structure of this type is the DIME file of the U.S. Census which encodes line segments, the names of the polygons to the left and right of each line segment, and the names of the two nodes at either end. The neighborhood relationships used in the DIME development are derived from the discipline of topology (Cooke and Maxfield, 1967).

Neighborhood relationships will be discussed in greater detail in the analysis of various existing data structures. Two types of geographic data bases will be discussed: those defining planar surfaces and those defining three-dimensional surfaces. For both types, a summary of their historical development will be presented, and it will be shown that, although presently at different stages of development, these two types can be treated as special cases of one topological data structure.

DATA STRUCTURES FOR PLANAR SURFACES

Types of Structures

The types of geographic entities on planar surfaces are points, lines, and area-enclosing lines or polygons. The latter are perhaps the most frequently encoded feature in geographic data systems.

The simplest data base system for planar surfaces is that of encoding *entity by entity* with little or no regard for entity overlaps or adjacencies (Fig. 1). In other words, every polygon in a polygon system is encoded and stored without any regard for contiguous polygons, and lines are encoded without regard for the fact that they may intersect or merge with other lines. The results of such an encoding are "sliver lines" (duplication of lines in slightly different positions). These sliver lines are confusing and unaesthetic and, hence, it is virtually impossible to do anything directly with such a data base except an extremely coarse graphic image.

To go beyond the use of such data for the production of coarse images, editing must be performed. This alternative has been attempted in several cases, one being the MAP-MODEL system (Arms, 1970). The editing in this system is guided by the assumption that every segment has to be represented twice except for segments on the outer boundary. For each segment, the editing program sorts through all remaining segments to find its complement (the identical line of the neighboring polygon). Those segments for which it has not found a complement are tagged to indicate potential errors to the user.

To overcome some of the limitations of independently encoded entities, systems have been developed based on a common *location dictionary*. This dictionary contains the coordinates of every boundary point on the map. Polygon boundary lists are then compiled which consist of the labels (location numbers) of these boundary points (Fig 2). Line information can be handled in the same way. Programs based on this structure include CAL-FORM from the Laboratory of Computer Graphics and Spatial Analysis. Other programs have subroutines to convert this point dictionary structure to the simple entity-by-entity line list described above. The data can then be used in programs such as SYMAP (Laboratory for Computer Graphics and Spatial Analysis) which are compatible with the entity-by entity structure. Programs have also been developed to simplify data input through automated polygon identification (Douglas, 1973).

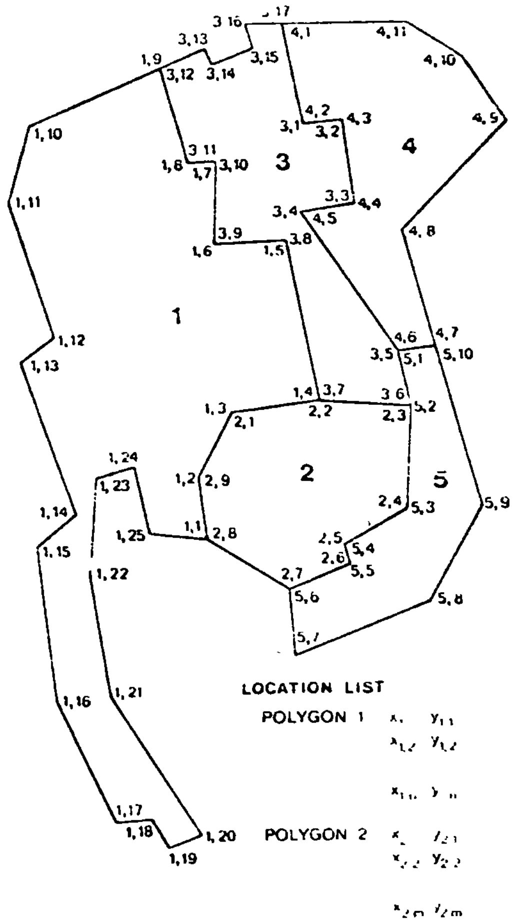


Fig. 1. The simplest encoding is areal entity by areal entity. In this system, most points are recorded twice and some (e.g., $P_{1,4}$; $P_{2,2}$; $P_{3,7}$) three times. A point does not necessarily have identical coordinates in all recordings.

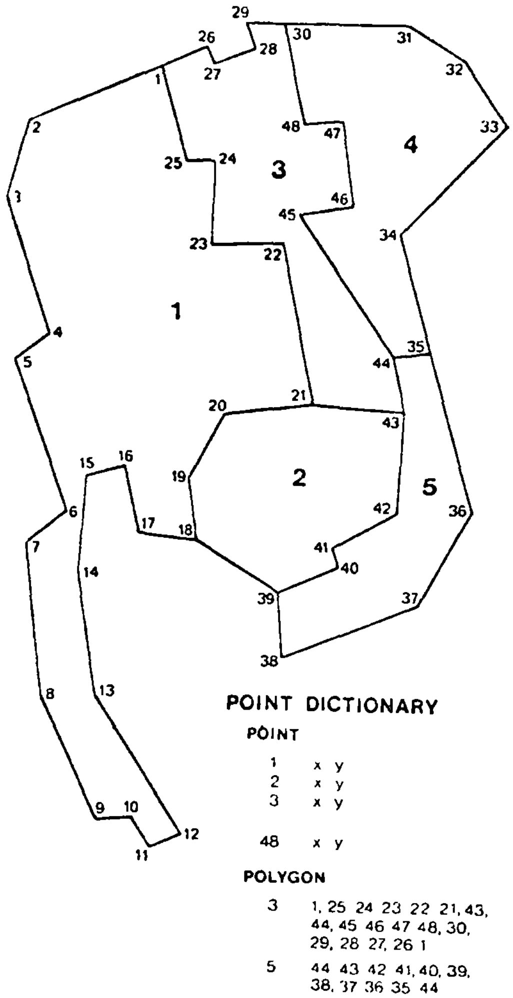


Fig. 2. In the second type of polygon system, the individual points are encoded only once and are stored in the Point Dictionary. For every polygon, a Polygon Boundary List is then established.

The point dictionary data base has the advantage that sliver lines do not occur. However, the problem of neighborhood relationships is not handled any better than with the entity-by-entity approach. The search for common lines is no longer according to the coordinates of points but by their labels; this brings us closer to a solution only by a little less computer time. It also creates difficulties. A point dictionary can and will be accessed in an arbitrary order, since there are no restrictions regarding point placement. The standard response to this problem is to make the dictionary core resident. Unfortunately, this will limit the complexity of the map that can be handled in this manner, since all points must be stored throughout the operation of the program. This shortcoming of such sharing of 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 simply *n* references to points.

Some of the objections to the ordinary point dictionary approach 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;* these segments 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 its 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 neighborhood relationships are not known. By adding the topological neighborhood 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 entity form* in the encoding stage. Many existing geographic information systems store land-use data in grids of rectangular cells (Hsu, 1975). However, a serious problem is encountered with such a regular discrete encoding of planar surfaces. According to the sampling theorem, the sampling interval has to be half the size of the smallest features to be encoded (e.g., Tobler, 1969). Hence, either the size of the cells has to be very small to enable encoding of detailed variations (e.g., urban land uses) or a grosser cell size can be sed for encoding only the very slowly changing variations (generalized land use). In the case of the small units, a very large volume of redundant information is created in an area of uniform land use and matrix reduction techniques such as run-length encoding (recording in each row only those cells differing from an immediately previous cell) create only physical and not logical compaction (Amidon and Akin, 1970). In the case of the grosser cell size, highly varying land use features are aggregated to a degree which reduces the usefulness of the whole system.

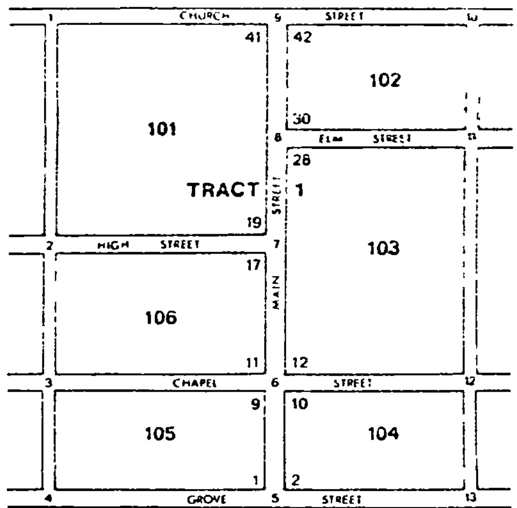
Beyond sampling problems, a grid structure imposes a bias toward specific orientations of features. Diagonals, although physically longer, are given the apparent cell relationships of unit distances along the major axes. The grid structure also creates the illusion of working with a discrete point space, rather than a regular areal partitioning.

Recent studies have developed parameters which determine the degree of inaccuracy of a given sampling mesh (Switzer 1975). In cases where some types of geographic features vary over small areas and others over very large areas, such as the case of statewide land-use patterns, the error is averaged over all features and will affect features with a high-frequency variation more than others.

The variation of the mesh size within the information system would be of only little help. The data management and the manipulation and display routines would become more complex, although probably less than many researchers imagine. Exact figures about the differences cannot be given since no literature about such a system is known to the authors

Structures Using 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 the Census (Fig. 3). The DIME files were originally developed as an automated topological error detection system for the Address Coding effort of the 1970 census.



CENSUS ADDRESS CODING GUIDE RECORDS

STREET	TRACT	BLOCK	LOW ADDRESS	HIGH ADDRESS
Main	1	102	30	42
Main	1	103	12	26
Main	1	104	2	10
Main	1	105	1	9
Main	1	106	11	17
Main	1	101	19	41

DIME STREET SEGMENT RECORDS

STREET	NODE START	NODE END	TRACT LEFT	BLOCK LEFT	TRACT RIGHT	BLOCK RIGHT	LOW ADDR	HIGH ADDR
Main	5	6	1	105	1	104	1	10
Main	6	7	1	106	1	103	11	17
Main	7	8	1	101	1	103	19	28
Main	8	9	1	101	1	102	30	42

Fig. 3. The DIME-file base. Represented is a portion of Tract 1 in a hypothetical city (Cooke and Maxfield, 1967). The small numbers at the street intersections are the nodes, the larger numbers on Main Street are the addresses. Two types of records are created as illustrated in the tables.

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 other line. 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 much information accessible to urban researchers, neighborhood relationships are not made explicit. Segments sharing a node, for example, must be found by laborious search procedures. Search is also required 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. For procedures in a one-time checking effort, as is the case of Address Coding and Address Matching in metropolitan areas, it is quite adequate For efficient computer storage and retrieval and for many applications, however, improvements are desirable For example, the reliance on the individual line segment makes the reduction of detail 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, POLYVRT, that is designed to contain all the information needed to construct any of the previously enumerated planar structures The basic object of POLYVRT is the "chain." Like a DIME segment, a chain has nodes at its two ends, separates two areas, and is assumed to be uncrossed It differs in that the POLYVRT chain may be made up of many points whereas the basic DIME unit has only two points. A boundary between two polygons can be referenced by a single chain no matter how complicated, because line detail is topologically unimportant (Laboratory for Computer Graphics and Spatial Analysis, 1974).

The coding of a complicated boundary as a unit is not unique to POLYVRT. The data bank used in the project "The Interactive Map in Urban Research" (Nake and Peucker, 1972; Peucker (ed.) 1973) as well as the World Data Bank I (Schmidt, 1969) are composed of "lines." In the latter case some of them contain over 4,000 points The chain based system of POLYVRT, however, is a different type of 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 except that 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 dependency on the number of boundaries.

In addition to the indication of the relative location of the chain with respect to its neighboring polygons, POLYVRT 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, 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 indicating the locations of the points in the point file, the chain record incorporates the name of the chain, the labels of the starting and ending nodes, and the left and right polygons Conversely, the polygon list consists of the bounding chains in proper sequence (Figs 4 and 5) A list linking nodes to chains could easily be constructed

DATA STRUCTURES FOR THREE-DIMENSIONAL SURFACES

Boehm (1967) describes in a very detailed analysis the advantages and disad-

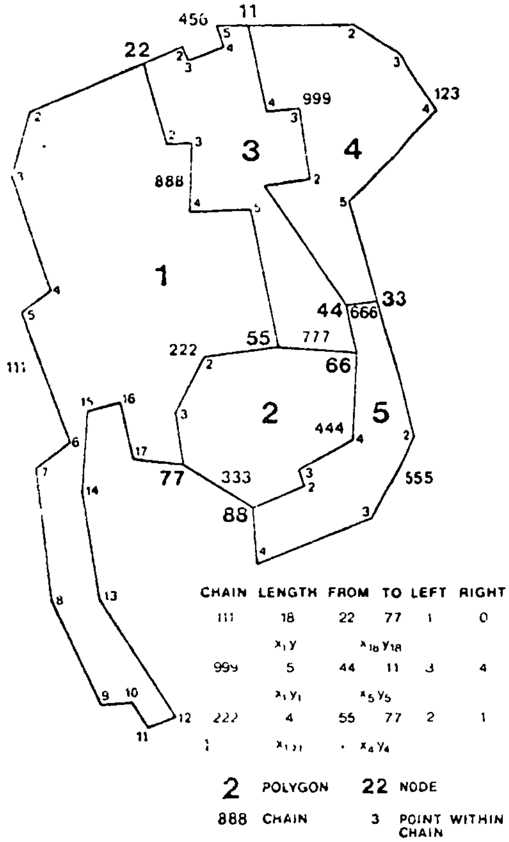


Fig. 4. The external representation of the POLYVRT chain-file. Every chain has a "name," the number of inner points (length), the two limiting nodes, the two boundary polygons, and a series of coordinates for the points.

vantages of different ways of encoding surfaces. He comes to the conclusion that the encoding of surfaces by contours minimizes the storage capacity, whereas a regular grid of surface points minimizes the computing time necessary for several types of manipulations. Boehm did not include in his study a data system with an irregular distribution of points but only contour encodings and regular grid structures of constant and variable mesh width. He did not reflect on the reasons why contoured data minimize storage capacity whereas a regular grid minimizes computing time. If he had done so it is quite possible that the development of geographic data bases would have taken different routes. Surprisingly 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, it is necessary to adapt the density of points to the variation in the local terrain. The question of how dense the points have to be can be answered by again using the philosophy of the sampling theorem. For the terrain within a typical map, the variation can change considerably, resulting in a need for frequent adjustments of the sampling interval.

In a contour map the density of contour lines changes with the density of relief variation. Therefore, it 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 length") S, then the grid spacing everywhere must be $S/2$ or less (note, for example, Mark, 1974). The regular grid again tends toward redundancy since smooth areas within 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 of f, the grid spacing must be decreased by

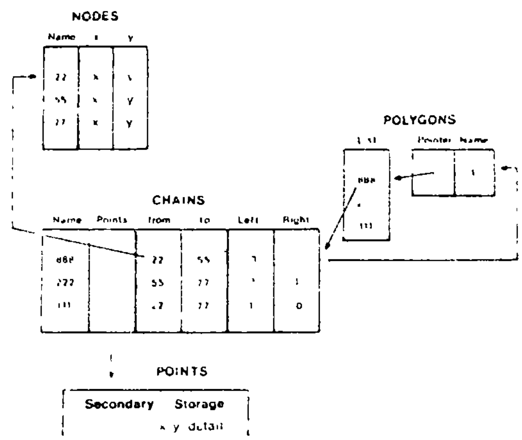


Fig. 5. The internal representation of the POLYVRT chain-file. Reference to polygon chains are positive or negative according to the direction of the chain used.

the reciprocal of this factor and the total number of points is increased by f^2.

The question has been raised many times in photogrammetry whether contours are the best representatives of topographic surfaces. It has been stated that contours do not detect many types of breaks which are frequent on terrain (e.g., Brandstatter, 1957). Therefore, the encoding of surfaces by vertical profiles has been attempted many times in photogrammetry in recent years. Points are encoded only when the slope of the surface changes (e.g., Kraus, 1973). In another case the break lines are encoded in addition to a regular grid (Sima, 1972), and in another approach only the break lines or only the ridges and channel lines of a surface are encoded (Grist, 1972). The amount of detail with these approaches depends 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. Most of the numerical computations on surfaces require some type of neighborhood function, either to compute some surface behavior, such as slope or local variation of relief, or to find the next unit for the drawing of a contour or a vertical profile. For a set of contours it is relatively easy to create a directory which indicates a sequence of contours in a type of tree in which the surrounding contour is the base and the other contours are the branches (Morse, 1968). To find the neighboring points on the two adjacent contours for a given point on an intermediate contour, however, one must search through all the points on the two adjacent contours and compute the distances to the point in question. This procedure is time-consuming if the contour lines are represented by very many points. A regular grid on the other hand has an implicit neighborhood function and finding a neighbor does not involve search nor extra computer time. For a set of irregularly-distributed points as they are represented in very simple data structures (e.g., SYMAP) the creation of a neighborhood function is usually performed by finding the closest small number of points where the number varies around six.

THE GENERAL CASE

It has been noted that although planar surfaces and three-dimensional surfaces look rather different it takes only a few assumptions to treat one as the subset of the other. While three-dimensional surfaces are always based on interval or ratio data, planar surfaces often involve ordinal and nominal data. However, it has been shown that ordinal and nominal data may be treated as interval data (Nordbeck and Rystedt, 1970; Rosenfeld, 1969) and even without this conversion we can combine the two types into one general case by simply using different assumptions about neighborhood.

Given a set of n data points for which x, y, and z coordinates are known, continuously defined surfaces (i.e., surfaces for which one and only one value exists at every point) can be created, using any of three different assumptions about the surface behavior (Peucker, 1972). The first assumption is that of a stepped 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). The second assumption is that each data point represents a sample of a single value on a constantly changing 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 from a constantly changing surface that may contain errors; thus the data point is not necessarily located on the surface, but close to it. This approach implies the further assumption that the actual surface is smoother than the surface constructed through the sample data points

With these assumptions about surface behavior, any point or areal distribution of a variable can be treated as a continuous function $z = f(x, y)$, and planar and three-dimensional surfaces can be combined into one type. This has already been done in some computer programs, the most notable being SYMAP. Many cartographic applications must treat both types of surfaces and, therefore, data structures must be developed which can handle both types at one time.

THE PROPOSED DATA STRUCTURES

Planar Surfaces

The need to incorporate different types, or hierarchies, of polygons uncovered one of the limiting assumptions made in POLYVRT. A chain plays a dual role: first, it is the boundary of two areal entities, and secondly, it is the unbroken unit of point retrieval. This is not a problem if one is limited to a single nonoverlapping set of polygons. The following proposed new system, to bear the name GEOGRAF, is an attempt at greater flexibility.

Because of the addition of many layers of complexity involving multiple polygon sets, 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 for a root object which is defined as an area uncut by any further partitioning. This object is termed the Least Common Geographic Unit (LCGU). The LCGU's are constructed as a POLYVRT polygon directly from chains. The relationship of the LCGU's to all other polygon types is hierarchical (Fig 6).

In turn, the existence of the LCGU allows for the creation of each class of polygon. In order to allow simple coding of the boundary relationships at these higher levels in the structure, the "chain group" was devised. A chain group is a set of chains which form a boundary of two areal units for a given polygon class. These polygons are constructed of chain groups which, in turn, are constructed from chains.

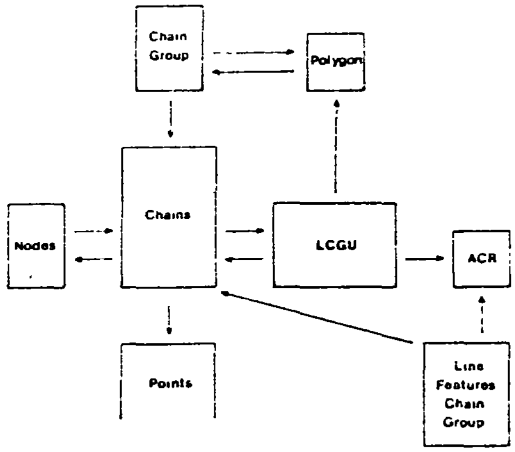


Fig. 6. The relationship between the various parts of the data structure.

Line features can be built up of chains in the same manner as chain groups. Note that the chain group listing for each level of polygon and each listing for a line feature reference only the chains themselves. This allows 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 can assign a string of data, collected for one polygon type, to a string of a second type.

Topological manipulation routines are central to the success of this structure. 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, built 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 uses nonlinear windows (which is often the case with geographic coordinates and map projections) and becomes quite elegant. Similarly, intersection procedures such as point-in-polygon, line-across-polygon, and polygon-over-polygon determination allow for gainful application of the topological principle. For point-in-polygon searches, chain groups are constructed which bisect the universe into parts to sort the points (or nodes of polygons) into three groups: left, within the band, and right. Only the second group needs more detailed treatment. The point set is then recursively partitioned until it reaches the level of the LCGU's. For line-oriented problems, the topological connections of the two sets compared would allow intersections to be limited to immediate neighbors.

Another 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. The search of a line through a set of polygons will use a graph-search algorithm developed for the GDS (Geographic Data Structure) project discussed in the next section. The neighborhood search routines create records of neighbors for every point or line or polygon at any specified depth of neighborhood.

Three-Dimensional Surfaces

To implement the ideas presented here, the senior author is developing a geographic information system for three-dimensional surfaces. Both systems, the three-dimensional and the planar, 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, a data set was generally tied to the application program that used it. Now the available data have become so voluminous and the application programs so varied that extra efforts in the preparation of data 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. The interpolation of an irregular grid to a regular grid of height-points provides a data structure through the implicit neighborhood function of the grid. And in the case where polygons are independently defined by a series of points, the neighborhood relationship is replaced by a search algorithm which finds the neighbors for every polygon by searching for matching segments in the boundary files of the other polygons. The idea is to spend computing time before any application has been performed in the anticipation of heavy uses of the data base. An efficient structure of the data base should not only speed up computations considerably, but should also simplify the production of application programs.

The data structure developed under the working title "Geographic Data Structure" (GDS) 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 form 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 triangle edge (Fig. 7). The second structure is produced by selecting those points of the surfaces which lie along lines of high information content, such as ridges and channel lines, and defining them by their nodes, which are peaks, passes, and pits (Fig. 8). This second data structure serves two purposes. First, it is a general representation of the surface for rough computations; second, it is a "directory" into the more detailed first structure.

In the first structure, the creation of the neighborhood relationship is based on the assumption that data-sets are usually of

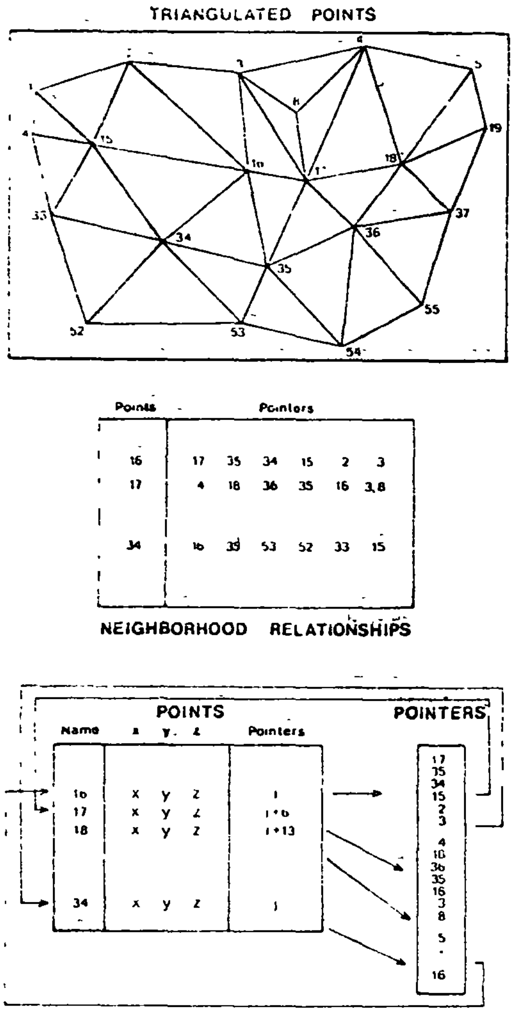


Fig. 7. The GDS-first data structure. The illustration shows the points on the surface with their links to neighbors (edges). The external representation is shown by the neighborhood relationships. The internal representation is composed of the point-file and the pointer-file.

two types: (a) sets of irregularly-distributed points which were digitized with the understanding that every point is significant, and (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, e g , 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 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 hundred. The first step in our approach therefore limits the links to a number of "potential neighbors," among which the shortest link is chosen and intersected with all other links originating in these potential neighbors. This procedure 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 between 8 and 14 depending on the density variation of these points. The procedure does not guarantee, however, 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

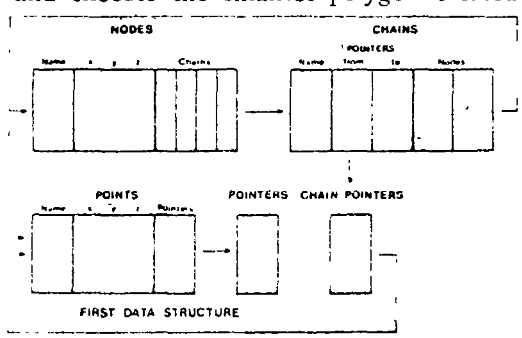


Fig. 8. The GDS-second data structure. Both the node-file and the chain-file have access to the first data structure, the node-file directly and the chain-file through a chain-pointer-file.

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.

Another approach which limits the number of necessary tests, but is mathematically correct at the same time, has been developed within the project by Kurt Brassel at Harvard University. 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, developed by Randolph Franklin of Harvard University and T. K. Peucker, 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). Since the record has to be long enough to include all possible numbers of neighbors, large parts of the pointer sections would be empty most of the time. The pointers are sorted, starting with the pointer the least East of North of a point (Fig. 7).

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 contouring, for example, the critical question is whether one point of the edge is above the contour level and the other below. A start is found and one point of the edge is 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 has only to have a reference and a subpoint. The third point is the next label in the pointer list of the reference point after the subpoint. To find all triangles one goes through the total pointer list leaving out all those edges connecting reference points with subpoints with a smaller label since they would create triangles which had been treated when the subpoint was a reference point.

Although the first data structure, as presented above, seems to be efficient in terms of storage capacity, it does not provide easy access to the data base which is often very large. It is for this reason that we are developing the second data structure to represent the general structure of the surface and to serve as a directory to reach into the first data base (Pfaltz, 1975) (Fig. 8).

The first step in the creation of the second data structure is to find the ridge, channel, and break lines on the surface. If one labels the highest point for every triangle, the unlabeled points are members of the channel line, at least on smooth surfaces. A subsequent search routine deals with the irregularities. Points along ridges are found by eliminating the lowest point of every triangle, using a routine developed for a regular grid by D. H. Douglas. The detection of break lines is somewhat more difficult.

Some theoretical studies of surfaces by Warntz (1966) show that ridge lines and channel lines cross at passes, a useful point of information relative to the development of the second data structure. Practical considerations suggest that in terrain and other surfaces, this regularity is not always present.

Once the topological structure is at hand, it can be used as a directory into the first structure. The line is therefore treated as a chain similar to the chain of the POLYVRT system. The nodes of the chain are the peaks, passes, pits, and other endpoints of chains on the surfaces. These points are stored with their coordinates and the names of the chains which terminate at the nodes. The chains are stored with the labels of the nodes and pointers into the chain lists which consist of labels of points in the first data structure (note that here the chain structure differs from that of POLYVRT).

A third component of the "Geographic Data Structure" should be mentioned since it illustrates very well the logical adaptation of a computer problem solution to geographical data. 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 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 very efficient data encoding even in terrain with sudden changes in the surface behavior as at the change from a mountainous area into a plain (Peucker, 1972).

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 an attempt has been made 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 system with little difficulty. This link between the two systems lends hope that eventually they may be merged.

It is an appropriate question to ask what such a data structure as the GDS will be able to accomplish. A number of display routines have already been developed (Cochrane, 1974) and a series of procedures for surface analysis based on heuristic searches are underway (Fig. 9). Since both levels of data structure are graphs, we will be able to rely on many of the developments connected with operations research, specifically network analysis, for the manipulative treatment of the data.

As both systems, GDS and GEOGRAF, have topological structures, it is possible to merge the two. The creation of polygons from points is the major link from the GDS project to 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 useful in quantitative geography and planning. Neighborhood searches are extremely expensive without the topological data structure, but they are usually a most important part of urban and environmental analyses once a general overview is obtained from the data.

Although basic research and application development are two sides of one coin and must go together to obtain lasting results, this paper has concentrated on the theoretical parts of the project since their development is ahead of the application routines, a fact which should be expected.

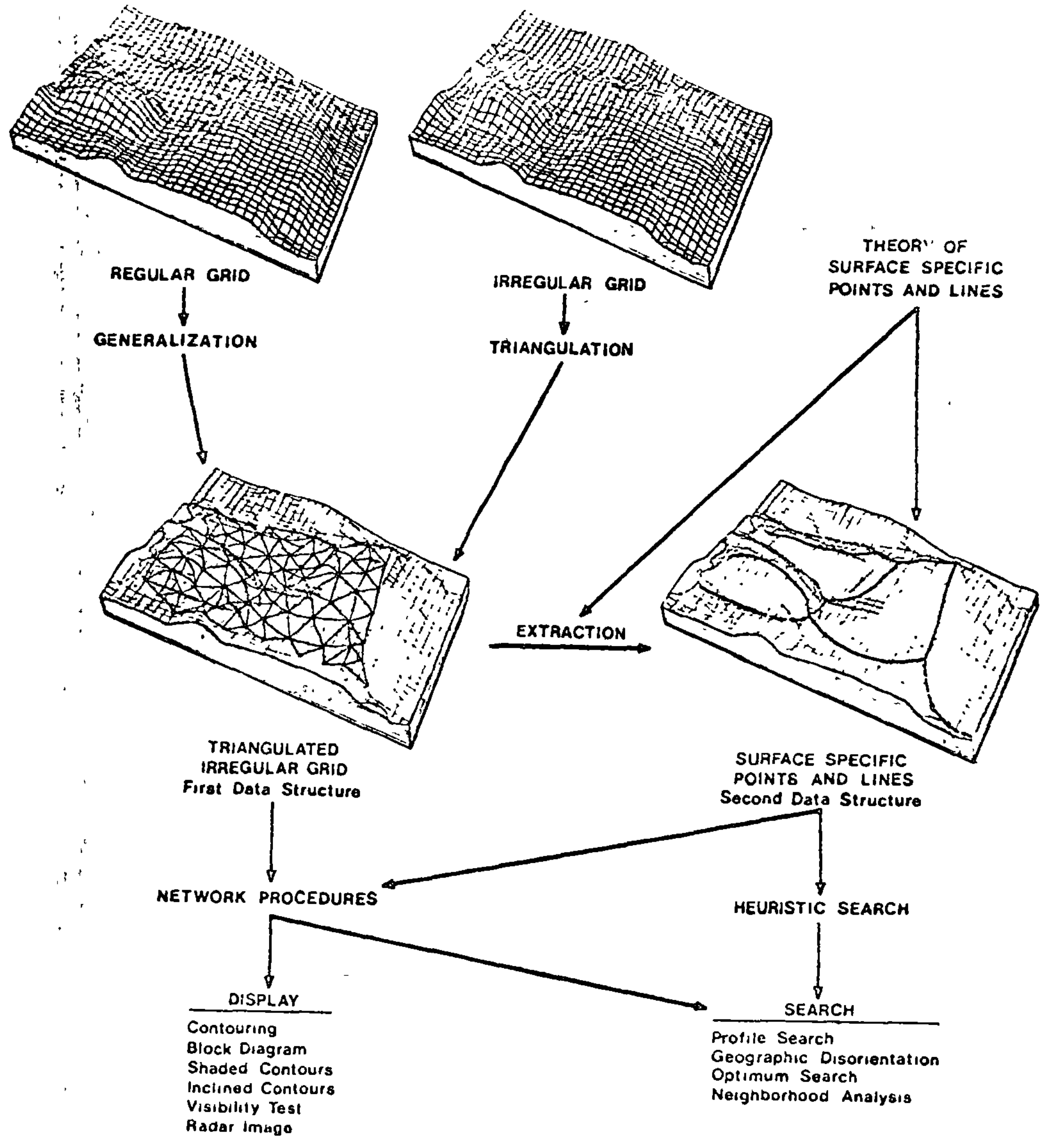


Fig. 9. The problem flow for GDS from the data base, via the extraction of points and chains for the two types of data structures, to their application.

The quintessence of the 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 and execution of application routines. We have some indication that the hypothesis is correct; the real test will come when the bulk of the application routines is completed.

REFERENCES

Amidon, E. L. and G. S. Aiken (1970), "Algorithmic Selection of the Best Method for Compressing Map Data Strings," *Communications of the Association for Computing Machinery*, Vol. 14, pp. 769–774.

Arms, A. (1970), *MAP/MODEL System: System Description and User's Guide*, Bureau of Governmental Research and Service, University of Oregon, 59 pp

Boehm, B. M. (1967), "Tabular Representation of Multivariate Functions with Applications to Topographic Modelling," *Proceedings*, 22nd National Conference, Association for Computing Machinery, pp 403–415.

Brandstaetter, L. (1957), "Exakte Schichtlinien und Topographische Gelaendedarstellung," *Oesterreichische Zeitschrift fuer Vermessungswesen*, Sonderheit 18, Wien, 90 pp.

Cochrane, Douglas (1974), "Cartographic Display Operations on Surface Data Collected in an Irregular Structure," M.A. thesis, Simon Fraser University.

Cooke, D. F. and W. F. Maxfield (1967), "The Development of a Geographic Base File and Its Uses for Mapping," *Urban and Regional Information Systems for Social Programs*, Papers from the 5th Annual Conference of the Urban and Regional Information System Association, pp 207–218.

Douglas, David H. (1973), "BNDRYNET," Peucker, T. K. (ed.), *The Interactive Map in Urban Research, Final Report After Year One*, University of British Columbia.

Dueppe, R D and H. J. Gottschalk (1970), "Automatische Interpolation von Isolinien bei willkuerlich verteilten Stuetzpunkten," *Allgemeine Vermessungs-Nachrichten*, Vol. 129, No. 10 (Oct.), pp. 423–426.

Grist, M. W. (1972), "Digital Ground Models. An Account of Recent Research," *Photogrammetric Record*, Vol. 70, No. 4 (Oct.), pp. 424–441.

Heiskanen, W. A. and H. Moritz (1967), *Physical Geodesy*, W. H. Freeman, 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.*, (1975), "Computer Applications in Land Use Mapping and the Minnesota Land Management Information System," Davis, J. C. and M. McCullagh (eds.), *Display and Analysis of Spatial Data*, New York, John Wiley and Sons.

Junkins, J. L., G. M. Miller and J. R. Jancaitis (1973), "A Weighting Function Approach to Modelling of Irregular Surfaces," *Journal of Geophysical Research*, Vol. 78, No. 11 (Apr.), pp 1794–1803.

Kraus, K. (1973), "A General Digital Terrain Model," translation from an article in Ackermann, F. (1973), *Numerische Photogrammetrie*, Sammlung Wiechmann, Neue Folge, Buchreihe. Bd. 5, Karlsruhe.

Laboratory for Computer Graphics and Spatial Analysis, Harvard University (1974), *POLYVRT Manual*, Cambridge, Mass.

Mark, David M. (1974), "A Comparison of Computer-based Terrain Storage Methods With Respect to Evaluation of Certain Geomorphometric Measures," M.A. thesis, University of British Columbia.

Morse, S. P. (1968), "A Mathematical Model for the Analysis of Contouring Line Data," *Journal of the Association for Computing Machinery*, Vol. 15, No. 2, pp. 205–220.

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 (1970), "Isarithmic Maps and the Continuity of Reference Interval Functions," *Geografiska Annaler*, Vol. 52, Ser. B, pp. 92–123.

Peucker, T. K. (1972), *Computer Cartography*, Association of American Geographers, College Geography Commission, Resource Paper No. 17, Washington, D.C.

Peucker, T. K. (ed.) (1973), *The Interactive Map in Urban Research, Final Report After Year One*," University of British Columbia.

Peucker, T. K. (1974), *Geographical Data Structures Report After Year One*, Simon Fraser University.

Pfaltz, J. L. (1975), "Surface Networks," *Geographical Analysis* (forthcoming).

Rhynsburger, D. (1973), "Analytic Delineation of Thiessen Polygons," *Geographical Analysis*, Vol. 5, No. 2 (Apr.), pp. 133–144.

Rosenfeld, A. (1969), *Picture Processing by Computer*, New York, Academic Press.

Schmidt, W. (1969), "The Automap System," *Surveying and Mapping*, Vol. 29, No. 1 (Mar.), pp. 101–106.

Shepard, D. (1968), "A Two-Dimensional Interpolation Function for Irregularly Spaced Data," *Proceedings*, 23rd National Conference, Association for Computing Machinery, pp. 517–524.

Sima, J. (1972), "Prinzipien des CS digitalen Gelaendemodells," *Vermessungstechnik*, Vol. 20, No. 2, pp. 48–51.

Switzer, P. (1975), in: "Sampling of Planar Surfaces," Davis, J. C. and M. McCullagh (eds.), *Display and Analysis of Spatial Data*, John Wiley and Sons, New York.

Tobler, W. R. (1969), "Geographical Filters and their Inverses," *Geographical Analysis*, Vol. 1, pp. 234–253.

Warntz, W. (1966), "The Topology of Socio-Economic Terrain and Spatial Flows," *Papers of the Regional Science Association*, Vol. 17, pp. 47–61. ◘

a DATAMATION. reprint

Interactive Graphics Comes of Age

Interactive Graphics Comes of Age

by Eric Teicholz

Continuing reductions in size, cost, and complexity are causing a population explosion in interactive graphics systems.

Imagine an architect first designing a building and then immediately being able to walk around and through it before the building is even built He could walk up to windows and doors, examine them and make appropriate changes if they did not meet his design criteria.

This story is not a fantasy. Dr. Ivan Sutherland, first at Harvard and then at the Univ. of Utah, designed and built a head-mounted display consisting of two miniature crt's mounted in a pair of goggles and mechanically connected to a computer. As the architect turns his head, the computer knows precisely what is being looked at and will generate stereo views of the building as if the designer were actually inside it.

Today, designers have a unique tool that makes it possible to realistically simulate a three-dimensional environment and to make design changes in a faster and more accurate manner than has ever previously been the case—interactive computer graphic systems. Whereas computer graphics had its origin in line drawing (pen on paper) machines, companies representing the "cutting edge" of graphic technology, such as those started by Dr. Sutherland (Evans & Sutherland, Computer Corp.), now make it possible to design three-dimensional figures dynamically using gray tone or color displays.

Some history

Interactive graphics has been with us since the early '60s when Ivan Sutherland developed "Sketchpad," the first interactive system for computer aided design. Early developmental efforts, like GM's DAC-1 system, tended to be based on large, expensive, and specialized hardware. Many early experiments in computer aided design were actually of greater value for promotional rather than practical purposes. The systems and applications software usually demanded dedicated central processors, and incidentally were in many ways incompatible with the newly emerging time-sharing services that required low cost graphics

Vendor	Applicon	Auto-Trol	Bendix	Calma	Computervision	Digital Equipment (Redac System)
1st delivery	Model 700—1970 Model 800—1974	1973	1972	1971	1969—1970	1971
Primary applications	Integrated circuits Printed circuits	Drafting Printed circuits	Drafting, Printed circuits, and Mapping	Integrated circuits Printed circuits	Integrated circuits, Printed circuits, and Drafting	Integrated circuits, Printed circuits, Architecture and Garments
Primary input & edit devices	• digitizer • tablitizer with optional plotter • crt with tablet and keyboard • magnetic tape	• digitizer with keyboard/display • interactive crt with cursor • Teletype ASR 33 • magnetic tape	• digitizer • crt with keyboard • Teletype ASR 33 • magnetic tape	• digitizer with keyboard/display • crt with tablet and keyboard • Teletype ASR 33 • magnetic tape	• digitizer/plotter • crt with tablet and keyboard • Teletype ASR 33 • interface to larger computer	• 17-inch crt with light pen • Teletype ASR 33 or Decwriter • magnetic tape
Maximum input stations	Model 700 4 plus two other devices Model 800 5 plus three other devices	6 stations and two plotters	4 stations	6 stations with three plotters and one tape	4 stations	4 crt stations
Processor	Basic station PDP-11/05 with 24K Additional stations PDP-11/05 with 8K	Varian 620L/200	Basic station Nova 1200 with 24K Additional stations Nova 1210 with 24K	Nova 1220	Nova	PDP-15/76 dual processor with PDP-15 and PDP-11
Typical system cost	$122 000	$126 895 with flatbed plotter	$91 290	$131 000	$110 000 ± 10% with digitizer/plotter but no drum	$109,800 hardware $ 50 000 software
Expansion cost	$18,000—$36 000 depending on size and display	• digitizer station $12 500 • crt station with thumbwheel x-y cursor & keyboard $11,500	Station with digitizer crt/keyboard, ASR 33 and Nova 1210 $35,000	• station with crt and tablet $24 000 • digitizer station $34 000	digitizer/plotter $32 000—$40 000	not applicable

Table 1 Typical minicomputer-based turnkey graphics systems.

**Most of the market estimates and tabular data used in this article are from International Technology Marketing, Inc, Newton, Mass, with which the author is associated* displays and low cpu overhead.

By the late '60s some changes took place in the computer graphics industry Computer manufacturers began to realize the economic and technical potential of the interactive graphics market and more readily supported graphic requirements in their hardware design. As computer memory technology advanced, hardware costs came down, resulting in the emergence of still smaller and faster machines.

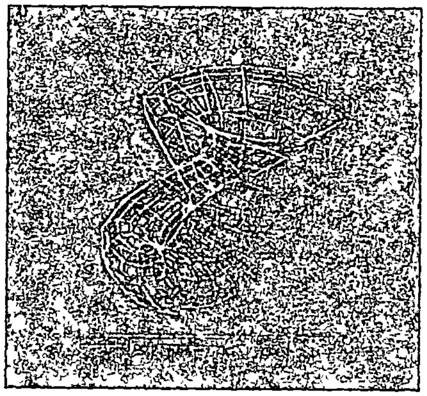
The greatest impetus of all for graphics came from the development of storage tube crt's that were both inexpensive (in the $3-$10,000 range) and could be used as terminals over telephone lines communicating with remote time-shared cpu's.

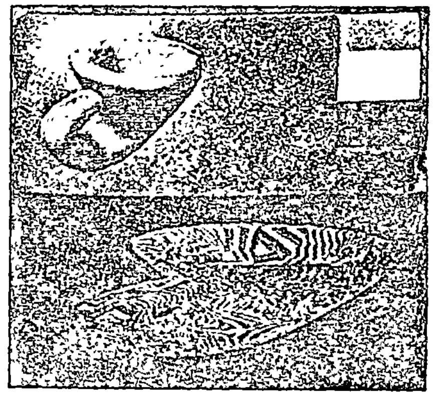
Storage crt's draw pictures on a display surface in a random fashion and the displays remain on the screen until they are erased. The storage tube, unlike its predecessor, the more conventional "refresh" crt, is not used with a display list or menu and can separate memory from display processing requirements—thereby freeing the graphics program from dependency on bandwidth, buffer size and phosphor decay rates. However, to be expected, the storage tube pays a price for these characteristics. because its pictures are not refreshed 30 to 60 times a second and are drawn in an unstructured (random) manner, storage crt's require high "driving" voltages to produce the required beam deflections, and therefore have relatively slow drawing speeds—especially if used over voice grade telephone lines to a remote central processor. Finally, a large class of interactive graphic procedures that are available on the refresh crt (such as selective erasure) are omitted from the storage tube user's repertoire.

In the past five years, the computer industry has been changing at an ever-increasing pace. We have seen the development of intelligent terminals (often containing their own graphic processors); the development of integrated turnkey systems that combine hardware, software, and service from a single source, a rapidly expanding minicomputer market (which will be about $1.24 billion in 1975); a less mature microcomputer market; and larger and faster mass memories.

It is difficult to establish categories for interactive graphic modes of operation. Rather the modes can be represented by a continuum. At one end is the user who communicates with a remote computer using only a storage crt as a terminal. This user pays a little over $200/month for the rental of the terminal and can perform only relatively simple graphic applications because of the data transmission limitations of telephone lines. A reasonably complex picture, for example, can take over 10 minutes to generate if sent over voice-grade telephone lines.

At the other end of the spectrum are





Evans & Sutherland's "Picture System" (from which these photos were taken) and others like it allow designers to work in gray tones or color or both. There are less than 200 such sophisticated (and relatively expensive) systems in use, primarily in basic research, aerospace modeling and simulation, and computer aided design.

Dimensional Systems	**Gerber Scientific**	**GCA/Hampshire**	**Macrodata**	**M & S Computing**	**United Computing**
1972	1973	1973	1970	1972	1974
Drafting and Mapping	Integrated circuits Printed circuits	Integrated circuits	Integrated circuits	Integrated circuits, Printed circuits Drafting and Mapping	numerical control mechanical design drafting
• digitizer with menu and ASR 33 • crt with tablet and keyboard • Teletype ASR 33 • magnetic tape	• digitizer/plotter • keyboard/display • crt/stylus • Teletype ASR 33 • magnetic tape	• digitizer • crt terminal • crt with cursor and tablet • Teletype ASR 33 • magnetic tape	• digitizer • crt, tablet, and keyboard • Teletype ASR 33 • magnetic tape • punched cards	• storage crt • stylus or cursor • data tablet • digitizer • keyboard	• storage crt • cursor control • keyboard • magnetic tape
4 stations	6 stations	8 stations	4 total two crt and two digitizer	8 total	4 stations
Lockheed SUE	Basic station H-P 2100A with 12K Additional stations H-P 2100A with 12K	PDP-11/40 with 24K	Interdata 70	PDP-11	General Automation SPEC 1665 with 32K
$114 500	$120 000 with crt/stylus station	$160,000	$130 000	$100 000	$150 000
40x60 tablet and crt $26 000—$28 000	• digitizer/plotter $50,000 • keyboard/display $40,000 • crt/stylus station $70 000	• edit station $30 000 • digitizer station $40 000	• digitizer station $15,000 • crt tablet, and keyboard $25,000	digitizer, two crts keyboard, and data tablet $25 000	$16 000/terminal

state of the art refresh-type systems consisting of sophisticated self-contained, standalone units, with two- and three-dimensional, and sometimes even color, graphic capabilities. These single-station systems contain large processors and are capable of continuous dynamic motion, zooming, perspective generation and other sophisticated functions. The costs of the display processors alone usually start in excess of $125,000.

Finally, in the middle of the spectrum, are the family of graphic systems called intelligent terminals. These contain various degrees of self-contained computational capabilities and cost anywhere from $8,000 to $75,000.

The leading manufacturer of the storage crt is Tektronix, which has over 10,000 terminals in the field representing a little less than 90% of the total market. Uses encompass just about every application but can be approximated as 75% scientific and 25% business. Many of the business applications are provided by time-sharing companies such as Cyphernetics which not only support the storage crt but offer valuable econometric data bases as well.

The leading manufacturers of state of the art systems are Evans & Sutherland, Adage, and Vector General. Together, there are probably less than 200 such systems in use. Because of their unique capabilities (and because of their price), most are used for basic research (in universities and research centers), modeling and simulation (in the aerospace industry) and, to an increasing degree, for computer aided design applications.

The graphic community has not yet reached a consensus regarding the direction of future technological developments Because of the developments mentioned above, in combination with emerging high speed digital telecommunications networks, however, raster scan or television compatible graphics (which structures data left-to-right and top-to-bottom) very likely will eventually predominate. Many research (MIT, Universities of Utah and North Carolina) and development efforts (Xerox, Datadisc, Evans & Sutherland) seem to point towards the primacy of tv-compatible graphics.

There are many factors that will facilitate movement in this direction television sets provide a low cost terminal, and there are approximately 120 million television sets in the U.S of which almost 50% are in color. Raster scan video memories are lowering in cost and have low power requirements. Gray tones and color outputs are readily achieved on raster scan (tv) displays Finally, raster scan technology has the potential of merging computer graphics with picture processing technology, thereby making possible the mixing and manipulating of photographic images with computer-generated displays.

Mini-based turnkey systems

One of the more successful recent achievements in the computer graphics industry has been the emergence of minicomputer based integrated turnkey systems. The typical system will cost approximately $125,000, and consists of a graphic input station (digitizer, tablet, function keys, joystick or keyboard), an output station (flatbed, drum, light beam, microfilm or electrostatic plotter), an interactive crt work station, a large secondary mass memory (disc, tape or drum) for storing large data bases, the mini and, in some cases, a communications interface to a remote processor.

Software for turnkey systems include both systems and applications capabilities for at least two-, and sometimes three-dimensional, graphic data bases. Table 1 compares some characteristics of turnkey systems as developed by their major producers. Although the hardware varies greatly from system to system, they are all alike in that both hardware and software support is provided by the same company.

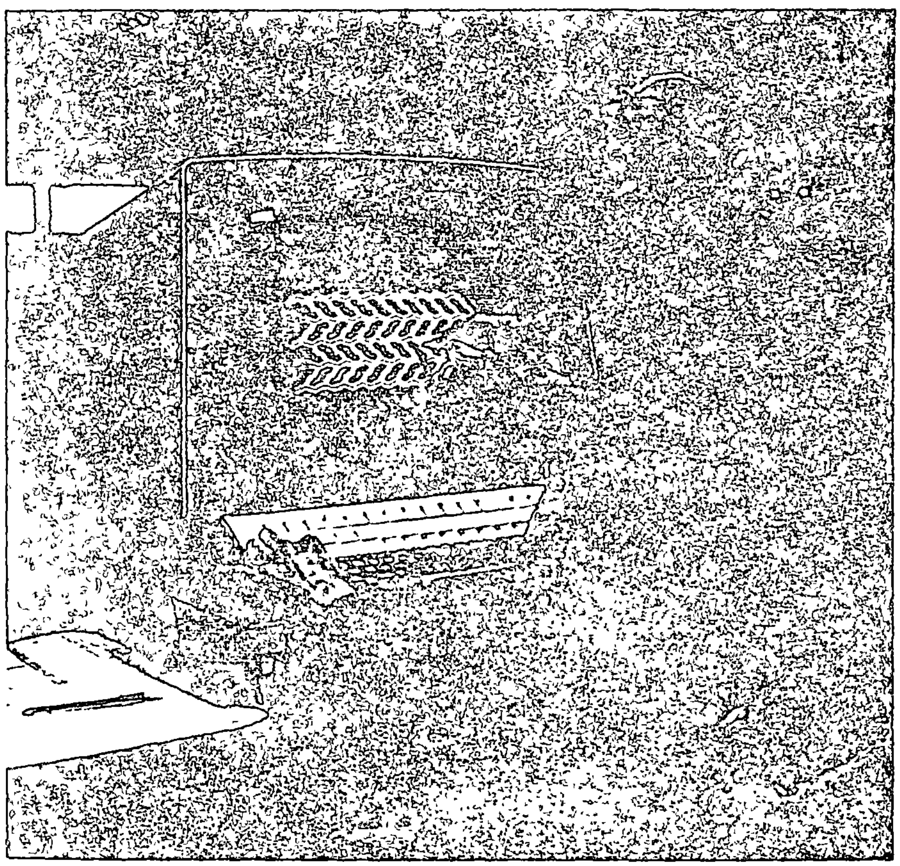
There are now about 500 systems in operation Of these, almost half will have been sold in 1975. In five years the number of systems in use will have doubled.

The current size of the turnkey system market (1974 figures) as split up by the same vendors is as follows:

	$ millions
Applicon	$10.0
Auto-Trol	1.0
Bendix	2.0
Calma	4.5
Computervision	13.0
DEC/Redac	2.0
Dimensional Systems	0.3
Gerber	1.0
Hampshire	0.5
Macrodata	1.5
M&S	0.2
United Computing	0.5
Total	$36.5

The figure is still relatively small, corresponding to only about one-third of the figure for non-interactive graphics products and services, but it is growing.

Also interesting in those figures is that only Digital Equipment, of all the major hardware manufacturers, makes



Firestone Tire and Rubber Co. uses a Sanders/900 system for applications like tire tread design Larger systems like the 900 are usually found in companies with sales over $50 million that can take advantage of multi-shift usage one of these systems, and that three vendors (Computervision, Applicon and Calma) share over 75% of the market. Computervision alone has over 200 systems in the field, reportedly.

Today's applications

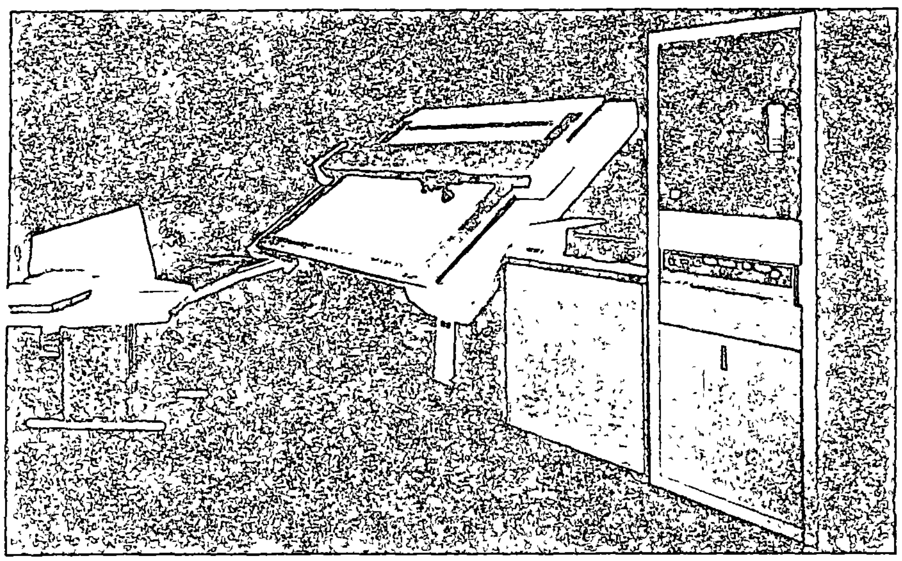
Most integrated systems are used for applications related to electronics (75%), drafting (15%), and cartography (5%), with architecture, engineering, plus university and government research making up the remaining 5% The latest published and forecast figures for sales of turnkey systems by industry are:

	$ millions 1975	1978
Electronics	$58	$150
Drafting	12 5	60
Mapping/cartography	7	20
Architecture/engr	2.5	3
Govt/univ research	3.5	7
Totals	$83.5	$240

Electronic applications encompass design and layout of wiring and circuits for printed circuit, integrated circuit, and hybrid circuit production The predominating application is the generation of artwork master and automated machine controls for the production of PC boards and for process masks used in IC production A principal advantage of graphic systems here (as for all application areas) has been the ability to create and store graphic data bases which can be easily recalled and revised by computer (the 'big eraser' concept).

Two-dimensional general purpose drafting for electrical, mechanical and piping systems has been commercially viable for some time. Most developmental work in drafting for electrical, schematic and mechanical work relates to three-dimensional data base systems as 3D drafting capabilities are required for manufacturers' applications in the ordnance, chemicals, refining, machinery and metal products industries. Unfortunately, such systems are extremely sophisticated and complicated to use Consequently there are only about 25 3D systems in use today. Response thus far has been a "wait and see" attitude. Progress is being made but it will be another year (or two) before the required system flexibility, file response and access criteria are achieved.

Computer mapping applications have enjoyed a rapid growth primarily because of the increased availability of geographic and statistical data bases such as census files. Furthermore, the forthcoming national conversion to the metric system will necessitate the redrawing of millions of base maps which can best be achieved by computer. General purpose mapping software is also readily available and, except for topographic maps, extremely high accuracy output is usually not required. At present, Applicon estimates that approximately 20% of its systems are used for mapping and engineering construction.



Computervision's "Designer System" includes a console, plotter/digitizer, disc, mag tape, and minicomputer. Approximately 200 of the turnkey systems have been put into the field, according to the manufacturer.

To date, most users of standalone interactive systems are large companies whose sales are over $50 million. These are the companies who achieve the greatest cost benefits from multi-shift usage Feedback from the user community indicates that benefits from integrated systems include shorter production times, design standardization, establishment of graphic data bases, improved accuracy and greater design flexibility, especially in terms of making design changes Of all these, design standardization and data base establishment seem to be the most commonly noted. (As an example, the article by W. Barkley Fritz and Charles R. Lansberry which follows these pages, describes how interactive computer graphics is being used at the Sun Shipbuilding and Dry Dock Co.)

Tomorrow

Integrated graphic systems of the future will become smaller (in terms of standalone capabilities) and less expensive than present systems. New systems will be designed for use with a host computer and will provide general local picture processing capabilities such as data base creation, graphic editing and interrogation, file formatting and the like. It will not only be easy to enter the geometry and topography of data base components, but attributes (such as cost, manufacturer, color, etc.) will be assigned to components as well. When this happens, and when such a system is available for under $50,000, entire new classes of users and new application areas will open.

Turnkey systems have proven to be a viable and lasting force in the graphics market. Psychological barriers to their use that existed just a few years ago have largely been overcome Skepticism has been replaced by respect and a sincere desire that machines will be able to assume an ever-increasing role in performing the drudgery of repetitive graphic bookkeeping chores Only 10 years ago designers were primarily interested in automating the design process and in "pushbutton" engineering Today, successful applications encompass a spectrum of activities from conceptualizing designs to production. Emphasis is on man-machine interaction rather than on man or machine action alone It is this approach that has produced the most successful results to date and will continue to do so in the near future. ✲

Prof Teicholz is associate director of the Laboratory for Computer Graphics and Spatial Analysis at Harvard Univ , where he teaches in the Graduate School of Design He was a member of the study team at International Technology Marketing, Inc , which produced the reports from which some of these findings were taken

GRDSR: Facts by small areas

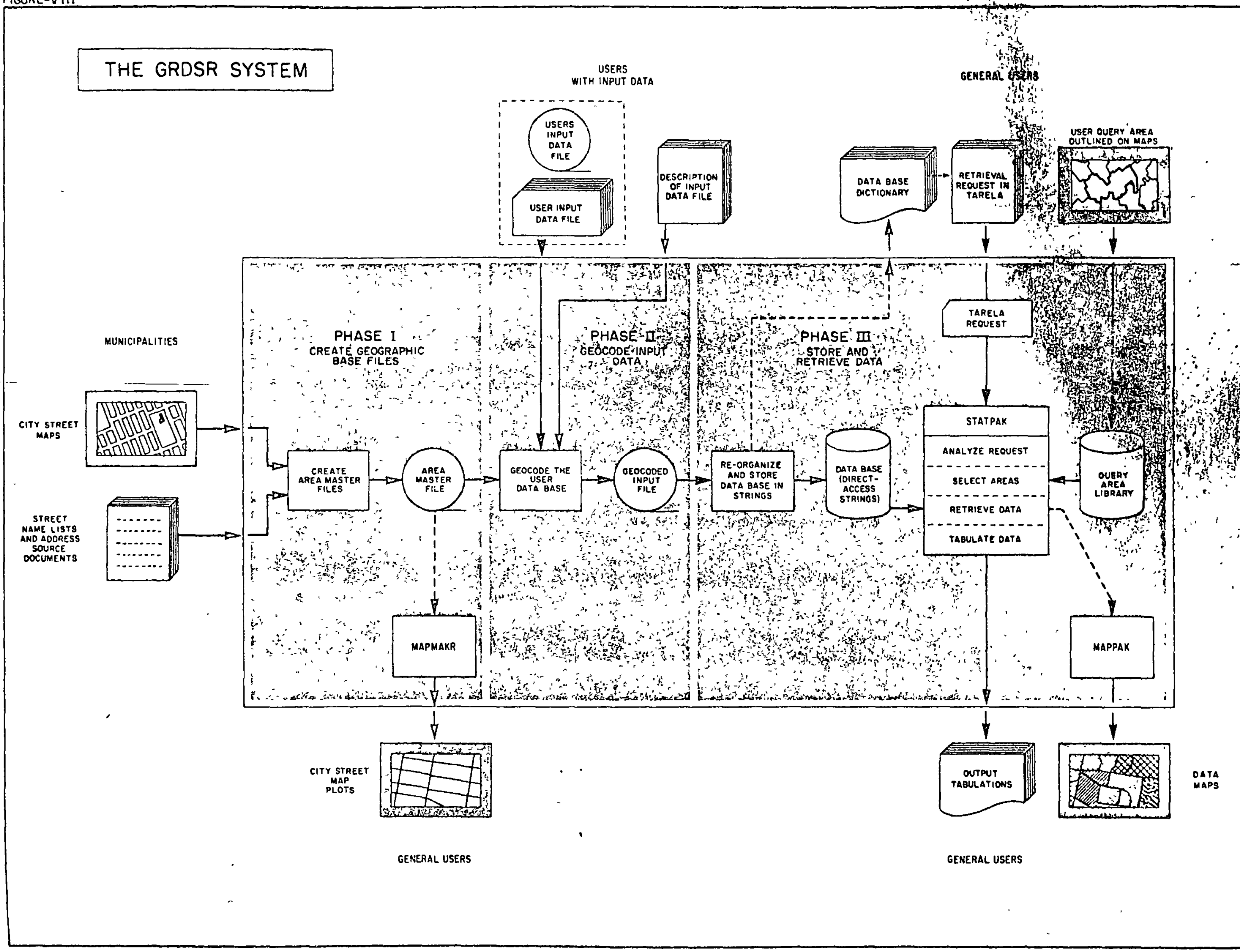
The Geographically Referenced Data Storage and Retrieval System

AN INTRODUCTION

JUNE 1972 i.e. GIS, Geographic Information System

Statistics Canada Statistique Canada

FIGURE-VIII


THE GRDSR SYSTEM
USERS WITH INPUT DATA
GENERAL USERS
USERS INPUT DATA FILE
USER INPUT DATA FILE
DESCRIPTION OF INPUT DATA FILE
DATA BASE DICTIONARY
RETRIEVAL REQUEST IN TARELA
USER QUERY AREA OUTLINED ON MAPS
MUNICIPALITIES
PHASE I
CREATE GEOGRAPHIC BASE FILES
PHASE II
GEOCODE INPUT DATA
PHASE III
STORE AND RETRIEVE DATA
TARELA REQUEST
CITY STREET MAPS
STREET NAME LISTS AND ADDRESS SOURCE DOCUMENTS
CREATE AREA MASTER FILES
AREA MASTER FILE
GEOCODE THE USER DATA BASE
GEOCODED INPUT FILE
RE-ORGANIZE AND STORE DATA BASE IN STRINGS
DATA BASE (DIRECT-ACCESS STRINGS)
STATPAK
ANALYZE REQUEST
SELECT AREAS
RETRIEVE DATA
TABULATE DATA
QUERY AREA LIBRARY
MAPMAKR
MAPPAK
CITY STREET MAP PLOTS
OUTPUT TABULATIONS
DATA MAPS
GENERAL USERS
GENERAL USERS

GRDSR (The geographically referenced data storage and retrieval system)

A new method of assembling statistical information by user-specified areas

AN INTRODUCTION

Published by Authority
of the Minister of Industry, Trade and Commerce

June 1972

CONTENTS

INTRODUCTION

What is GRDSR?

A unique and flexible system now makes it possible, for the first time, to provide information by user-specified areas in Canada's larger urban centres. Fully computerized, the GRDSR (Geographically Referenced Data Storage and Retrieval) system is the outcome of five years' research by Statistics Canada into solving the many problems associated with the storage and retrieval of statistics about small areas.

Through GRDSR, statistical information can now be quickly and inexpensively obtained about retrieval areas that range in size from a few city blocks to an entire urban centre.

Retrieval is made possible through a technique called geocoding, whereby urban areas are divided into many small building blocks or micro-areas. The blocks must be small enough that they can be assembled to approximate most retrieval areas required by users. Each building block is assigned a unique identifying coordinate number which, in turn, allows files of households, persons, or events to be coded to appropriate building blocks in the city area. The appropriate building block is usually the place of residence or location where the event occurred. At this point, the files are said to be geocoded. When an interested user needs information from a geocoded file, he outlines his area of interest (or "query area") on a map and makes a request. GRDSR then identifies all the building blocks contained within his query area and, using the corresponding coordinates, automatically retrieves all data belonging to the blocks. The statistics are then tabulated in a convenient report.

Background to GRDSR development

The need for concise, timely statistics is constantly pressing in many sectors of the economy. Diverse planning and decision-making efforts are often frustrated by a lack of relevant and timely data; the socio-economic benefits of more fully informed decisions may, as a result, be diluted or lost. Today, it is obvious that the pressure for diverse and specialized statistical information cannot help but increase. The comprehensive and thorough use of data already collected is now, therefore, more relevant than ever.

At Statistics Canada, this pressure is evident not only in the mounting volume of special data requests but by their changing nature. The trend consistently points to the need for flexible information systems fully capable of retrieving data on a specialized, often one-time basis. The basic requirement is, essentially, for an integrated information service – not just a data collection facility.

The development of GRDSR thus focuses upon an important trend, and the nature of this trend is clear: user requirements will increase, in terms of the amount of data required, types of aggregation and manipulation available, ease of retrieval, the format of the final statistical product and, of course, the response time. The evolution of the GRDSR system is now at the point where each of these requirements is substantially met.

Nature of the Problem

The gathering of small-area data has presented a difficult problem for some time. Urban planners, municipal agencies, school administrations and governments each impose different zoning patterns or jurisdictions over settled areas of land. Many agencies maintain records and use their own jurisdictions to collect and identify statistical information. At some point these records may attract general interest. But problems arise when outside groups try to use this information, because their requirement is for facts related to different area breakdowns.

Today, special-interest areas such as marketing zones, census tracts, school districts and land-use areas are in everyday use in major cities. However, these areas usually overlap and have little in common but the land area they reference. Thus, it is difficult to relate information from one source to outside areas of interest (see Figure I, page 2.)

In the past, when the sole means of disseminating statistics from the census was through published volumes, the statistics had to be summarized in terms of enumeration areas, census tracts or other standard areas. The standard census areas did not, however, coincide with many query areas for which data were required. Consequently, the requirements of many census data users could either not be met, or met only with great difficulty, at considerable cost and with considerable delay.

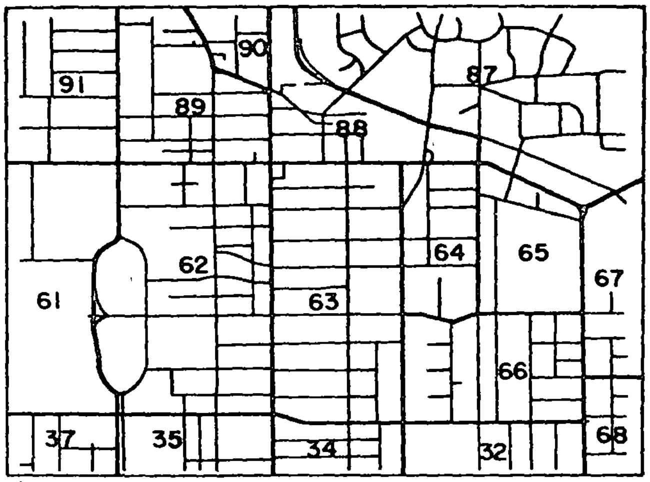
Where can GRDSR be Applied?

Given that the fundamental purpose of GRDSR is to allow users more flexibility in obtaining information about special-purpose areas, it is significant that the first major application of GRDSR has been the 1971 Census of Canada.

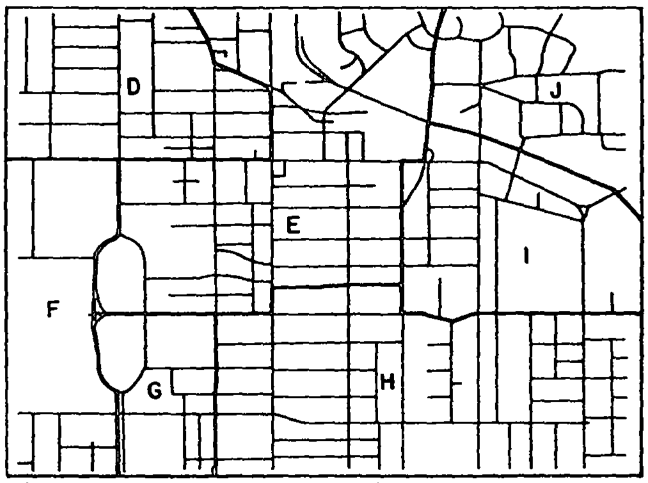
Originally conceived in anticipation of special census requests, the system has since been developed for general-purpose applications Municipal assessment files, fire and accident reports, marketing surveys and hospital records are among several applications discussed in the next section.

FIGURE - I

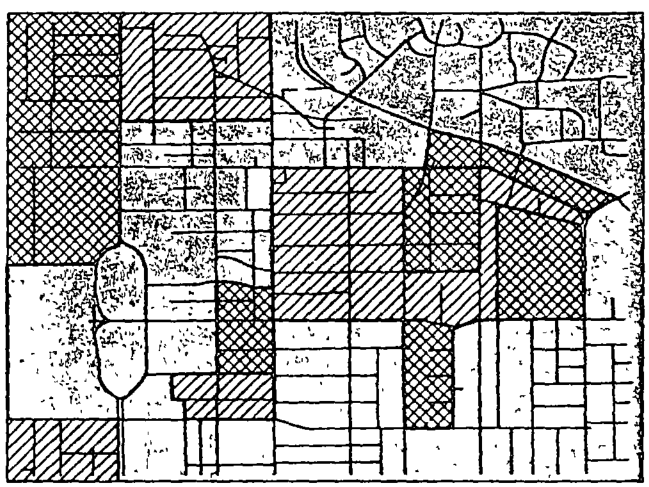
HOW USERS IMPOSE DIFFERENT ZONING PATTERNS OVER A CITY AREA



(1) CENSUS TRACTS



(2) SCHOOL DISTRICTS (HYPOTHETICAL)



(3) PLANNING ZONES (HYPOTHETICAL)

RESIDENTIAL OR PARKLAND — COMMERCIAL — INDUSTRIAL

APPLICATIONS AND POTENTIAL OF GRDSR

While Statistics Canada expects to serve many requests for statistical information from geocoded 1971 Census data, GRDSR is designed to handle the majority of address-bearing files and survey data which originate in larger urban centres in Canada (see Figure II, page 5.) Extensive geocoding applications are now possible in both the public and private sectors. Potential users include municipalities, planning and research groups, industrial and commercial firms, public utilities, social agencies, universities and governments – in short, any group using geographically-based information for research, planning or decision-making.

Noteworthy features of the Census application are outlined in part (i). Next, in part (ii), a number of other specific GRDSR applications are discussed. Finally, some aspects of a possible geocoding application, health services planning, are described in part (iii).

The Geocoded Census

GRDSR will provide a new dimension in census retrieval services: the facility to provide statistical data for user-specified areas anywhere in Canada.

Confidentiality

While Statistics Canada attaches great importance to meeting the need for custom-made, user-oriented data on a uniform, national basis, it can only do so within the confidentiality constraints imposed by the Statistics Act (1971). As a result, no information can be disseminated in such a way as to identify an individual respondent. Automatic routines within the system ensure that no such disclosure of information is possible.

Query Areas

In 14 larger urban centres (see Figure II, page 5), users may request data for areas as small as a few city blocks. Users should not, however, expect to receive representative data for smaller areas, such as one side of a block. There are two important reasons:

First, the results would be subject to high response and sampling errors, due to the small number of cases on which the statistics would be based. The usual process of compensating errors for larger samples could only take place to a limited extent.

Second, a carefully-controlled amount of statistical error is purposely introduced to all retrieved data so that no census respondent can be identified from the final tabulations. This random error, while of little or no significance to normal tabulations, would further obscure any information obtained about very small areas.

Outside the major urban centres, statistical information will be available at a coarser level of geographical detail. Here, query areas will be assembled using traditional census enumeration areas (EA's), which contain approximately 150-200 households each. As a result, extensive census data will be available for more than 27,000 EA's – either individually, or in any aggregation of interest to the user.

In either case, the desired areas are simply outlined on a suitable map, named clearly and submitted to Statistics Canada along with the tabulation request.

Request Language

Users may request census tabulations using an English-like language called TARELA (Tabulation Request Language), which can be learned in a few hours without previous programming knowledge. TARELA allows subject-matter specialists to write requests in terms familiar to their work. With this language users can create cross-tabulations of any combination of 1971 Census variables (which number more than 120) and generate tables having up to 10 dimensions Users who are not familiar with TARELA can, of course, submit their request in precise narrative form or in the form of "dummy" tables. The required TARELA coding will then be generated by Statistics Canada.

Data Mapping

In addition to supplying census data in tabular form, GRDSR also includes a facility for data mapping. MAPPAK, which incorporates the Harvard mapping package SYMAP, is a remarkable feature in that it can accurately depict the distribution of data values over any area in graphical form. This type of map is particularly useful in locating areas where extreme values of some factor occur, and can be used to reveal problem areas at a glance.

Both TARELA and MAPPAK are general-purpose features of GRDSR, by no means limited to census applications. TARELA and MAPPAK are further described in Features and Components, page 10.

General Applications

Geocoding applications can be served by many data bases in addition to the Census The system is designed to geocode many types of address-identified files, provided they originate in one of the larger Canadian urban centres (See Figure II, page 5.) The GRDSR programs will be available to municipal users wishing to geocode local data bases. Examples of suitable data bases include assessment rolls, traffic surveys, hospital and welfare records, marketing surveys, school census data and certain accident, fire and police records.

Municipal Administration and Government

Public Services: Research studies are being conducted, using geocoding, to determine the frequency of accident, fire and police reports originating from various sections of large cities. Such statistics would clearly be a significant aid in planning or re-allocating municipal resources and services; the use of GRDSR is possible whenever records of such incidents are address-identified.

Education: A new method for planning the location of new schools and school districts is now possible through geocoding. Facts related to this application may include the concentrations and age distribution and projected growth rates of school-aged children within the community.

The routing of school buses is another application where geocoding offers considerable promise. GRDSR is ideally suited to provide statistics such as the geographic distribution of school-aged children.

Other applications include analysis of districts by such socio-economic factors as country of origin, language, religion, occupation and income as an input to planning of day-school curriculums and adult-education programs.

Urban Planning

Interests in the urban planning area include study and analysis of planning zones, optimizing the location of city services and facilities, planning of mass transit and analysis of potential urban renewal areas, land values and housing data.

In planning the route of a new city transit system, for instance, the starting points and destinations of potential users form a definite network or pattern. Subject to further analysis, such as transportation modelling, this network can have decisive impact on the final route chosen.

Further possibilities include planning of municipal services according to socio-economic factors such as population density, language, and income within selected urban zones. New approaches to planning the nature and location of welfare services may also become possible.

Medical Services

Typical problems include planning the location of hospitals, out-patient clinics and medical centres, and the establishment of a geographically-referenced inventory of nurses.

Industry, Commerce and Utilities

Geocoding has played a part in the allocation of facilities and services such as telephone exchanges and banks. Other applications include population and demographic studies of city areas, the planning of marketing zones and radio and television coverages, the optimization of retail store location in terms of customer proximity and resource allocation problems faced by oil, hydro and gas utilities.

A number of simulation and modelling techniques exist for solving network problems in the commercial transportation/distribution area. Typically, data related to some grid pattern constitute an essential requirement for this approach. GRDSR is an ideal research tool to help meet this need.

Universities

Interests include economic, political and social studies of neighbourhoods, electoral districts and socio-economic research into city areas defined by such factors as country of origin, language or income.

Health Services Planning: A Potential GRDSR Application

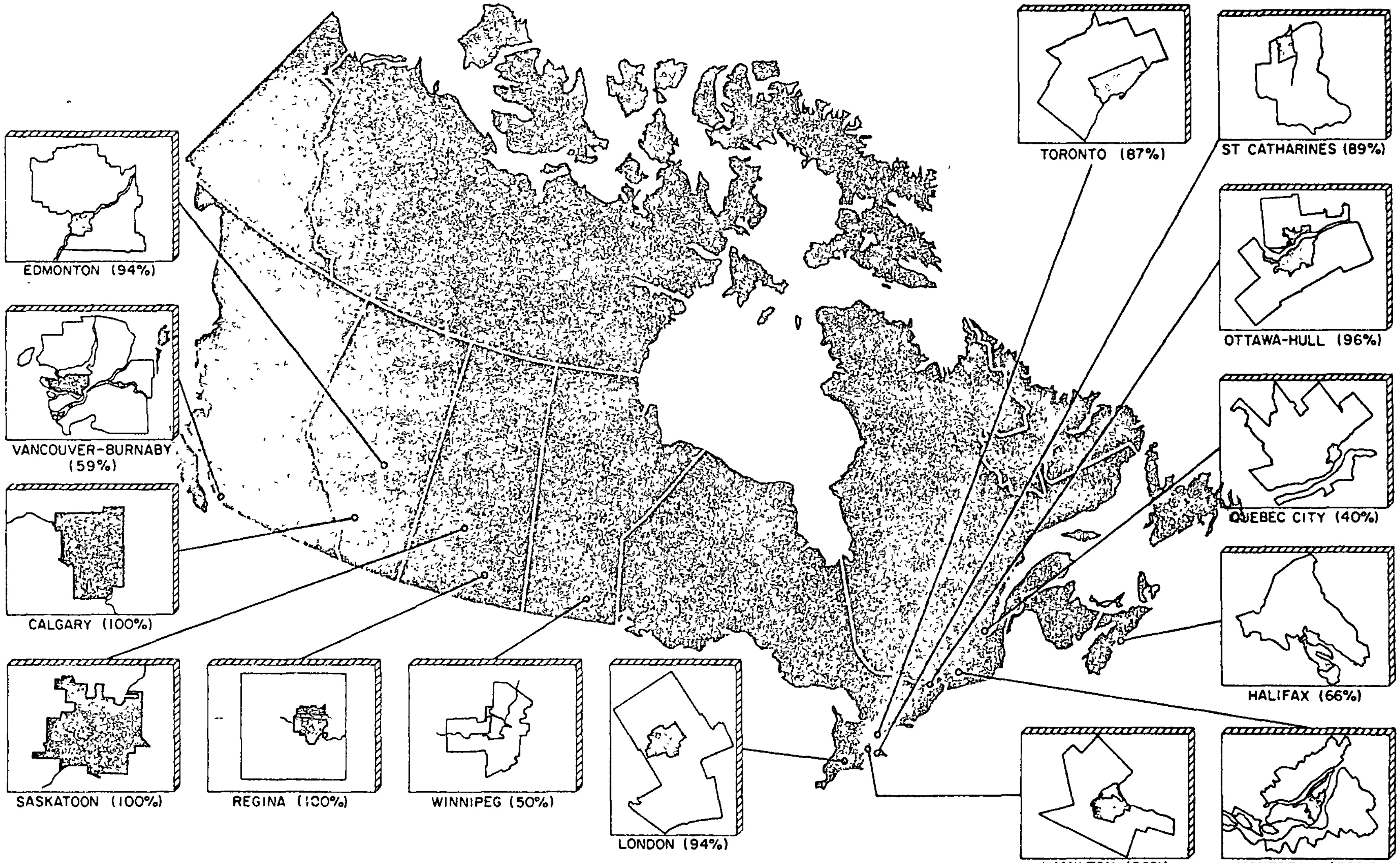
A number of factors influence the choice of location for a new hospital or health services clinic in a major city, such as accessibility through major traffic arteries, availability of professional staff, areas most in need of services.

GRDSR can be particularly helpful in deciding which city areas are most in need of proximal medical facilities. One approach is to find out where past patients have lived and what medical services and equipment they required, using city hospital records.

Hospital visitation records bear, in addition to medical content, an address identifier for every patient. Therefore geocoding operations can, in most cases, be carried out on the visitation records. Through GRDSR, considerable statistical information can then be generated (for instance, the incidence of hospital visits originating from each and every portion of the city). The retrieved information can be cross-tabulated by the type of medical services required or by any other item of information contained in the original records. For example, the incidence of various diseases, illnesses, or special health problems in certain city areas can be ascertained. Such statistics can prove to be an invaluable aid in determining which city areas would best be served by neighbourhood medical services or a new hospital.

FIGURE- II

GEOCODING COVERAGE AT THE BLOCK-FACE LEVEL (JUNE, 1971)



IN EACH DIAGRAM, THE OUTLINE DEPICTS THE BOUNDARY OF THE CENSUS METROPOLITAN AREA WHILE THE SHADED PORTION REPRESENTS THE AREA OF BLOCK-FACE COVERAGE THE PERCENTAGE OF POPULATION RESIDING IN THE SHADED AREA IS ALSO SHOWN DIAGRAMS ARE ALIGNED IN THE NORTH-SOUTH DIRECTION

APPROXIMATELY 7 MILLION, OR 34% OF THE POPULATION OF CANADA ARE NOW COVERED AT THE BLOCK-FACE LEVEL

CONCEPTS AND METHODS

Review of small-area problems

An urban planner, faced with comparing the expropriation costs of several expressway routes, might attempt to use municipal assessment files and find that records were identified by address, city wards or in some other way. To obtain statistics about land values and dwelling types, the file must be inspected one record at a time to determine which data to include in estimates for the proposed expropriation area. The expense of this approach has been prohibitive but, until recently, few alternatives were available.

Another type of requirement, now directed to the census, might be phrased as follows:

> "*A tabulation of the number of people resident in the Toronto area bordered by Summerhill Avenue, Yonge Street, Mount Pleasant Cemetery, and the boundary for East York is required. Break this tabulation down by age, sex, income, country of origin and occupation.*"

Alternatively, another request might read:

> "*Provide the same statistics for the area named Ward Five, as outlined on the attached city map.*"

Such requests have been difficult to service, since census data have been summarized by census tracts and enumeration areas, which may not coincide with the required boundaries.

To solve small-area data requests economically, Statistics Canada required an efficient system to repeatedly assemble and tabulate information according to arbitrary special-interest areas. Before describing the conceptual aspects in detail, let us expand upon the operational steps in GRDSR.

Before a data base can be geocoded, each record must be assigned some reference code which identifies the record to its proper geographical source. In GRDSR, the source of each record or data observation is precisely located using a comprehensive geographical coordinate system. Each record is assigned a coordinate value, or "key", which actually becomes part of the record during the geocoding operation. The geocoded file is then stored for later use. Ultimately, at retrieval time, GRDSR automatically identifies each query area with a list of coordinate values and, using the coordinate values as keys, retrieves the precise set of data records required. The retrieved information is then summarized according to the tabulation request; the user receives statistics in the form of a convenient report.

The UTM System

The coordinate system chosen for GRDSR is known as the UTM (Universal Transverse Mercator) System. UTM is an established international convention for specifying point-locations on the globe, and is shown on the popular National Topographical Map Series produced by the Department of Energy, Mines and Resources. This system divides the globe into 60 vertical zones. Altogether, 16 zones cover the land mass of Canada. Each zone has a width of 6 degrees longitude and a central meridian which becomes the vertical axis for the zone. The horizonal axis is formed by the earth's equator.

In UTM, point-locations within a zone are based on two distances in metres (one easting, one northing) from the zone axes. The central meridian is assigned an artificial value of 500,000 metres easting; the equator is assigned the value 0. Distances are measured on a plane rectangular grid onto which the zone's surface features have been projected. The two values are combined with a zone number to arrive at a unique coordinate value for every point on the land mass of Canada.

For example, the UTM coordinates of the Peace Tower, Ottawa, are:

Zone	X	Y
18	445177	5030250

In this way, the UTM coordinates seem to define a point-location to the nearest metre, although the projection of the earth's surface onto a plane grid introduces minor distortions.

Basic definitions

Points at which streets intersect or curve sharply in the city pattern are referred to as *nodes*. Every street is represented by a series of nodes connected by straight-line *segments*.

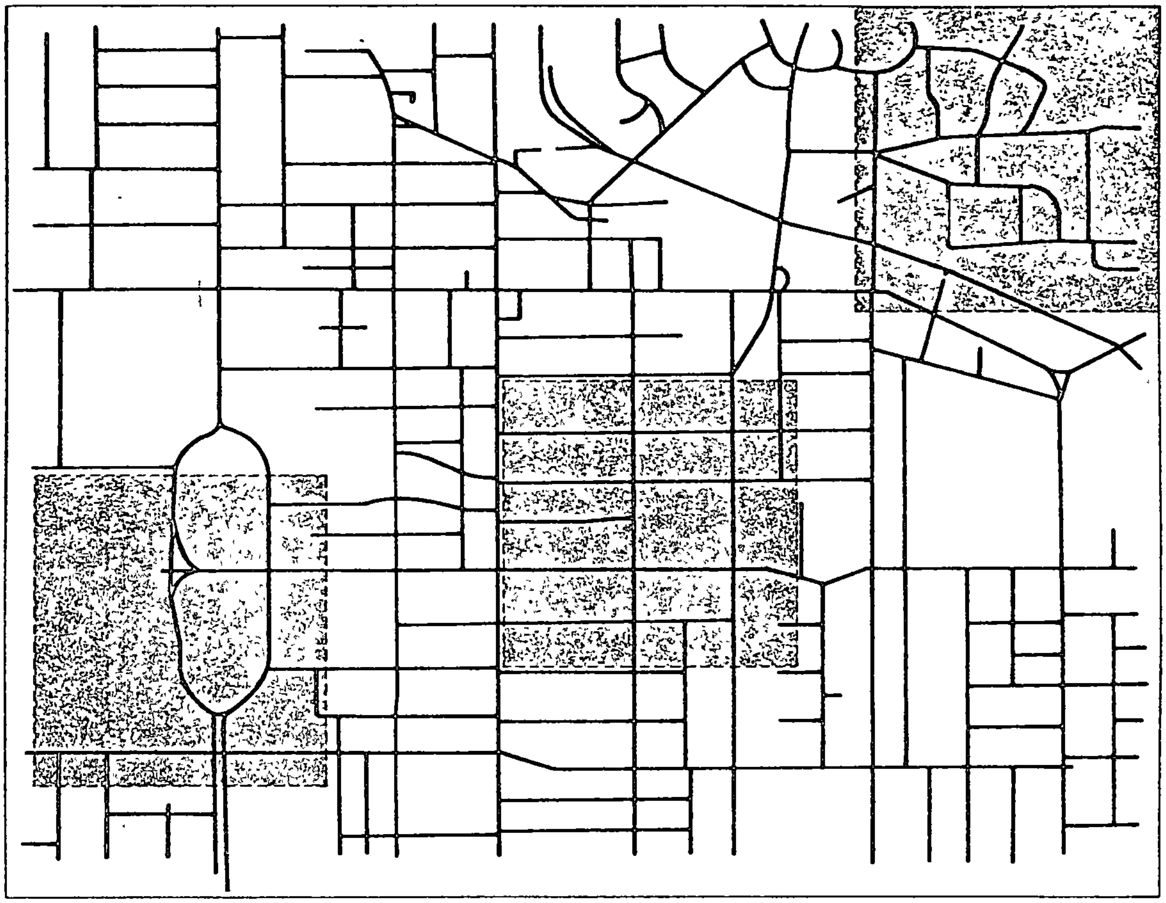
A *block-face* is defined as one side of a city street between consecutive intersections with other streets. Thus, up to two block-faces can be formed by a pair of adjacent nodes, each located at a four-way street intersection. However, a block-face can also encompass several nodes. For example, a block-face may contain one intermediate node marking a change in direction and another node representing an intersection on the opposite side of the street only (see Figure III, page 7).

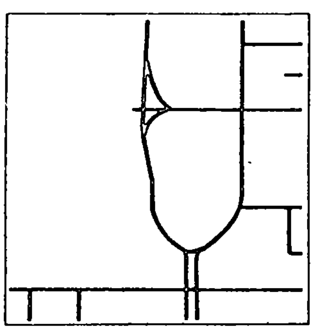
Whenever a block-face is to be formed by a pair of nodes, these nodes must constitute the beginning and end of a valid civic address range.

In this way, block-faces become the basic building blocks used in the GRDSR System.

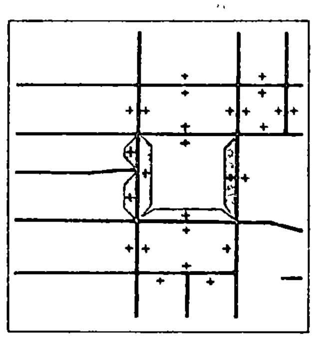
FIGURE III

HOW BLOCK-FACES AND CENTROIDS (+) ARE CHOSEN

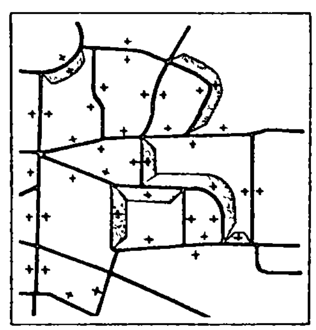




(1) STREETS THROUGH RECREATIONAL OR PARKLAND --NO CENTROIDS



(2) STREETS IN REGULAR (GRID) PATTERN



(3) STREETS IN IRREGULAR PATTERN

 THE SHAPES OF SEVERAL BLOCK-FACES ARE SHOWN BY SHADED AREAS

Why addresses are necessary

The GRDSR System is partly based on the premise that most agency records and survey responses are identified, geographically, by the addresses of respondents. An address is the starting point in coordinate assignment, because every street address in an urban centre can be identified as belonging to some block-face.

How addresses are converted into coordinates

In GRDSR, all street addresses along a block-face are assigned, and share, the coordinates of the block-face centroid, which is simply a reference point offset from the street midway between the two nodes forming the block-face. During the conversion operation, the address of each record or data observation is matched to a block-face (using a list of valid street names and address ranges). From there, the correct centroid is known and its coordinates can be added to the record.

The Area Master File

The actual geocoding operation (or assignment of coordinates to data) is carried out using GRDSR components known as Area Master Files (AMF), which will be described in detail in Features and Components, page 10.

Area Master Files contain a logical representation of all city streets, plus some other features, in computer-readable form. An AMF references every street, address range, block-face and centroid coordinate in the covered area. Also itemized are other features (such as railroad tracks, rivers, and municipal boundaries), which help users to choose query areas. During the geocoding operation, centroids are obtained by matching addresses against street names and address ranges within the Area Master File. (In this way, address ranges can be thought of as representing the actual building blocks, rather than block-faces.)

Area Master Files have been created for major portions of 14 Canadian urban centres, which include a total of 16 cities (see Figure II, page 5). These files reference more than 225,000 block-faces, corresponding to a population figure of approximately seven million.

Rural Geocoding Coverage (1971 Census)

The 14 Area Master Files have already served to geocode certain urban portions in the 1971 Census For the remainder of Canada not covered at the block-face level (urban and rural), census geocoding, as already noted, has been carried out using standard enumeration areas, with one centroid assigned to the approximate population centre of each. Enumeration Areas outside Area Master File coverage number more than 27,000. Tabulation requests for query zones in rural areas or in the urban shadow of developing urban areas are easily (and automatically) handled using centroids at the EA level, the block-face level, or both. During retrieval, inaccuracy in data selection at the EA level is minimized by the choice of centroids near the population centre and by a process of compensation, whereby errors from including or missing centroids in a query area are self-cancelling.

ADVANTAGES, LIMITATIONS OF CONCEPTS

Choice of block-faces

Block-faces become the finest level of resolution possible when each cluster of data observations or survey records belonging to one block-face share the same centroid coordinate. This is a logical outcome of the building block principle adopted by GRDSR.

As one alternative, geocoding to the land parcel or household level achieves higher resolution which may be desirable for some purposes. This approach requires extensive local research. Since block-face resolution is expected to satisfy the vast majority of geocoding requests, land-parcel geocoding could not be justified for a Canada-wide system such as GRDSR.

A second alternative was to identify data by city block, a poorer resolution. However, this approach would not have allowed users enough flexibility, since the integrity of city blocks would have to be respected in specifying query areas. For instance, it would not be possible to obtain tabulations for one side (or both sides) of a city street.

The choice of block-faces as basic geocoding building blocks has several implications. All observations originating from one block-face bear the coordinate of its declared centroid. As a result, the integrity of block-faces should preferably be respected in specifying query areas for retrieval. They should not be split: observations referenced to a split block-face will appear in the results only if the query area includes the block-face centroid. If not, the observations are missed entirely. Another implication is that geocoding to the household or land parcel level (each individual property bears a centroid) is not possible using this system. This may pose definite restrictions on municipal services, engineering and land-banking applications where higher resolution is required.

Identification by street address

Statistics Canada recognizes that a majority of statistical surveys and agency records are address-identified and provides for this with a System component known as the Postal Address Analysis System (PAAS).

In geocoding a file, addresses are analyzed and converted to centroid co-ordinates. Because the conversion is done by computer, complete addresses must be decomposed into separate, clearly identified components (such as street name, type, house number and municipality name). Because PAAS achieves a high efficiency and success rate, address specifications of relatively poor quality can still be geocoded. This feature clearly extends the scope of GRDSR applications. More information about PAAS is provided in Features and Components, page 10.

However, GRDSR cannot perform the geocoding operation on records which, by their nature, are not identified to street addresses. Certain city facilities, such as sewers, gas and hydro lines, traffic signals and overhead structures may be of interest from a geocoding standpoint. In this case, the user must geocode the file before submitting it to GRDSR.

Choice of coordinate system

While UTM is ideally suited to geocoding at the block-face level it has some limitations in land survey and civil engineering operations where the 3 Transverse Mercator System is more accurate, and thus a frequent choice. However, programs are available to convert files geocoded with the UTM system to 3 TM and vice-versa.

FEATURES AND COMPONENTS

The Area Master File

How an AMF is created

Geocoding starts with an accurate street map. A large-scale, current map showing block-face address ranges is required, together with an up-to-date street index. After the map is divided into sections a node is assigned to each street intersection. Nodes are also assigned to points where streets begin, end or curve sharply. A non-distorting overlay is prepared for each map section and the position of each node is marked on the overlay.

Once serial numbers have been assigned to the nodes, descriptive codes for every street segment are transcribed onto a specially-prepared form. The codes include feature names, types, directions, node numbers, and addresses at the intersections. Then the overlay is placed on a digitizing table. The digitizing equipment measures node positions relative to control points on the overlay, and generates one horizontal and vertical "table" coordinate for each node. Since the UTM coordinates of the control points are known in advance, the UTM coordinates for the nodes can then be calculated from the table coordinates. During subsequent computer processing, centroid coordinates are calculated for each block-face using the coordinates of the two nodes bordering the block-face. Finally all items are merged to create an Area Master File for the city (see Figure IVa, page 11).

How the AMF is used

Three operations, each related to address conversion, require files of information contained in the Area Master File. To eliminate the maintenance and updating of three separate files, each is derived from a clean, up-to-date AMF as required.

(i) Street name lists are used by PAAS to verify input addresses prior to the assignment stage.

(ii) The Address Conversion File (ACF) is used to obtain centroid coordinates for input addresses once the PAAS stage is complete. Addresses are matched against block-face address ranges and the corresponding centroid coordinates are selected from the ACF. Geocoding is complete once centroid coordinates replace addresses in the original file.

(iii) The Block-Face File was created specifically to geocode the 1971 Census. This file makes it possible to link parcels of census data, which are not otherwise address-identified, to block-faces and centroids.

As a geographic base file, the Area Master File design is unique. The central concept is to provide a geographical framework that is as practical as possible for a variety of potential users, but efficient from a file creation/update standpoint.

A series of error-handling and correction procedures comes into play whenever Area Master Files are being built or updated. Extensive computer checking is done to ensure that each node is linked to the correct street segments, and vice-versa. This process locates the majority of clerical errors. When each section file is complete, it is plotted at the same scale as the original map. The two maps are then compared to verify node locations. Usually, further plotting followed by two to three update cycles, will produce a clean Area Master File.

Local area breakdowns, such as census tracts, electoral wards, city wards, and other extra codes were purposely excluded from the AMF. Its design is such that these jurisdictions are easily constructed independently of the AMF, but using the identical building-block technique. Because areal boundaries are constantly changing, their inclusion would have seriously prolonged the operation needed to build and maintain an accurate, up-to-date base file.

Urban Street Maps

Computer-plotted street maps are an important by-product of building an Area Master File. Because the AMF is a logical representation of city features, its contents can be used, in reverse, to create facsimile maps at any scale. Plotting is accomplished using the GRDSR component, MAPMAKR (see Figure V, page 15).

These maps have several purposes:

- The best way to edit or validate an Area Master File is to recreate the original map, using the plotter. Errors and inconsistencies are clearly highlighted.
- They provide a return service to municipalities who in turn are aware of what updates are required.
- The maps are supplied to users for outlining query areas and depict city features as seen by the AMF.

Using MAPMAKR, maps can be produced to suit a variety of purposes. It is possible to pre-specify the area to be plotted and the scale required. Parameters are used to determine whether various options, such as nodes, feature names, centroids, address ranges and control points, will appear on the final plot.

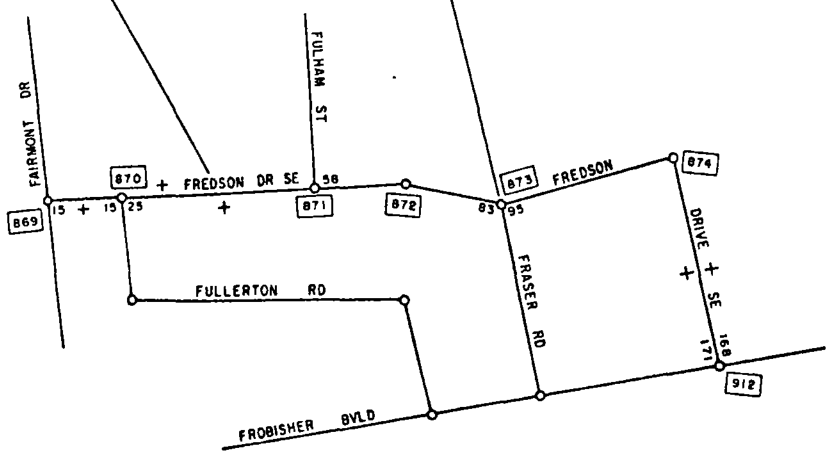
FIGURE- IV a

THE AREA MASTER FILE (SPECIAL FORMAT)

MUNIC CODE	FEATURE CODE	SEQ No	STREET NAME	TYPE	DIR	NODE NUMBER	NODE COORD X	NODE COORD Y	ADDRESS BEFORE L	ADDRESS BEFORE R	ADDRESS AFTER L	ADDRESS AFTER R	CENTROID LEFT X	CENTROID LEFT Y	CENTROID RIGHT X	CENTROID RIGHT Y	INTERSECTING FEATURES
4835	59600		FREDSON	DR	SE												
		040E				08912	706954	5651540	168	171	—	—	706868	5651755	706861	5651673	FROBISHER BV
		035				08874	706844	5651755	—	—	—	—	—	—	—	—	
		030				08873	706771	5651722	—	83	—	95	—	—	706635	5651700	FRASER RD
		025				08872	706733	5651729	—	—	—	—	—	—	—	—	
		020				08871	706631	5651722	—	—	58	—	706516	5651734	—	—	FULHAM ST
		015				08870	706497	5651711	—	15	—	25	—	—	706452	5651687	FULLERTON RD
		010B				08869	706406	5651707	—	—	—	15	—	—	—	—	FAIRMONT DR

SECTION 08, NODE 873

FEATURE CODE 59600, FREDSON DRIVE



(+) CENTROIDS FOR FREDSON DR SE

Postal Address Analysis System

Addressing conventions vary according to locality, language and post office regulations, but few comprehensive systems are available to digest and organize a file of street addresses. The PAAS system is a flexible and inexpensive device for accomplishing this job. For geocoding applications, addresses can originate from any city having an Area Master File at Statistics Canada (see Figure II, page 5). Otherwise, PAAS can re-structure and organize virtually any address file in use today.

While the number of addressing conventions across Canada is considerable and many conventions often appear in one file, PAAS consistently demonstrates a high success rate at exceptionally low cost. In its current version, it accepts street addresses (including municipality names) in completely free format and decomposes each address into several elements such as street name, street type and direction. The addresses are then matched against a subset of the Area Master File (the Address Conversion File) and, if the match is successful, a centroid coordinate is assigned to each record in the original file (see Figure IVb, page 13).

The flexibility of PAAS is enhanced through parameters which are passed to the program when geocoding starts. These parameters improve PAAS efficiency by indicating the nature and characteristics of the incoming addresses.

Significantly, the entire conversion process is accomplished at an average cost of less than one cent per address.

The Query Area Library

Many users are expected to submit special-purpose areas for data retrieval and refer to them repeatedly in making requests. Statistics Canada also expects continuing requests for census statistics arranged by the traditional standard areas — provinces, counties, census tracts and enumeration areas. (Altogether, there are 13 distinct sets of standard census areas, each set covering most of the settled area of Canada. The 13 sets comprise more than 53,000 separate areal units.)

Before information about any query area can be retrieved, GRDSR must define the query area in terms of the geocoded data base. Definition is accomplished by associating the area name with "pointers", which indicate precisely where the desired elements can be found. Pointer sets for each standard census area and for special-purpose areas are kept in a system component called the Query Area Library. A QAL is opened specifically for each new data file stored in GRDSR.

In normal practice, users outline the boundary of a special query area on a map. Vertices along the boundary are located using a digitizer so that their positions can be converted to UTM coordinates. A computer-programmed algorithm is used to test whether each successive centroid coordinate in the data base belongs in the query area. Finally, the coordinates selected are converted to pointers, which serve to locate the corresponding data elements required.

To avoid repeating this process, frequently-used area names are stored in the Query Area Library. Each area name is associated with a set of pointers. Areas that will be requested often and by different users are stored in a portion of the QAL reserved for permanent areas. Other area definitions will be stored for a limited time in the temporary QAL. Several other methods for designating query areas are described in Operations, page 18.

STATPAK

STATPAK was developed for GRDSR as a generalized program to retrieve statistics efficiently by arbitrary areas. Users communicate with STATPAK through the problem-oriented language TARELA and receive statistics in the form of convenient, easy-to-read tables.

How a file is stored

STATPAK's efficiency is made possible by changing the structure of an incoming file after it has been geocoded. Instead of keeping all data characteristics for a respondent together in one record, each data characteristic is handled separately. The entire set of responses for one data characteristic are assembled and stored as a continuous string. Because more than one record is usually attached to each centroid coordinate, an index is built to locate precisely where the responses for each distinct coordinate value are found. The index is then used to provide pointers for new query areas before they enter the QAL.

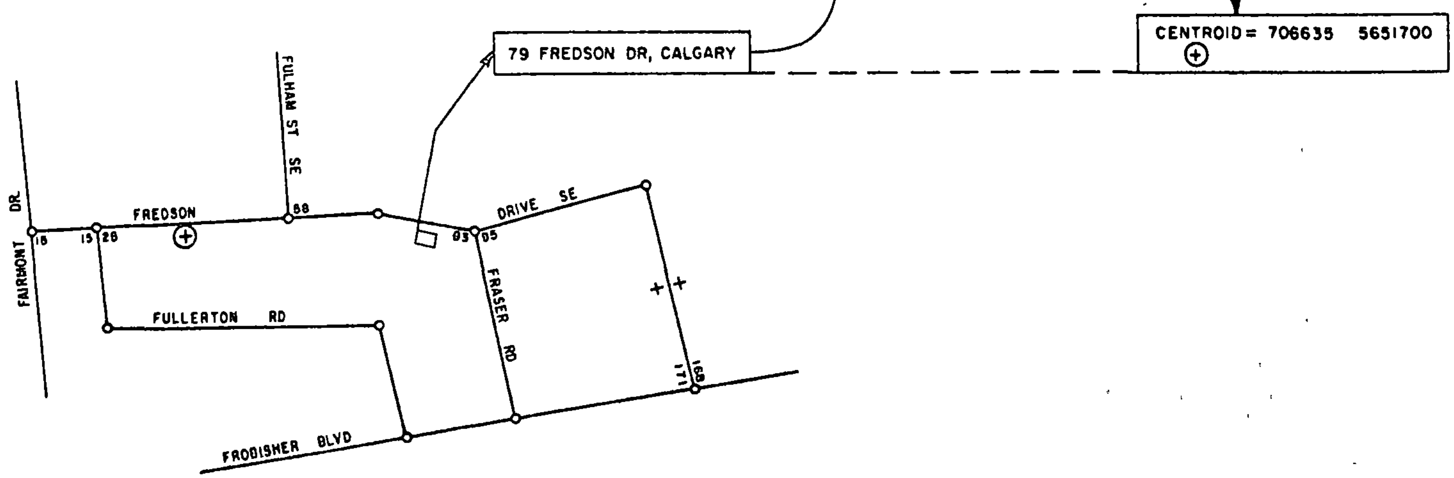
Complete strings of data characteristics are finally stored on direct-access devices and a Query Area Library is established for the file.

To visualize the final geocoded data base, imagine a huge matrix. Using the 1971 Census as an example, 21.6 million people counted in the census are arranged vertically in order of their centroid coordinates along the left side. Approximately 120 data characteristics form vertical parallel strings suspended from the top of the matrix. Instead of storing all characteristics for a person in a self-contained record, one string is created for each characteristic (such as age, sex, marital status, income or occupation). This method allows each string to be compressed to occupy the least possible space for

FIGURE - IV b

THE ADDRESS CONVERSION FILE (SPECIAL FORMAT)

STREET NAME	TYPE	DIR	FEATURE No.	ODD/ EVEN	ADDRESS RANGE LOW	ADDRESS RANGE HIGH	CENTROID ZONE	CENTROID X	CENTROID Y
FRASER									
FREDSON	DR	SE	59600	ODD	15	15	18	706452	5651687
"	"	"	"	ODD	25	83	18	706635	5651700
"	"	"	"	ODD	95	171	18	706861	5651673
FREDSON	DR	SE	59600	EVEN	58	168	18	706868	5651755
FROBISHER									



CENTROIDS (+)

the information contained. As a result, the use of costly direct-access storage space is minimized.

During the storage operation a name for each data characteristic is retained along with code names for the values the characteristic can assume. The names appear in a document called the Data Dictionary which is used, in turn, to code TARELA requests. The problem-oriented nature of TARELA rests on these names, because they are chosen by subject-matter specialists when files are submitted to GRDSR for geocoding.

How information is retrieved

After STATPAK accepts and analyzes a TARELA request, it generates a tailor-made program to retrieve the data. The program is then executed.

The operating advantage rests on large files where only a small portion is accessed at one time, that is, whenever tabulations are requested for small areas or relatively few data characteristics. Because data are retrieved in direct-access mode, the actual execution cost is strictly dependent on the extent of the query area and on the nature of the tabulations required, not on the size of the whole file. In a file of 1.5 million records, the costs of a tabulation vary from $30 to over $100, depending on the number of records to be retrieved.

Any file which has fixed or variable-length records with geographic identification (ideally UTM coordinates) may be reorganized into a form acceptable to STATPAK. Written in PL/I, STATPAK is a set of modules assembled into a tailor-made source program for each new TARELA request. The tailor-made program is exceptionally efficient because it immediately locates the required data string and accesses only those portions belonging to the user-defined query area. It is erased when the final tabulation is complete.

STATPAK is implemented at Statistics Canada on the System/360-65, under OS/MVT and HASP, and occupies roughly 150 K bytes of core storage.

TARELA

Tabulation requests are coded in a highly user-oriented language called TARELA, requests in which are submitted directly to the system and will normally be returned within one or two days, depending on the computer workload and the size of tabulations requested.

As a data retrieval language TARELA offers significant advantages It spares non-programming users the trouble of writing retrieval requests for subsequent analysis and programming, and it frees programmers and analysts for more complex work, such as refining the GRDSR System. Programming, debugging and testing delays are bypassed. Finally, potential communications problems inherent in dealing with different professional groups are avoided because the ultimate user can himself communicate directly with the data base.

To write a TARELA request, users must have access to the appropriate Data Dictionary created when their data base was geocoded. A standard data dictionary for the census files will be available to interested users. Using the dictionary, each response characteristic (i.e. age, sex, occupation) is selected by name and code words representing numerical values appearing in the data base. The user can also specify appropriate functions (such as a COUNT of persons satisfying some criteria, or SUM and AVERAGE of a set of retrieved data values). As the request is formulated, coded information is simply written after each TARELA keyword as shown in Figure VI, page 17.

Data Mapping by Computer

MAPPAK is a facility to display spatial distributions of a statistic in the form of a map. MAPPAK operates as an interface between STATPAK and SYMAP, a mapping program developed at the Laboratory for Computer Graphics, Harvard University.

Reading statistical data from a map often has compelling advantages over having the same information tabulated in report form. Inspection of the map can instantly show where extreme values of some function occur. A map can highlight problem areas at a glance.

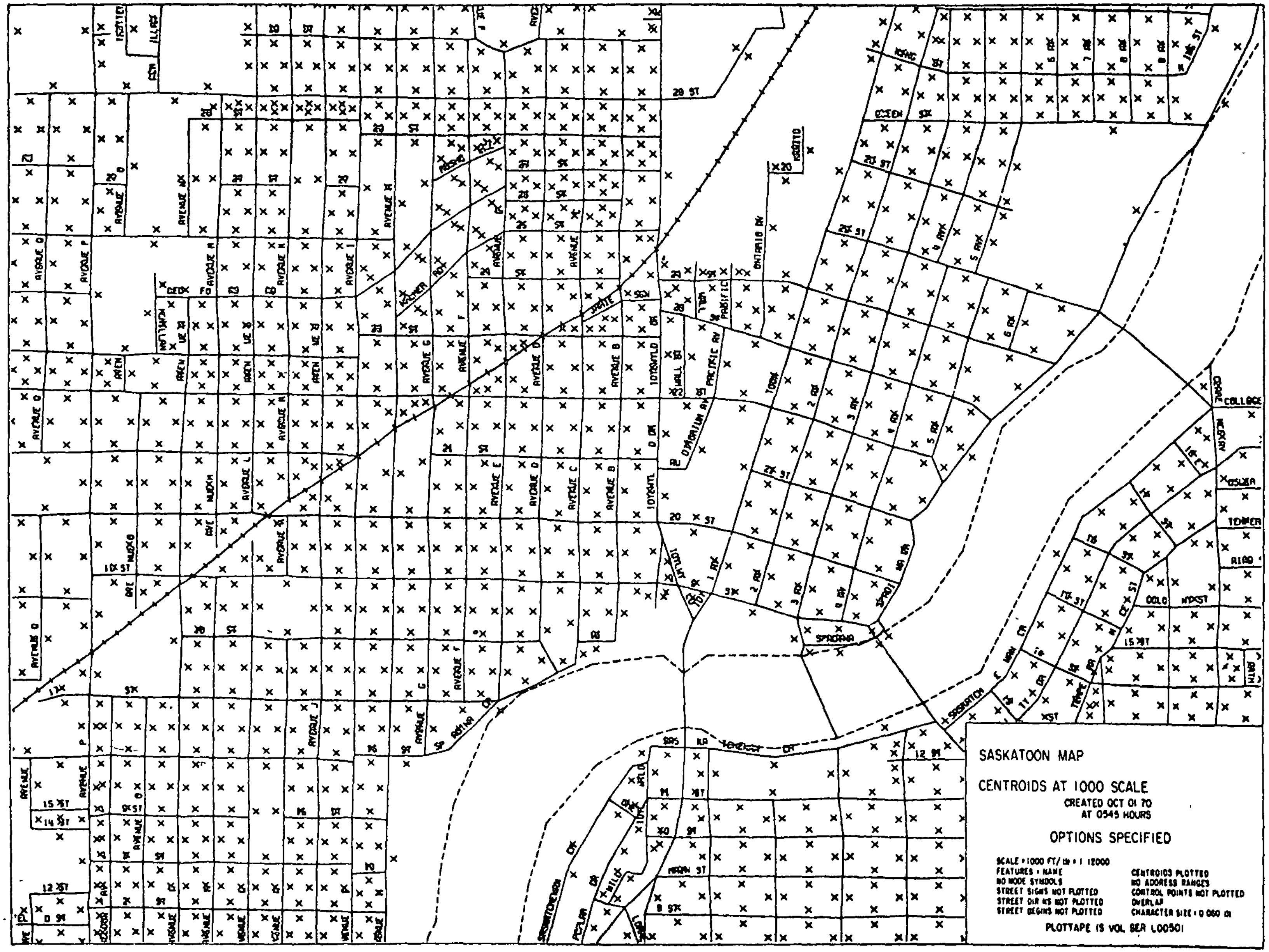
MAPPAK can be used to stratify data values into several classes or to filter a data characteristic. The results are shown as numbers or as shaded areas on the paper surface.

For instance, if a MAPPAK user is interested in census data, the distribution of average income can be depicted in many levels of shading over a city area. Or, a user can specify that an area be subdivided into 400 by 800 foot rectangular cells, with the average number of children per family shown as a number within each cell. To illustrate the filtering characteristic, MAPPAK can be requested to shade city areas where half the population is of foreign origin, or where a majority of families rent rather than own homes

The uniform data areas generated by MAPPAK can take many forms. Users can request data relative to any grid cell pattern, by rectangles of any size, or in terms of concentric circles At one extreme, a data value can be mapped for every centroid point in the city area (subject to confidentiality constraints) At the other extreme, a single data value for some arbitrary area sketched on a street map can be obtained.

FIGURE-V

COMPUTER-PLOTTED STREET MAP OF SASKATOON (PORTION)



MAPPAK incorporates all SYMAP facilities including contour mapping of surface data, classification of data values within arbitrary, pre-defined areas and summing the distribution of a set of data values. It has the flexibility to display detail down to the finest level on the data base (the block-face) and can generate maps to any desired scale (see Figure VII, page 19).

Again, it must be pointed out that the routines for confidentiality checking will be applied when MAPPAK is used to retrieve data from the geocoded census files. The routines will operate in the same manner as for regular statistical tabulations.

FIGURE-VI

PREPARING A TARELA REQUEST

RESEARCH PROJECT PLANNING ZONES

AREA SET ID--
TS01674000

AREA ID---	DESCRIPTION ---
TA01674001	PLANNING ZONE 1
TA01674002	PLANNING ZONE 2
TA01674003	PLANNING ZONE 3
TA01674004	PLANNING ZONE 4
TA01674005	PLANNING ZONE 5

TARELA DATA DICTIONARY FOR FILE
•••• GEOCODED 71 CENSUS FILE SHORT FORM
FILENAME:GE71CFSF:

PART3 : VARIABLE AND CODE DESCRIPTION
VARIABLES ON LEVEL POPULATION

VARIABLE INTERPRETATION	VARIABLE NAME	CODE OR CODE RANGE	CODE NAME	CODE INTERPRETATION	STUB TEXT
RELATIONSHIP TO HEAD OF HOUSEHOLD	RELTOHD	1	HEAD_OF_HOUSEHLD	HEAD OF HOUSE HOLD	HEAD OF HO
		2	WIFE	WIFE OF HEAD OF HOUSEHOLD	WIFE OF HE
		3	SON	SON OF HEAD OF HOUSEHOLD	SON OF HE
		4	DAUGHT_IN_LAW	DAUGHTER-IN-LAW OF HEAD OF HOUSEHOLD	DAUGH.TER-
		5	SON_IN_LAW	SON-IN-LAW OF HEAD OF HOUSEHOLD	SON-.IN-.L
		6	DAUGHTER	DAUGHTER OF HEAD OF HOUSEHOLD	DAUGH.TER
		7	GRANDSON	GRANDSON OF HEAD OF HOUSEHOLD	GRAND.SON
		8	GRANDDAUGHTER	GRANDDAUGHTER OF HEAD OF HOUSEHOLD	GRAND.DAUC
TOTAL INCOME OF PERSON	TOTINCOM	0-99999		TOTAL INCOME IN $	

```
FILE: GE71CFSF;
AREA: TA01674001;
DEFINE HEAD_INC FØR HØUSEHØLD AS TØTINCØM IF RELTØHD=1;
HEADING: NUMBER ØF HØUSEHØLDS
        AVERAGE INCØME ØF HEAD ØF HØUSEHØLD
        EACH BY NUMBER ØF RØØMS IN DWELLING AND NUMBER ØF PERSØNS
        IN HØUSEHØLD;
CHARACTERISTICS:
   PERSINHH 1,2,3,4,5 ØR MØRE;
   RØØMS 1, 2,3,4-5, 6 ØR MØRE;
TABULATE:
   CØUNT,
   AVERAGE(HEAD_INC) /#AVERAGE HEAD INCØME#/;
```

NUMBER OF HOUSEHOLDS
AVERAGE INCOME OF HEAD OF HOUSEHOLD
EACH BY NUMBER OF ROOMS IN DWELLING AND NUMBER OF PERSONS
IN HOUSEHOLD
2 FEBRUARY 1972 PAGE 1

AREA: PLANNING ZONE 1

	ROOMS 1	ROOMS 2	ROOMS 3	ROOMS 4-5	ROOMS 6 OR MORE
PERSINHH 1					
COUNT...............	15	15	40	40	30
AVERAGE HEAD INCOME.	3,555	3,035	1,810	1,515	850
PERSINHH 2					
COUNT...............	5	15	40	180	115
AVERAGE HEAD INCOME.	3,010	3,020	2,725	3,250	3,500
PERSINHH 3					
COUNT...............	-	5	15	140	75
AVERAGE HEAD INCOME.	-	3,050	3,615	4,380	3,990
PERSINHH 4					
COUNT...............	5	-	5	135	90
AVERAGE HEAD INCOME.	2,550	-	3,305	4,835	4,910
PERSINHH 5 OR MORE					
COUNT...............	-	-	10	200	285
AVERAGE HEAD INCOME.	-	-	3,190	5,350	5,150

IN THE REQUEST, THE AREA ID IS PROVIDED BY THE AREA LIST (TOP LEFT) THE FILE ID, VARIABLE NAMES AND CODE NAMES ARE PROVIDED BY THE DATA DICTIONARY SHOWN (IN PART) AT THE TOP RIGHT IN THE OUTPUT TABULATION (ABOVE) THE AREA DESCRIPTION, "PLANNING ZONE 1" HAS BEEN RETRIEVED FROM THE QUERY AREA LIBRARY

OPERATIONS

Handling User Surveys

Many geocoding applications are of interest to municipal administrations. GRDSR can be used to access information of significant importance to urban planning and administrative processes Several possible applications were described on page 4.

Geocoding and Data Storage

The geocoding operation can now be carried out in 14 larger Canadian urban centres having Area Master Files at Statistics Canada (see Figure II, page 5). Since GRDSR is fully computerized, input files can be in machine-readable form (such as punched cards or magnetic tape). Any unusual address structures in the input file may require definition prior to submission. Once the user has provided a description of the file (including record length, variable names and values, address location, etc.) geocoding can begin. This can happen in one of two ways. Users in the public sector (municipal, provincial, and federal governments) who have suitable computing facilities may obtain the GRDSR System for their own use. In other cases this operation as well as subsequent data retrievals may be carried out by Statistics Canada under contract.

In either case the actual processing phases are as follows:

- To geocode the input file, address identifiers are removed, analyzed and used to assign a centroid coordinate to each record.
- Data characteristics are gathered together and arranged in strings. The re-organized file, together with control information describing the strings, is stored on direct-access devices, ready for data retrieval.
- A Query Area Library is opened for the file.
- Finally, a Data Dictionary for preparing TARELA requests is created.

Data Retrieval

Any number of retrievals can be carried out once the storage operation is complete. Definition of query areas for the retrieval phase can also be done in several ways Initially, it will be necessary to outline the desired query areas on a city map, name them clearly, and submit these specifications along with the tabulation request. For instance, query areas for a municipal retrieval might be defined as "Planning Zones I, II, III, and IV". Mapped query area boundaries are then digitized and converted to UTM. At this point, the area definitions are stored in the Query Area Library opened specifically for this file. Subsequent tabulation requests for these areas can then be referred to the QAL for definition rather than repeating the UTM conversion operation.

From this point onwards, it is possible to obtain data tabulations through the GRDSR System. TARELA is used as a vehicle for the tabulation request, which is coded using the Data Dictionary. For instance, if the input data base was an assessment file, a tabulation of assessed value for various dwelling types could be specified using convenient characteristic names (such as "VALUE", "DWELLTYPE"), area names (such as "PLANNINGZONE4"), file names (such as "ASSMFIL6") and parameters indicating the format of tabulations desired.

Requests for further tabulations can be handled in a similar manner. Tabulation requests can be processed with exceptionally quick turnaround once files have been geocoded.

How users can specify areas

An important feature of GRDSR is that it accepts area descriptions in several convenient ways.

Outlines on maps

Users will probably find it most convenient to outline query areas on a map. In most cases any convenient map can be chosen. Statistics Canada is producing copies of computer-plotted city maps, which are particularly appropriate for graphically displaying block-faces in each city area. Outside the urban areas of block-face geocoding coverage, users will be advised to use the National Topographic Series (NTS) maps produced by the Department of Energy, Mines and Resources. The important thing is that users choose an appropriate map scale, then mark out and name query areas as clearly and as accurately as the problem demands.

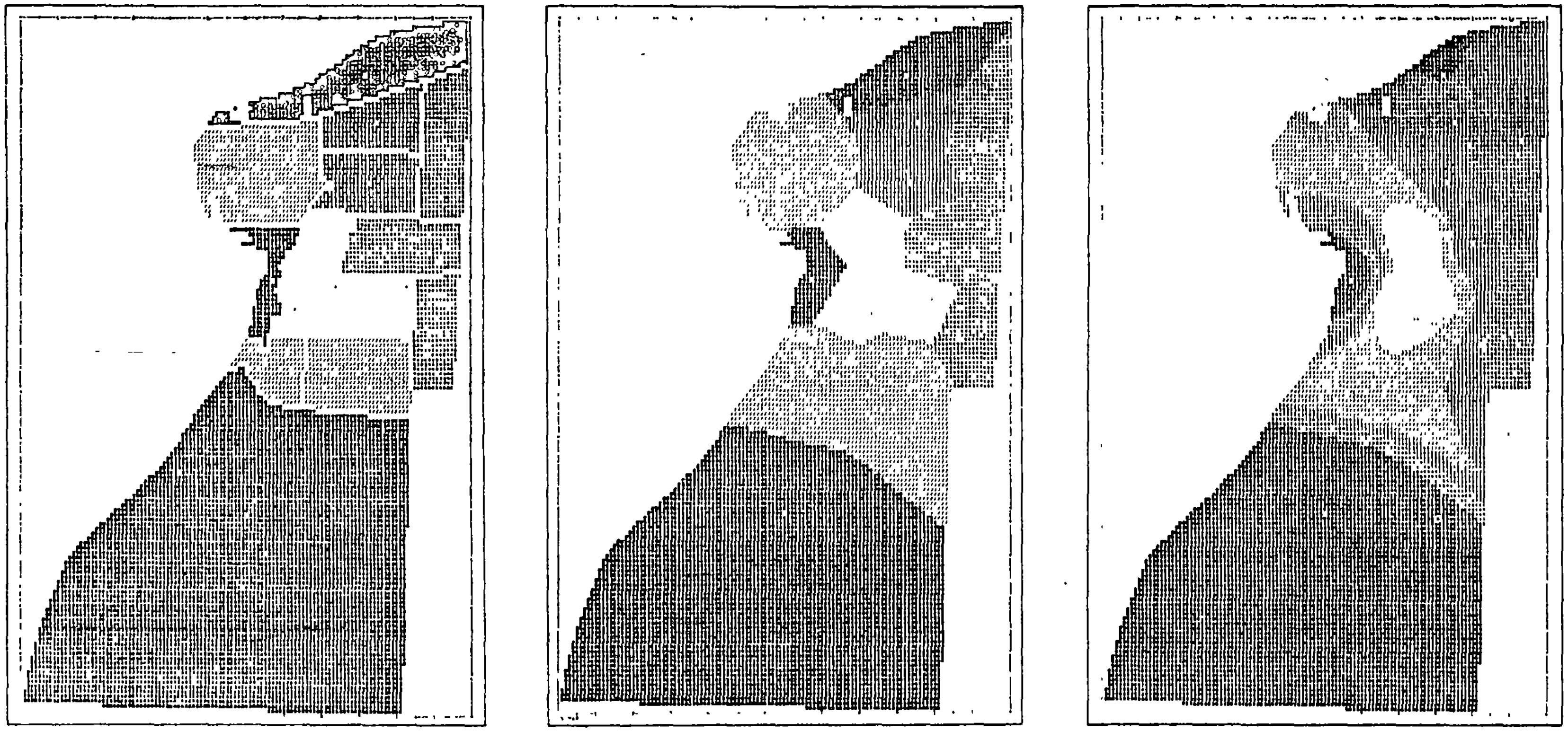
Defined by features

It will be possible to specify a query area in terms of known features (rivers, streets, railroad tracks). For instance, an Ottawa user could describe in writing that Research District No. 5 consists of an area bounded on the north by St. Patrick Street, on the east by Chapel Street, on the south by Templeton Street and on the west by Nelson Street. It is possible to request data for street-oriented query areas in the same manner. For example, a user could request statistics for one side of Rideau Street, in the municipality of Ottawa; from Sussex Drive to King Edward Avenue (odd-numbered side, even-numbered side, or both).

Rather than submit a list of feature names, users can define areas by a set of node numbers copied from the computer-plotted city maps. Nodes are chosen at

FIGURE VII

MALE/FEMALE RATIO FOR THE CITY OF SARNIA



(1) CONFORMAL (2) PROXIMAL (3) CONTOUR

MAPS ARE ALIGNED IN THE NORTH-SOUTH DIRECTION BOUNDARIES FOR THE CONFORMAL MAP ARE FORMED BY CENSUS TRACTS. FOR THE OTHER TWO MAPS, BOUNDARIES ARE DETERMINED FROM THE DATA RATIO VALUES ARE DIVIDED INTO FIVE CLASS INTERVALS BETWEEN EXTREME VALUES OF 0 89 AND 1 18 EACH HIGHER CLASS INTERVAL IS REPRESENTED BY A PROGRESSIVELY DARKER SHADE points where the boundary features intersect. Thus, the area perimeter is defined by the nodes, which are matched to the Area Master File before storing the area in the QAL.

Using grid coordinates

UTM coordinates can be used to specify query areas in two ways. Data can be retrieved according to a list of individual centroids chosen from the Area Master File. Or, a set of coordinates along a boundary can be used by the system to calculate an enclosed area.

Using area names

Once an area has been submitted using one of the above methods, its name and description are entered and stored, temporarily, in the Query Area Library. For subsequent references the QAL description will be referenced directly by area name, bypassing the map conversion operation.

Of course, all requests for census data by traditional standard areas will also be serviced through the Query Area Library. The QAL contains a pointer set for each province, county, municipality, census tract, enumeration area, and all other standard geostatistical areas used in the 1971 census.

Using other areas

The system permits addition and subtraction of query areas to form a new query area. For instance, a user can outline and request statistics for six areas on a map, naming these areas Area 1, Area 2, ... Area 6. He can then request further data for a new zone, defined as follows:

QZONE1 = Area 1 + Area 2 + Area 3

If Area 5 is contained within Area 6, the following specification would result in a doughnut-shaped query zone:

QZONE2 = Area 6 - Area 5

FURTHER INFORMATION

Users who are primarily interested in census statistics using GRDSR may obtain further information by contacting:

User Inquiry Service

Census Division

Statistics Canada

Ottawa, K1A 0T6

Statistics Canada is prepared to provide assistance and further information to users who wish to geocode their own data files. Detailed system documentation will be available in response to technical requests. This information will be provided by a manual entitled *A Technical Description of the GRDSR System*, followed by User Manuals for certain components. For further information of a specialized or technical nature, please contact:

General Survey Systems
Methodology and Systems Branch
Statistics Canada
Ottawa, K1A 0T6

Comput & Graphics, Vol 2, pp 155-161 Pergamon Press, 1977 Printed in Great Britain

THE USE OF GRAPHICS IN A COMPUTER-BASED URBAN INFORMATION SYSTEM

R. S. BAXTER
Applied Research of Cambridge Ltd, 4 Jesus Lane, Cambridge, U.K.

(Received 22 October 1976)

Abstract—This paper describes the development of an interactive graphics system for the display of planning information in urban areas It outlines the objectives of the system and explains the content and structure of the various data files It goes on to explain the user operating environment and the provision of suitable graphics facilities By way of examples it demonstrates the potentialities of the system

INTRODUCTION

As part of a research project to explore the effects of different urban forms in British new towns[1, 2], a team in the Department of Architecture at Cambridge University developed a computer-based information system. An account of the preparation of the data bank and its contents is documented in two papers produced by the group[3, 4]. A visual comparison of the land uses and road networks[5] acts as a graphic record of the contents of the data bank.

On the completion of the research project, the author took the opportunity to re-structure the data bank and to develop a computer-based system for the manipulation and output of the data This paper describes the outcome, and to allow the reader some degree of organization of ideas, is structured in the following manner. Firstly, a brief description is given of the data files for the land uses and road network of each town (see Appendix 1 for a list of the towns) This is followed by an explanation of the command language which allows the user to communicate with the urban information system The next section identifies a need for computer graphics as an output style and discusses the software language adopted The paper concludes with some applications which indicate a likely operating environment and detailed program requirements based upon user needs

DATA FILES

For each of the towns there are two data files The first of these contains information relating to the land use and the second contains information describing the road network.

The fundamental spatial unit in the land-use file is the parcel. Within all towns each parcel is uniquely numbered on the map The spatial location of the parcel is given, in the case of small parcels, as the coordinate point reference of the parcel's visual centroid and, in the case of large parcels, as a string of boundary coordinates defining the parcel's perimeter. Once the town map is completely marked up, the parcel number together with the spatial location is recorded on a digitiser The digitiser output file for each town is subsequently processed by the computer and all coordinate points are transformed to a user-defined grid (National Grid, say)

A second file contains land-use characteristics for each numbered parcel. For each land-use entry associated with the parcel (and parcels can have mixed land uses and thus more than one entry) there are four data items The first of these gives the development stage (year or proposal phase) of the specific land use on the parcel. The second item is a three-digit hierarchical land-use code. The most significant digit, for example, denotes broad categories of offices (0), industry (1), storage (2), shops (3), residences (4), education (5), public buildings (6), open space (7), transportation uses (8) and health services (9). The third item is an attribute code which relates in some way to the land use For example, attribute codes 13 and 14 refer to the number of school places and forms of entry to the school and must, necessarily, be related to land-use codes 501–514 which include various categories of nursery, primary and secondary schools. The fourth data item gives the specific parcel value for the previous attribute code. Hence a typical record might be 5 511 13 250 which would indicate a phase 5 development on the parcel of an independent primary school (land-use code 511) with 250 pupils (attribute code 13)

The parcel characteristics file is merged with the parcel location file by matching the parcel number The single land-use file which results for each town has a structure as illustrated in Fig 1.

The fundamental spatial unit in the network file is the node Within all towns each node is uniquely numbered on the map. Once the map is comprehensively marked up, the node numbers with their point locations are recorded on a digitiser and the output file is subsequently processed by computer and all coordinate points are transformed to the user-defined grid

A second file contains information about the nodes and links of the network For each numbered node a code indicates its type· level junction (1), roundabout (2), multi-level exchange (3), cut-off (4), public transport stop (5) and cell centroid (6). The connectivity of the network is defined by a string of node numbers to which the origin node is connected and following each destination node number is an attribute code indicating the link type secondary (1), secondary doubling as public transport (2), connecting to cell centroid (3), primary (4), primary doubling as public transport (5), public transport only (6). A typical record might be 7 2 8 1 10 4 which would indicate a roundabout node numbered 7 which is connected to node 8 by a secondary road link and to node 10 by a primary road link

The node and link characteristics file is merged with

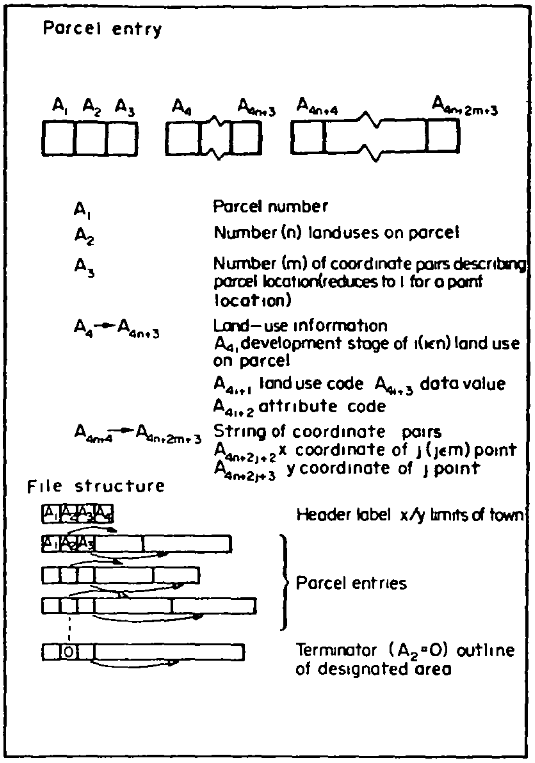


Fig 1 Structure of land-use files

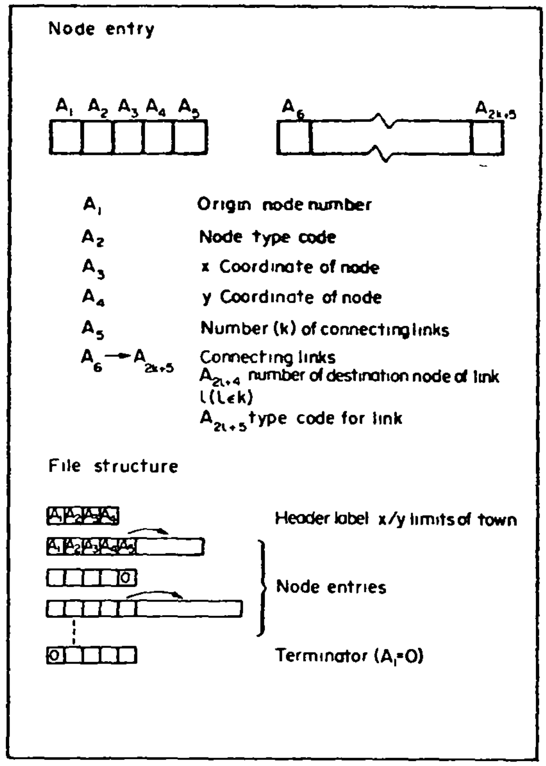


Fig 2 Structure of network files.

the node location file by matching the node number The single road network file which results has a structure as illustrated in Fig. 2.

COMMAND LANGUAGE

One of the problems attendent with the development of computer-based information systems is that access to the information must necessarily involve the computer. Since the majority of users will not be computer specialists (or wish so to be) there is an evident need to provide a means whereby the user can interrogate the information system with the minimum of effort. With multi-access on-line systems this can be achieved through a command language. Maximum transparency is provided by reserving a few commands which allow the user to invoke operations of varying levels of sophistication.

For the particular application being described in this paper, three simple commands suffice: LOGON, LOGOFF, ANALYSE The first two are system commands which allow the user to wake up the machine from a terminal to gain access to computing resources and to terminate a session of on-line working The third is a macrocommand (in turn calling system commands) which allows the user to interrogate, manipulate and output urban information held in the files previously described.

The command ANALYSE takes four parameters separated by spaces which are substituted in the appropriate slots in the macro·

Parameter 1 Town code (see Appendix 1);

Parameter 2: LAND or ROAD depending upon the type of data to be interrogated;

Parameter 3: Output device name (plotter, Tektronix, etc);

Parameter 4: Integer denoting the type of analysis required and hence the program module to be loaded (see later).

Invoking this command will open the required data file from the magnetic disc (parameters 1 and 2), reserve an output channel to the named device (parameter 3) and load and execute the numbered program module (parameter 4).

On completion of the command the user is free to repeat the command with different parameters settings or to LOGOFF.

The files are held on magnetic disc with access provided through a directory. The entries in the directory have three components the first of which is the file-owner name (BAXTER, say) The second component is the town code (see Appendix 1) and the third component is either LAND or ROAD depending upon the information type of the file For example, a user might request dataset BAXTER 15CU LAND in which case the land-use file for Cumbernauld would be opened

Whilst the data originates as card images, the internal computer structure comprises a sequence of binary records Because of the variable length of records, each parcel or node entry is held as two logical records with the A_1 to A_3 (land-use file) or A_1 to A_5 (network file) constituting the first record and thus providing the necessary information (attributes A_2 and A_3 in the land-use files, attribute A_5 in the network file) with which to read the second entry correctly. No attempt has been made to provide direct access pointers since most applications require a comprehensive scan of the entire file.

GRAPHICS

Output from any information system has to be communicated to the user. The nature of the output dictates, to a large degree, the method of presentation In urban information systems there are two common types of enquiry. Firstly, there are those queries which produce a single simple answer This approach is typified by interactive-communication languages in which the user is encouraged to type a request on a terminal and the system responds with a typed answer to the question based upon the interrogation of one or more data files The second type of enquiry involves the aggregation, manipulation and presentation of large data sets. Traditionally this approach was undertaken in a batch mode environment with voluminous lineprinter output as the end product. The type of application being discussed in this paper falls into the second category and attention is now given to rectifying constraints in its user interface.

Arguably, the user concerned with matters of urban governance requires answers from any information system at a speed which cannot be satisfied by a batch processing environment (even given a sympathetic scheduler). The first requirement, therefore, is for a method of interactive enquiry in line with that adopted for interactive-communication languages. Solutions to this problem lie with the command language already described. The second requirement is to provide a more satisfactory method of output presentation. In this context computer graphics can play a decisive role in two ways. One of the intrinsic characteristics of the land-use and network data files is the spatial attributes associated with the basic entities. The retention of space in the output suggests the use of suitable mapping facilities. Secondly, the aggregation and manipulation of spatial attributes produces data structures which might be better presented in diagramatic form rather than as tabulations. In both cases there is an evident need for suitable graphics software

A comprehensive software system for graphics has to be conceived with three priorities in mind. Firstly, the programs written in the system should be transportable

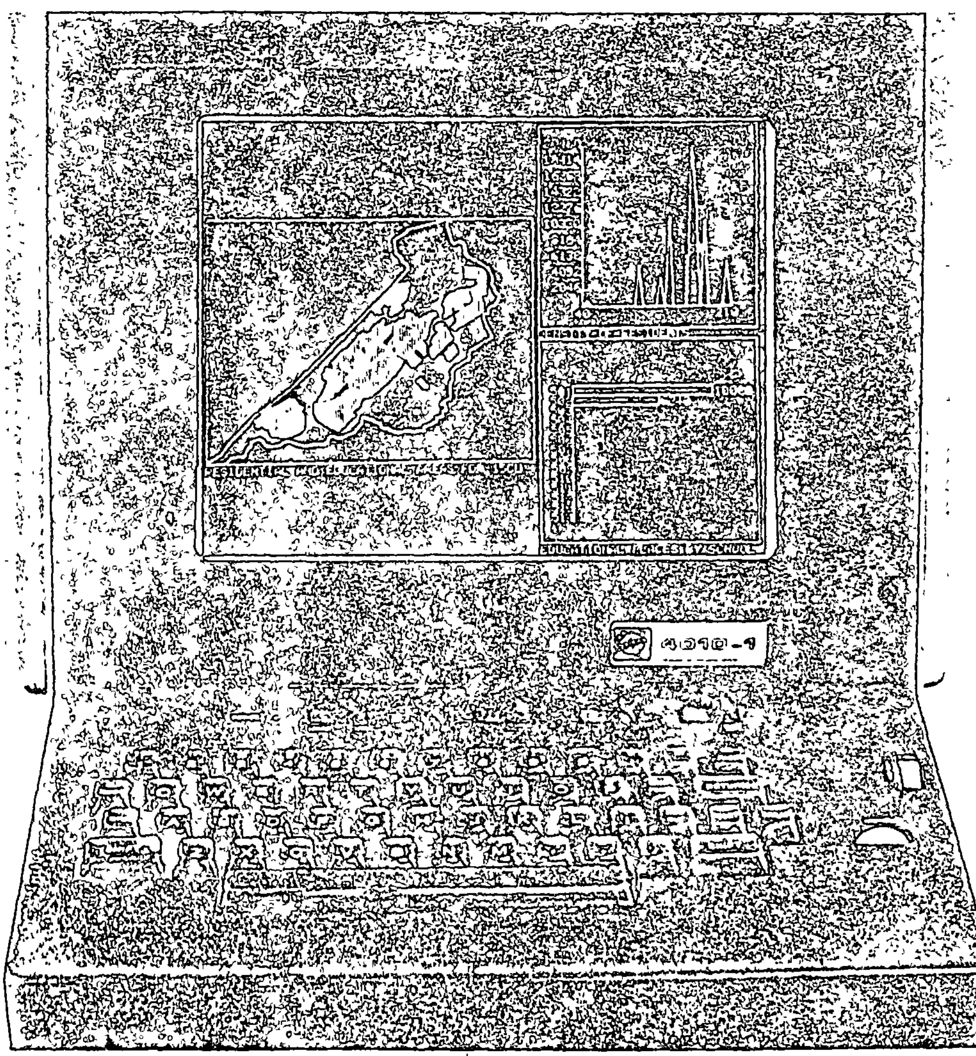


Fig. 3. Output of land uses for Cumbernauld on Tektronix tube

between computer installations. Secondly, the graphical output generated within the program should be device independent. Thirdly, the system should allow the user to generate two- or three-dimensional objects and then manipulate and interpret them prior to display. GINO (Graphical Input and Output) is such a language[6, 7]. It was designed and implemented by the University of Cambridge Computer Laboratory and has since been distributed widely as GINO-F by the Computer Aided Design Centre in Cambridge.

GINO satisfies the stated requirement of any graphics language by allowing the user to access the system from a high level programming language and by providing a series of initialization routines for each graphical device which can be called from the user's program. The third priority (to generate, manipulate and interpret objects) is possible because GINO takes the form of a library of sub-routines accessible from the user's program.

In the GINO system the unit in terms of which pictures are generated is the picture segment—the smallest unit that may be independently displayed or plotted. Picture segments are generated sequentially by the user who controls the size and content of the segment. Any number of segments can be viewed together. A picture segment consists of a sequence of picture parts, each added by a call to a GINO routine. Routines are available for a group of built-in picture parts which include points, lines, characters and circular arcs. This is supplemented by a facility to create user-defined picture parts and the object thus created can be called as many times as required. Picture segments can be generated in two or three dimensions, the latter being displayed as two-dimensional projections. Windowing or clipping routines exist for both two and three dimensions, enabling the suppression of picture parts falling outside a specified region so that selected parts of a large picture can be viewed separately. Transformations can be applied to all or any subset of picture parts making up a segment. These take the form of translation, rotation, scaling, shearing, reflection, parallel or point projection. This means that the user is able to define his picture parts to any scale or orientation that is convenient and

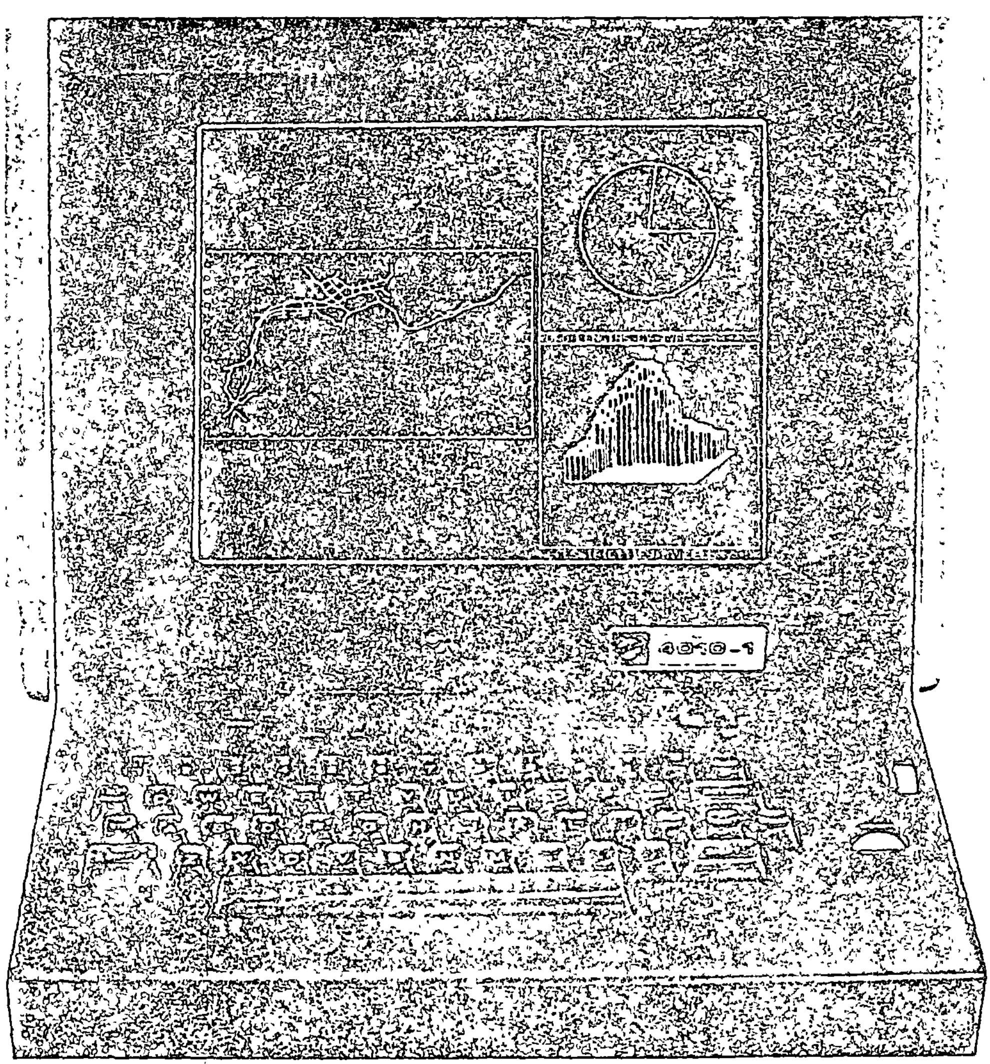


Fig 4 Output of road network for Newtown of Wales on a Tektronix tube.

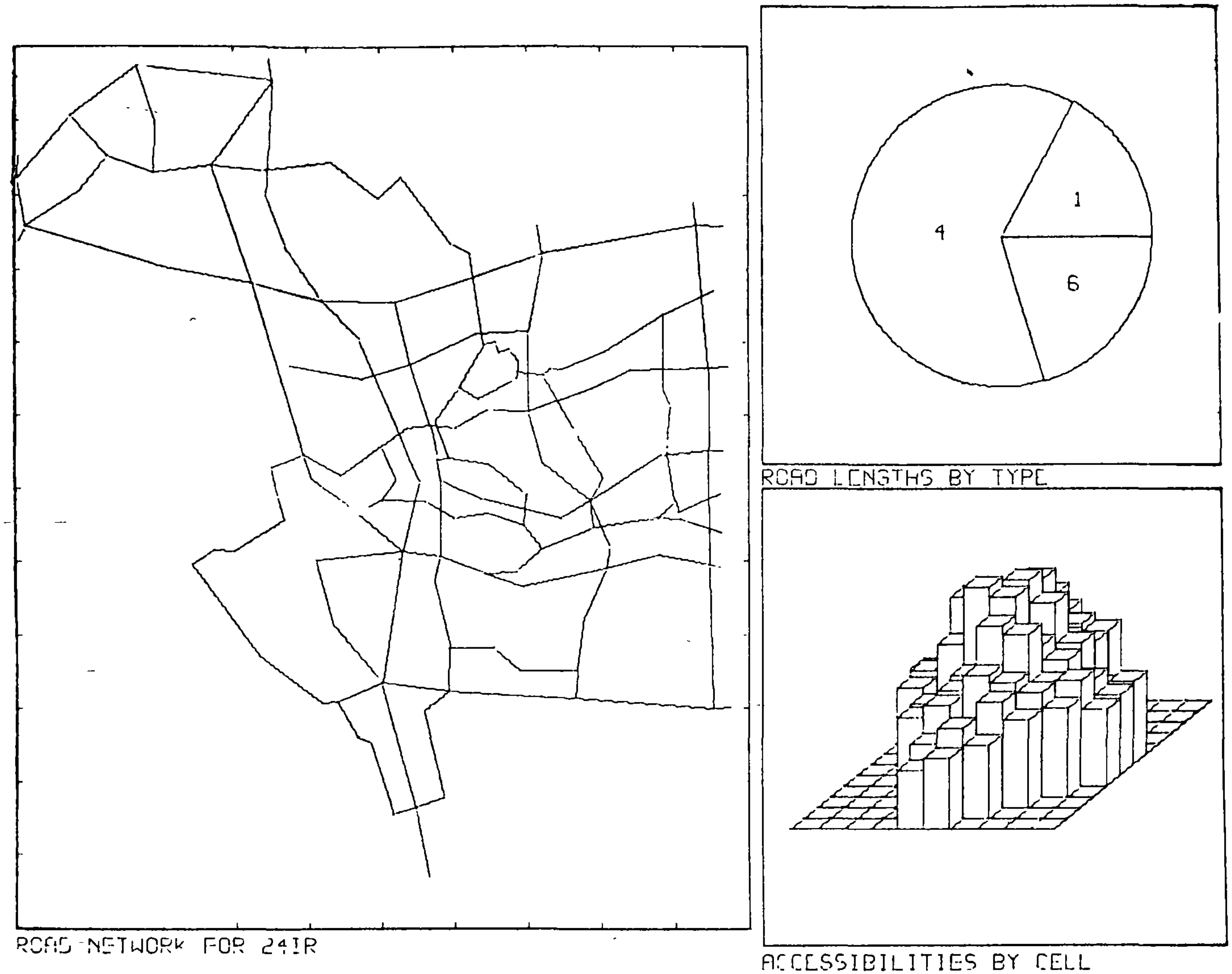


Fig 5 Output of road network for Irvine using a drum plotter

then use transformation to position the resulting display as required.

APPLICATIONS

The system as described was implemented on the Atlas computer in Cambridge. It is a large multi-access computer to which are attached a selection of on-line teletype terminals and graphical devices as well as the normal facilities such as card, paper tape lineprinter and plotter peripherals with backing store provided by magnetic tape and disc units. The multi-access environment operates under the Cambridge command system. The graphical devices available to the user (and designated through parameter 3 of the ANALYSE command) range from small satellite computers with cathode ray tube displays, storage tube displays such as the Tektronix, off-line drum and flatbed plotters, Complot on-line plotter and the HRD-1 laser plotter.

An integral component of the information manipulation system is a suite of programs which read the opened data files (parameters 1 and 2 of the ANALYSE command) and process the data according to user requirements This suite of programs is not static but evolves to match user needs. A particular program is invoked through parameter 4 of the ANALYSE command after consulting a dictionary of available programs and their reference number. Each of the load modules in the library has a source program written in ANSI FORTRAN with calls to the GINO library for all graphical output. Let us see how the system operates.

Assuming that a user is on-line via a Tektronix storage tube to the computer through invoking the LOGON command, then by typing

ANALYSE 15CU LAND TEKTRONIX 1

the land file for Cumbernauld will be processed by program 1 and the results sent to the users console. Program 1 retrieves all residential and educational land-use parcels from the file and maps their distribution within the designated area of town From the output in Fig 3 it can be seen that all educational areas are horizontally hatched and residential areas are vertically hatched. The scale of the map is denoted by the kilometre square pecks along the borders. The residential densities for each of the land-use parcels are calculated and presented in graph form (density against frequency of occupation) together with the calculated mean residential density for comparative purposes Also the school spaces are broken down into categories (nursery, primary, secondary, etc). and shown in the form of bar diagrams.

If the user now types

ANALYSE 28NT ROAD TEKTRONIX 2

then the road network file for the Newtown of Wales is opened and processed by program 2 This displays the road network on the Tektronix display together with certain analyses as shown in Fig. 4. The first of these analyses presents a breakdown of link types in the form of a pie diagram divided according to the total road length in each type. The second diagram is a projection of a three-dimensional histogram with hidden lines suppressed representing the accessibility surface for the town. This surface is generated by summing the minimum road distance (by applying a minimum path algorithm to the network) from each kilometre cell to all the other cells connected to the network.

If the user were to type

ANALYSE 24IR ROAD PLOTTER 2

the same analysis would be performed on the Irvine data file but with the output going to a plotter as in Fig. 5.

This, hopefully, has served to illustrate the environment available to users who wish to examine urban information about the towns contained in the data base. Each of the requests described has a response time of about 90 s between typing the request and obtaining a picture on the Tektronix screen. This is sufficiently fast to satisfy the majority of users. Furthermore, through the use of a general purpose graphics language the development of new applications programs involves the minimum of effort. For example, in the case of the two programs described, the combined length of FORTRAN code is only 411 executable statements.

REFERENCES

1. P Cooper, W Lindsay & E Taylor, *New Towns· Analysis of Land Uses* LUBFS, Cambridge (1974)
2. P Cooper, W Lindsay & E Taylor, *New Towns Analysis of Activities and their Densities* LUBFS, Cambridge (1974)
3. R Cheesman, W Lindsay & M Porzecanski, *New Towns The Data Bank, its Construction and Organisation* LUBFS, Cambridge (1971).
4. E Taylor, P Cooper & W Lindsay, *New Towns Development of the Data Bank* LUBFS, Cambridge (1974)
5. W Lindsay, R Cheesman & M. Porzecanski, *New Towns. A Comparative Atlas* LUBFS, Cambridge (1971)
6. P A Woodsford, The design and implementation of the GINO 3D graphics software package *Software—Practice and Experience* 1, 335–365 (1971)
7. R S. Baxter & B B Shah, GINO—a general purpose graphics system for cartography *Cartographic J* 10, 99–104 (1973)
8. D F Hartley, *The Cambridge Multiple-Access System* University Mathematical Laboratory Cambridge (1968)

APPENDIX 1

Town codes

01ST Stevenage
02CR Crawley
03HH Hemel Hempstead
04HO Harlow
05AY Aycliffe
06EK East Kilbride
07PE Peterlee
08HA Hatfield
09WE Welwyn
10GL Glenrothes
11BA Basildon
12BR Bracknell
13CW Cwmbran
14CO Corby
15CU Cumbernauld

16CU Cumbernauld extension
17SK Skelmersdale
18LI Livingstone
19TE Telford
20RE Redditch
21RU Runcorn
22WA Washington
24IR Irvine
25MK Milton Keynes
26PB Peterborough
27NT Newtown of Mid-Wales 2nd. Proposal
28NT Newtown of Mid-Wales 1st. Proposal
29NO Northampton
30WR Warrington
31CG Craigaron
32AN Antrim
33BM Ballymena
34HK Hook.