Centro de Educación ontinua
división de estudios superiores
de la facultad de ingeniería unam
Tacuba 5 Primer Piso México 1, D. F. Tels. 521-40-20

1.a/78. MODELOS DE SIMULACION PARA LA PLANEACION DEL USO DEL SUELO Y DEL TRANSPORTE.
II

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

MODELOS DE SIMULACION PARA LA PLANEACION INTREGRAL DEL USO DEL SUELO Y EL TRANSPORTE.

TEMA: VARIOS.

PROF. ING. STEPHEN H. PUTMAN.
PROF. ING. FEDERICK W. DUCCA.

ENERO, 1978.

Palacio de Minería Calle de Tacuba 5, primer piso. México 1, D. F.

DRAM - Disaggregated Residential Allocation
Model: Forecasting Procedure

Program documentation for users

December, 1977

Urban Simulation Laboratory
Department of City and Regional Planning
University of Pennsylvania
Philadelphia, PA 19104

This work was supported by U.S. Department of Transportation, Urban Mass Transit Administration, via National Science Foundation Grant APR73-07840-A04.

SUMMARY

DRAM forecasts the spatial location of employed residents based on place of work, zonal land use patterns, zonal income distributions, and the zone-to-zone travel impedances. In addition, land use patterns are updated, and the work to home, home to shop and work to shop trip matrices are estimated.

DRAM is meant to be used with a trip assignment model as part of an iterative procedure to forecast future locations of urban activities and traffic patterns. It can, however, be used without this interconnection.

Input Requirements - DRAM requires a basic employment forecast, the base year distributions of land use by zone and income distribution by zone, and an impedance matrix for the forecast period. The basic employment forecasts must either be obtained exogenously or, if possible, forecast using EMPAL. Data on land use and income distributions must be obtained from published or survey data for the base year, but later runs of DRAM in a forecast series can use output from previous DRAM runs. The impedance matrix can be obtained either from observed data or a trip assignment model for the base year, but future impedance matrices must be obtained from the trip assignment package.

Control totals must be obtained exogenously. Parameters for the residential allocation procdure are normally obtained from the gradient search procedure CALIB.

Output - DRAM outputs two files to mass storage and prints several reports. The files on mass storage are land use, income distribution and service employment, and the trip matrices.

Reports printed are the input data, land consumption coefficients, the intermediate calculations in the household allocation procedure, final population and land use measures, and a summary of trips originating and arriving in each zone.

Operation - The operation of DRAM consists of 3 parts; data input, residential allocation, and land consumption and trip generation.

Data input is done in the main routine and an input subroutine. Base year land use and income data, and a basic employment forecast are read, and land consumption factors are calculated. The household allocation procedure is a three step process. First an initial allocation is made using the base year data. With this initial allocation, a forecast of retail employment location is made. Then a second household allocation is made using the new retail employment location. Land consumption and trip generation are performed last. Land is consumed based on the allocation of residents and previous density. Commercial land is consumed based on commercial employment and previous commercial density. Trips are generated using probability matrices from the residential and service employment allocations.

ACKNOWLEDGEMENTS

The DRAM model was specified by Stephen H. Putman and programmed by Frederick W. Ducca. Both of the Urban Simulation Laboratory, Department of City Planning, University of Pennsylvania.

REPORTS

What is shown below is the title of each report, headers and the first line of output.

1. Size and parameter data

```
100 ZONES
30 INTERVALS IN PROBABILITY FUNCTION 3 MINUTE INTERVAL WIDTH
PRINT OPTION 1

TITLE
RETAIL TRADE COEFFICIENTS
10.0
 5.0

                              PARAMETERS

STRATUM          LAND USE AND IMPEDANCE              INCOME DISTRIBUTION

   1        1.05  -.45  .77  1.24  0.0            2.22  1.11  -.6  -1.71
   .
   .

   .
   .
```

The parameters are in order:

1,2 - impedance parameters.

3 - residential land area.

4 - percentage of developable land which has been developed.

5 - vacant land.

6 - attraction to income group 1

7 - attraction to income group 2

8 - attraction to income group 3

9 - attraction to income group 4

```
          CONVERSION MATRIX

0.319    0.297    0.126    0.258

0.159    0.192    0.168    0.481

0.133    0.133    0.190    0.544

0.318    0.295    0.125    0.262
```

note: This matrix converts employment by type to employment by income group. The entry 0.297 indicates that 29.7 percent of the employees in employment group 1 are in income group 2.

2. Base year data

BASE YEAR--HOUSEHOLD DISTRIBUTION

HOUSEHOLDS 1 2 3 4	TOTAL HOUSEHOLDS	RESIDENT POPULATION	GROUP QTRS RESIDENTS	EMPLOYED POPULATION	NONWORKING POPULATION	TOTAL POPULATION
873 344 134 78	1429	7500	979	1429	6071	8479

BASE YEAR--LAND USE ACTIVITY

TOTAL LAND	UNUSEABLE LAND	STREETS, HIGHWAYS	BASIC	RETAIL	RESIDENTIAL	VACANT AVAILABLE	TOTAL AVAILABLE
1 705	83	205	250	108	56	3	417

IMPEDANCE MATRIX, UP TO FIRST 20 ENTRIES

1	66	272	250	249	265	193	257	152	92
2	171	343	266	251	177	123	205	116	181

BASE YEAR-- EMPLOYMENT DISTRIBUTION

BASIC EMPLOYMENT			TOTAL	RETAIL EMPLOYMENT	TOTAL EMPLOYMENT
1	2	3			
10	21456	7771	29237	10988	40225

LAND CONSUMPTION

RETAIL	RESIDENTIAL
.006	.003

note: .006 indicates that for each retail employee, .006 units of land are required.

3. Forecast year data

FORECAST YEAR--HOUSEHOLD DISTRIBUTION

HOUSEHOLDS 1	2	3	4	TOTAL HOUSEHOLDS	RESIDENT POPULATION	GROUP QTRS RESIDENTS	EMPLOYED POPULATION	NON-WORKING POPULATION	(con
57245	45429	28820	23102	157792	620580	979	154597	465982	

TOTAL POPULATION

621559

FORECAST YEAR--LAND USE ACTIVITY

TOTAL LAND	UNUSEABLE LAND	STREETS, HIGHWAYS	BASIC LAND	RETAIL LAND	RESIDENTIAL	VACANT AVAILABLE	TOTAL AVAILABLE
705	83	205	49	267	101	0	417

Note: Total available land is the sum of basic land, retail land, residential land and vacant available land.

FORECAST YEAR--EMPLOYMENT DISTRIBUTION

BASIC EMPLOYMENT 1	2	3	TOTAL	RETAIL EMPLOYMENT	TOTAL EMPLOYMENT
0.	21456	7771	29227	240525	269752

4. Trip distribution

TRIP GENERATION

WORK TO HOME	WORK TO SHOP	HOME TO SHOP
389785	155228	307241

TRIPS FROM ORIGIN ZONES TO FIRST 9 DISTRIBUTION ZONES

1	53314.	20374.	10554.	5185.	6191.	12170.	7268.	11952.	41342.

	TRIP ORIGINS			TRIP DESTINATIONS		
	WORK-HOME	WORK-SHOP	HOME-SHOP	WORK-HOME	WORK-SHOP	HOME-SHOP
1	91196	36318	41197	42348	36445	74076

TOTAL TRIPS 852251

In addition to the reports above, there are 2 optional reports which may be printed if parameter IPRNT=1. They appear between the Land consumption report and the Forecast year - household distribution report. They are:

```
                    INTERMEDIATE ALLOCATION OF RESIDENTS
 1                  5437        39674       23633       19548


   PROBABILITY OF TRAVEL FOR 3 MIN. TIME INTERVALS PROBL-BETA = 5.0
       .1997        .2598        .1471        .0904        .0606
```

MESSAGES FROM DRAM

DRAM prints only two messages, both in the main program. These messages concern the amount of the array area requested compared to the amount available. They are in the following form:

```
Main Program Array M Allocated XXXX words, uses YYYY
Main Program Array NU Allocated VVVV words, uses WWWW
```

If either YYYY is greater than XXXX or WWWW greater than VVVV the program will stop. Otherwise the messages are for information only.

FILE TABLE

The files below are used for intermediate data storage during execution.

File name	DD Name	Contents
ITLUPIV. WKHOME	FT02F001	Probability matrix of work to home trips
ITLUPIV. HMSHOP	FT08F001	Probability matrix of home to shop trips
ITLUPIV. WKSHOP	FT09F001	Probability matrix of work to shop trips

The files below are input files for DRAM. Some of the files described may be in card input for the sample problem but should be on mass storage devices for applications.

File Name	DD Name	Function
Optional	FT04F001	Basic employment and basic land use forecasts.
Optional	FT11F001	Base year land use and income distribution data.
Optional	FT12F001	Impedance matrix

The following are output files:

File Name	DD Name	Function
Optional	FT13F001	Matrix of origin to destination trips, to be used in trip loading procedure.
Optional	FT14F001	Forecast year land use and income distribution data. To be used in next iteration of projection package.

KEYWORD TABLE

Dimension Parameters

Keyword	Type	Value or Purpose
N	I	Number of zones in area.
NINT	I	Number of intervals in retail allocation function.
INT	I	Interval width, retail allocation function.
NG	I	Number of income groups.
NK	I	Number of employment groups.

Controls

Keyword	Type	Purpose
EAMPRO	R	Forecast volume of retail employment.
PROPOP	R	Projected population total.
WH	R	Total work to home trips.
WS	R	Total work to shop trips.
HS	R	Total home to shop trips.

Mathematical parameters

Keyword	Type	Purpose
CNV	R	Factors to convert employment by type to employment by income group.
PLT	R	Land use and impedance parameters of residential allocation function.
PID	R	Income distribution parameters of retail allocation function.
B(1)	R	Impedance factor, work to shop trips
B(2)	R	Impedance factor, home to shop trips

Other

Keyword	Type	Purpose
TITL	A	Run title
PPI	R	Factor to scale employment, used in policy testing.
IPRNT	I	Print option. 0 - Omit intermediate reports I - Print intermediate reports

---

CORE REQUIREMENTS

Core required for an execution of DRAM is a function of program size and array size. The array size in turn depends on the number of zones in the areal system. The formula for the required array size is $\frac{N^2}{2}$ + 60N+ 200 where N is the number of zones. The program length is 50796 bytes of core. Thus, if there are 100 zones in the areal system, the core requirement is $(\frac{100^2}{2}$ + 60 (100) + 200) X (4) + 50796. This is equal to 95596 bytes. In applications, it is wise to add 4K bytes to allow for buffer size.

EXECUTION TIME

Execution time increases as the square of the number of zones. For a nine zone sample problem, execution totals 22 seconds of CPU time on a 360/65. For a 291 zone application, execution time was 9 minutes CPU time using a 360/65.

DATA INPUT

Data Input - DRAM

Card 1 - Format 7 I 5

Variable

1 - N - the number of zones

2 - NINT- the number of intervals in retail probability function

3 - INT - interval width, retail probability function

4 - IRND- projection round i.d. number.

5 - IPRNT - print option for intermediate calculations

6 - NG - number of household groups

7 - NK - number of employment groups

Card 2 - Format F5.1

Variable

1) -TITL- title of run, 0 to 80 characters

Card 3 - Format F5.1

Variable

1) Beta (1) - impedance parameter, retail allocation function, worksite demand.

Card 4 - Format F5.1

Variable

1) -Beta (2) - impedance parameter, retail allocation function, residential demand.

Card 5-8 - Format 5X, 5F10.5

Variable

1-5 PLT - Land use and transportation parameters, residential allocation function, one card for each income group, 5 parameters per card.

Card 9-12 - Format 5X, 5F10.5

Variable

1-4 PID - income distribution parameters, residential allocation function; 1 card for each income group, 4 parameters per card.

Card 13-16 - Format 4F10.5

Variable

1-4 CNV - conversion matrix, converts employment by employment type to employment by income group. Four parameters per card. (Note: the following data may be input from cards or mass storage devices. The documentation listed here is for the 9 zone sample problem).

File 11 -1 record per zone.

Variable

1 - zone ID

2-5 - HH - employed residents in each income group

6 - AP - population

7 - TEPR - total employed residents

8 - GQPS - Group quarters population

9 - AHU - housing units

10 - TAA - total land area

11 - AU - unusuable land

12 - AAAB - land used for basic employment

13 - AAC - land used for commercial employment

14 - GAAR - residential land

15 - AAV - total useable land

16 - A9 - streets and highways land

17 - AVI - vacant land

18 - EMP(I,4) - base year service employment

File 12 - 1 record per zone.

Variable

1-9 - impedance from each zone to each of the other zones

File 4 - 1 record per zone.

Variable

1-3 - EMP - employment in three categories, basic employment forecast

- AAAB- basic land use

Card 17 Format 5X,3F10.0,F10.5

Variable.

1 - EMFR 2 - dummy variable

2 - EMFR 3 - dummy variable

3 - EAMPRO - total service employment, target year

4 - PPI - factor to increase or decrease total employment, used in policy testing

Card 18 - Format 5X, F10.0

Variable

1 - PROPOP - total projected population

Card 19 - Format 3F10.0

Variable

1 - Total number of work to home trips

2 - Total number of work to shop trips

3 - Total number of home to shop trips

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

MODELOS DE SIMULACION PARA LA PLANEACION INTREGRAL DEL USO DEL SUELO Y EL TRANSPORTE.

TEMA: VARIOS.

PROF. ING. STEPHEN H. PUTMAN.
PROF. ING. FEDERICK W. DUCCA.

ENERO, 1978.

Palacio de Minería Calle de Tacuba 5, primer piso. México 1, D. F.

INTRAURBAN EMPLOYMENT FORECASTING MODELS: A REVIEW AND A SUGGESTED NEW MODEL CONSTRUCT

Stephen H. Putman

Urban employment location has frequently received short shrift in simulation studies. Although the lack of good data is largely responsible for this, few attempts have been made to develop a model framework which might help specify what data were required. A review of prior urban employment modeling efforts is presented along with a classifying principle which helps to clarify the problems, and a structure for a comprehensive model is proposed. Although it is not likely that the model can be implemented in the near future, discussion of the structure can, hopefully, lead to further attempts to resolve the problems of urban employment forecasting.

The processes which result in the spatial distribution of employment in urban areas are among the more important phenomena in the urban system. To date, with the possible exception of the retail sector, attempts to explicitly describe these processes have rarely been successful. Many attempts at urban modeling sidestep the issue by assuming that most or all urban employment activity is determined exogenously (to the model). The basis for this separation is frequently not well developed, often being the result of intuitive feelings of the modelbuilder about basic and nonbasic employment characterizations. This article reviews a number of the more important modeling attempts and suggests some appropriate steps leading to improved forecasting ability.

The article is divided into three main sections. The first section discusses the problems of classifying employment types. Such a classification is important in discussing the adequacy of existing and proposed models of employment location and growth. The second section discusses existing models of intraurban employment location for both non-market-sensitive employment and market-sensitive employment. The third section describes a suggested new model of intraurban employment location which attempts to combine features of existing models.

Classification of Employment Types

A key decision, usually made early in employment modeling efforts, is the determination of which types of employment are to be included in the model, and of how they are to be handled. Consequently, the question of employment types is critical: depending on their definitions, some sectors may be inappropriately modeled or even excluded from consideration. Many of these classifications stem from economic base theory and implementation of economic base studies, and as such are defined in terms of the industry's output. Although this method of classification was quite appropriate for early attempts to model urban employment behavior and was entirely consistent with most other economic analyses, a more rigorous and comprehensive classification can and should be attempted. Previous classifications have also been based on observed locational characteristics or on very specific attributes of a given sector in a given area. Such classifications, while an improvement over the earlier economic base schemes in that they are based on "industrial" input characteristics, are still somewhat restricted in scope. A new classification would be based on a much more general, though explicitly described, examination of input requirements. This method, in effect, produces a classification based on production function attributes which, in turn, relates to the theoretical construct on which the model(s) is (are) based. One such scheme of employment classification is discussed in Hamilton et al. (1966, pp. H-1 -I-30), and another, somewhat different procedure is discussed in Rose (1968, p. 34). A third scheme, based (though not explicitly) on differences in production functions and oriented toward intraurban analyses is discussed below.

If it is assumed that every employment type may be considered as performing three primary activities: procuring; processing; and distributing (Hoover, 1948), then by characterizing the ways in which these functions are performed by different types, some bases for classification can be derived. For example, certain types of manufacturing industries requiring bulky inputs (procurements) which are best delivered by rail transport will be likely to locate on a site near a rail terminal or siding. The nature of these industries' outputs may require additional facilities, for example, good access to main intercity highways. Another type of manufacturing industry, having much smaller (physical) input requirements, might be most influenced by relative ease of access to the local airport for priority shipment of its outputs. A research facility whose primary input was skilled labor might tend to locate in a suburban area convenient to the employees' residences.

The primary activities may be defined as follows:

1. The procurement aspect includes everything associated with obtaining the required material inputs and getting them to the location where they are to be used.
2. The processing aspect includes all the costs of performing whatever processing function the employment performs. For example, in a manufacturing sector it would include costs of land, labor, and capital for production. In a service sector, it would include the same costs but would undoubtedly be dominated by the labor cost.
3. The distribution aspect includes all the costs associated with the sale and delivery of the product or function performed.

This classification scheme encompasses a wide variety of factors: for example, availability of labor, the cost of which presumably increases as the supply of it decreases, and proximity to markets, which is important to certain employment types because consumer behavior may dictate high distribution costs for a production site located far from a market (that is, set of consumers) whose members do not traditionally travel far to obtain its specific product.

This classification is developed under the following conditional statement: *If* a facility (that is, a group of employees) is to be located in a given urban area, *then* these are the factors which will determine its locational preference. A preliminary classification scheme may thus be developed based upon the procurement-processing-distributing trichotomy. At the same time, for each employment type, it is also necessary to consider whether any or all of these elements of activity are *intraurban location* determinants. That is, to classify an employment type according to this scheme it is first necessary to determine whether the particular aspect of employment activity (procurements, processing, or distribution) is of locational importance; then it must be determined if there are sub-area differences within urban areas which will affect the performance of this activity.[1]

A Review of Existing Models

The following discussion assumes that the problem being addressed here is that of small area intraurban employment location and that citywide totals of each employment type are given. In addition, it is assumed that within the urban area, the size of the sub-areas being considered is relatively small, being on the order of census tracts. The first assumption has a substantial impact on the nature of the problem. By restricting the problem to *intra*urban location, interurban or interregional comparisons are obviated. That is, many location decisions consist first of selecting a region or a metropolitan area in which to locate, and then of selecting a specific site within the chosen area. The discussion presented here focuses only on the latter part of the problem. The first part of the problem is an important aspect of *regional* employment models and is discussed elsewhere (Putman, 1969, 1970).

A most important aspect of the proposed employment classification scheme is the determination of intraurban differences with regard to the distribution aspect. This subset of the proposed classification scheme refers to the location of "markets" and the importance of their location to the employment sector being considered. Existing models of urban employment activity are frequently classified by the sectors which they attempt to model, for example, retail or manufacturing. The discussion above indicates that an alternative classification of existing models can be made by analyzing the means by which they consider "markets." Given this, *there are some employment types whose location within the urban area will be market-sensitive and some whose location will not.* In many urban modeling efforts, those types whose intraurban location decisions are not market-sensitive have been treated as being exogenously determined. In a few cases, attempts have been made to model the locational behavior of these types of employment. These attempts have had some, albeit limited, success but in all cases they have required substantial additional work.

In contrast to the few attempts at modeling non-market-sensitive location, there have been a number of attempts to model market-sensitive intraurban employment location. Employment

BIOGRAPHIC SKETCH

Stephen H. Putman is Research Investigator, Department of City and Regional Planning, University of Pennsylvania. He heads the Planning Sciences Simulation Laboratory, a research facility which makes available for student and faculty use more than a dozen of the well known urban and regional simulation models. Dr. Putman has been a consultant to a number of urban and regional simulation studies and has developed and tested several such models.

sectors which in some studies were considered non-market-sensitive, however, were considered as being market sensitive in other studies. Although this difference was sometimes due to differences in geographic level of detail, in other cases it was due to differences in interpretation or assumptions or both on the part of the modelbuilders. Typically these models develop measures of the "potential market" for the sectors of interest in each sub-area being considered. The growth of each sector in each sub-area is then proportional (via more or less complex functions) to the market potential measure. Although these models have met with greater success, work is needed here as well.

MODELS OF NON-MARKET-SENSITIVE LOCATION

First it must be recognized that the models discussed here are not necessarily non-market-sensitive in their constructs; rather, they are considered by their authors to be capable of producing forecasts of non-market-sensitive employment location.

There are several discussions of the determinants of intraurban plant and facility location in the literature (Carrier and Schriver, 1968; McMillan, 1965; Stefaniak, 1962; Townroe, 1969). These studies were basically empirical efforts to determine what factors locators considered important in making the location decision. Along with other studies of a similar nature (Creamer, 1963; Fuchs, 1962; Luttrell, 1962), they provide useful information leading to the formation of models, but they are not themselves models of location.

A second body of literature focuses on various mathematical programming treatments of the problem (Stevens and Coughlin, 1959; Cooper, 1963; Lefeber, 1958; Goldman, 1958). Although these articles provide interesting theoretical frameworks for speculation, the loss of realism incurred by forcing them to fit the somewhat restrictive linear programming form makes them operationally useless. Goldman (1958, p. 97) does not list intraurban location forecasting as one of the possible uses of his model. Specifically, the problems here are the required linearity of model equations and the rather limited consideration of location determinants other than transport costs.

The largest class of intraurban employment models is that of the "continuous function allocation" models. These models, with few exceptions, have primarily market-sensitive constructs; they are mentioned here because they sometimes include non-market-sensitive employment types as well as market-sensitive types. One exception is the LINTA model (Seidman, 1969, pp. 135-138) developed to forecast the location of manufacturing employment (which is predominantly non-market-sensitive) by small area. This model first spreads the regional growth rate of a given activity over all sub-areas. Then, based on sub-area specific measures of "attractiveness," the model reallocates the growth to different sub-areas in proportion to their attractiveness. Although tests of this model were moderately successful, its high level of sectoral aggregation and its inability to consider discrete facility location are important deficiencies. This is particularly the case for urban regions where the economic attributes (that is, determinants of growth) of sub-areas are sufficiently similar as to result in the model's being capable of forecasting only small regular changes in all sub-areas rather than the discrete irregular changes which so often occur.

Another continuous function allocation model designed specifically for forecasting non-market-sensitive employment is Nathanson's BEMOD. This model (for which a complete technical description is not available) was developed by the Bay Area Transportation Study Commission to provide basic employment inputs to their PLUM population and nonbasic employment model (Bay Area Transportation Study Commission, 1968). Based upon certain attributes of each sub-area, a sector (employment type) specific score or index is calculated for the sub-area. When all sub-area scores have been computed, the regionwide growth or decline for each employment type is allocated to the sub-areas as a function of the "score" and available land. Excess allocations are handled by repeating the procedure with modified land availabilities (Stevens and Wheaton, 1970, p. 3). This model probably suffers from the same problems as LINTA, while performing somewhat better overall, but lack of published documentation makes further evaluation impossible.

The EMPIRIC model is another continuous function allocation model (Hill, 1965). This model is similar in principle to the LINTA model (which was developed from the EMPIRIC and POLIMETRIC model efforts). The EMPIRIC model's construct is described by Hill in the following way:

> The change in the subregional share of a located variable in each subregion is proportional to one, the change in the subregional share of all other located variables in the region, two, the change in the subregional share of a number of locator variables in the subregion, and three, the value of the subregional shares of other locator variables. [p. 113.]

The comments made above regarding LINTA are equally applicable here.

The Loewenstein model

> combines a trend-based analysis of industrial sector growth within districts in the Philadelphia region with an allocation technique based on the attractiveness of each district relative to all other districts in the region for each industry under consideration.
>
> . . . the critical notion of this process is that it ascertains the importance of a number of locational determinants for a variety of industries as well as the availability of these determinants in a number of areas. The areas are ranked according to the degree to which they possess these characteristics and an apportionment is made. [Lowenstein, 1967, pp. 3-6.]

In this model, which was never made operational, the employment types are two-digit S.I.C. classes. The question of simultaneous versus sequential location of different industry types is not discussed in the model description. Instead, the problem is bypassed by assuming that the location of each industry will be independent of the location of each other industry. Each of the location factors was weighted according to its importance to the locating industry. The model deals with allocation of employment increases or decreases as a continuous function allocation, and therefore there is no constraint on the allocations except that of employment density by sub-area. The absence of such a constraint presents the possibility that over longer projection periods this model may forecast a uniform density by industry types as the predominant characteristic of industrial location (although operationally this may never occur). There is no consideration in the model of the possibilities of discrete facility location. The useful aspect of this model construct is the list of factors it postulates as influencing intraurban industrial location.[2] Although it would not be useful to try to implement this model, consideration of these location factors for use in another model might prove rewarding.

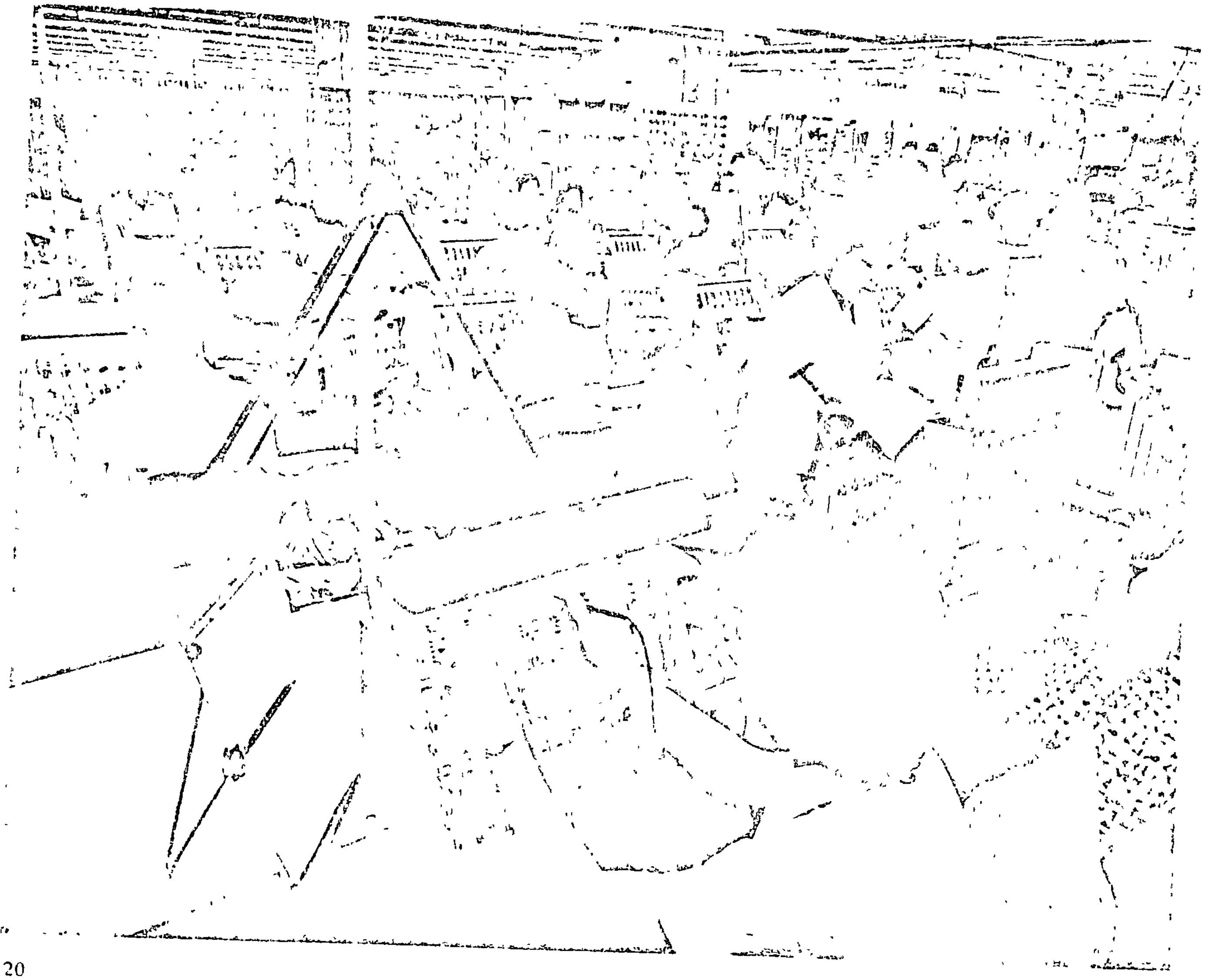
In summary, some continuous function allocation models are capable of producing reasonably good forecasts of non-market-sensitive employment location in some circumstances. However, they are unable to deal with discrete facility location. This inability is likely to be a significant problem in dealing with non-market-sensitive employment types. When these models attempt to forecast significant changes in the location of non-market-sensitive industry types, they are open to this criticism. Thus, although these models are capable of producing reasonable forecasts for market sensitive employment types, as will be discussed below, their capability to deal with non-market-sensitive employment is restricted to situations where new or relocating facility location is unlikely to take place and where growth or decline (or both) are more or less regularly distributed.

The most promising models of non-market-sensitive location appear to be those which attempt to forecast discrete facility locations. A few such models have been developed, including one by Goldberg (1967).

Goldberg describes his model in the following way:

> [It] functions by identifying for each industry group a set of location factors it needs in finding a site. For each site (in the case at hand for each census tract) the factors are measured and added together to give a score for each census tract for each industry group. The tract with the highest score for a given industry group receives a unit of employment. The scores are then usually recalculated and the allocation continues as above until all projected employment is allocated. [p. 3.]

In this model the employment classes are two-digit S.I.C. classes aggregated into six manufacturing subgroups. One of the problems of this type of model is its sequential, rather than simultaneous, treatment of industry location. In this model the problem is sidestepped by grouping the industry types in such a way that a different set of relevant location factors (of a total set) can be defined for each. This avoids the problem of competition between sectors by precluding it. At the same time, it vitiates the need to weight the



location factors because only the "relevant" factors are considered in any case. The model does not provide for weighting of these factors, however. The most serious problem with the model construct is its unit of allocation. When there is a predicted regionwide increase in a given employment type, the model allocates this increase to sub-areas in the region in clumps called "units of employment." Each unit is equal to the average facility size for that type employment in the region. This method of allocation represents a relatively unsatisfactory compromise between continuous function allocation (that is, any number of employees of a given type to a sub-area) and discrete facility location (that is, location of discrete facilities in a sub-area according to a known size distribution). This compromise is unsatisfactory for two reasons. First, from the point of view of theory, the model considers location from an allocative (that is, continuous function allocation) standpoint in calculating "tract" scores, but then adopts a location (that is, discrete facility location) standpoint by using clumps of employment (some of which might be quite large) as the unit of allocation. Second, from a practical point of view, the "lumpy" allocations of employment which could result from this method may well force heavy reliance on arbitrary scaling or smoothing procedures in order to produce reasonable results.

Finally, this model deals with employment decreases in a rather peculiar fashion. The regionwide percentage decrease for a given employment type is multiplied by 1.2 (that is, 120 percent), and a continuous function allocation of this decrease is made, where the function used is a total activity density index, weighted by the specific employment types, of the sub-area. An excess decrease of 20 percent results from the above. This excess is then allocated back to the sub-areas as a 20 percent regionwide increase of that employment type. This procedure implies that those sub-areas which would grow the most if a region's employment of a given type were growing are also the ones that, if the region's employment of that type were declining, would decline the least. Although this is a reasonable assumption, the documentation of the model is unclear as to why this particular procedure was adopted. Despite these shortcomings, some of which stem primarily from data inadequacies, the model has been made operational and has produced reasonable forecasts that have been used as a part of the Bay Area Simulation Study (BASS). For these reasons, further and more detailed evaluation of this model is justified.

Another discrete facility location model, Putman's INIMP model,

> uses an allocation procedure for distributing a portion of projected citywide employment changes (positive or negative) among existing facilities and, upon reaching certain critical values of the magnitudes of allocated increases or decreases, switches to a location technique for the addition or deletion of facilities comprising the remainders of the citywide projections. [Putman, 1967, p. 203.]

In this model, too, the employment types were two-digit S.I.C. classes. The areal unit was the census tract

The allocation portion of the model is a simple continuous function allocation procedure. Increases or decreases dealt with in the allocation portion of the model are proportional to the portion of the city's employment of each type in each area. Although this procedure produces reasonable results when the model is used, it is definitely a weakness of the model. This portion of the model ceases to function after certain capacity constraints in terms of S.I.C. specific levels of employment density have been met.

The location portion of the model, which takes over when the capacity constraints are exceeded, locates clumps of new employment in areas according to a set of measures which indicate the likely preference of the industry for the area, based upon known past location behavior. The size of these clumps is determined stochastically from the prior distribution of facility sizes for each industry. Location of clumps of employment-decrease occurs when the citywide percentage decline of a particular industry reaches certain magnitudes called "shutdown percentages." When this occurs, facilities are selected stochastically for deletion, depending upon the magnitude of the percentage decrease. Again, in the discussion of the model no mention is made of the question of sequential versus simultaneous location. Another weakness of the model is that it follows a sequential (by S.I.C. number) location procedure.

The flexibility of this model construct is greater than the flexibility of other constructs. As a consequence, however, it is significantly more difficult to establish the parameter values (for example, capacity levels and shutdown percentages) for this model. In addition, the model uses only four measures of area unit acceptability. This number should probably be increased. Despite these failings, the model is conceptually quite complete and offers significant promise for future development.

The Rose model,

> rather than analyzing only the net change in employment in a zone, attempt[s] to explain that change specifically in terms of its components: natural growth of existing firms, "death" and "birth" of firms and movement of firms within the region.
>
> . . . Only "births" and "moves" are locationally relevant decisions—size changes (including "deaths") are a function of the industrial "climate" of the region, the type of industry, the size of the firm—with few exceptions (e.g. Urban Renewal) the characteristics of the locations have little to do with the behavior of locationally static firms. [Rose, 1967, pp. 5-6]

The employment sectors used in this model, a total of fifty, are approximately the same as those delineated in the national input-output tables. Although the model is not described in an operational form, it appears to attempt simultaneous location. In a later document (Rose, 1968) the allocation portion of the model is described as a linear programming formulation. However, the number of sectors (fifty) and the number of areal units (647) indicate that operational simultaneity in this model may be very nearly impossible.

In this model all employment forecasts are treated as discrete facility locations. Of particular interest is the means by which the facilities to be located are generated. First, based upon regional employment projections and the base-year firm size distributions, regional projections of the firm size distribution are made. The base-year firm size distribution is then "aged" via a Markov process,[3] yielding an updated firm size distribution. The difference between this distribution and the exogenous regional projection is the estimate of firms which were "born" or which "died" during the period. If the data and model construct were available, firms making intra-area movements would be estimated here also. Then, via a mathematical programming algorithm, the new or relocating firms are allocated to zones.

Although this model is not described in an operational form, and appears to have only discrete facility location (that is, no continuous function allocation of even small employment increments), the idea of facility "birth" and "death" via a Markov process is worth considering as an area for future research.

The empirical studies of market-sensitive location, analogous to empirical studies of non-market-sensitive location, have often been studies of shopping center locations. Like the similar non-market-sensitive location studies, these empirical studies provide information which is useful in the development of models, but they are not themselves models.

The development of normative models of market-sensitive location was one of the earliest topics of quantitative spatial economics (Hotelling, 1952, Losch, 1967). However, as was the case with normative (for example, mathematical programming) models of non-market-sensitive location, normative models of market-sensitive location are too restrictive in their assumptions or data requirements or both to be of any practical use in the near future.

Virtually all existing operational or potentially operational models of market-sensitive location are continuous function allocation models. Future progress in this area is most likely to come from further development of one or more of these models. As many of these models are rather similar in their theoretical constructs, the reviews which follow attempt to cover the principal constructs rather than to describe all such models.

The retail market potential model (Lakshmanan and Hansen, 1965) is based on the premise that the size and number of retail establishments in an area is a function of the number of consumers as represented by the magnitude of their retail expenditures. This model is descended from an earlier one (Lakshmanan, 1965) and has as its central theoretical construct a measure of market potential derived from the traditional gravity model construct (Huff, 1962). This construct, in the words of the authors, states that

> the retail center of zone j attracts consumer dollars (from consumers from zone i)
> a) in direct proportion to the consumer expenditures
> b) in direct proportion to its size
> c) in inverse proportion to distance to the consumers
> d) in inverse proportion to competition. [Lakshmanan and Hansen, 1965, p. 135.]

Although the model construct implicitly assumes that the location of retail activity is determined solely by the above market potential function, it does produce reasonably good forecasts. Two important considerations are omitted from the model, however. First, there are other influences on retail activity location (for example, accessibility to markets and sources of supply, and considerations of land availability and cost) that will cause location deviations from those forecasted by the above construct. Such factors should be included in a realistic model of retail activity location. Second, and perhaps of less importance in already developed areas, is the question of minimum facility size. That is, certain types of market-sensitive activities have important scale economy considerations that will dictate minimum sizes for new facilities. Future models of retail activity location should include these considerations.

The EMPIRIC model (Hill, 1965), discussed above in the section on models of non-market-sensitive location, is also a model of market-sensitive location. The sets of variables affecting location tested in the model were:

Set I: intensities of land use (densities), zoning practices, and automobile accessibilities (within the region);
Set II: same as Set I plus transit accessibilities
Set III: same as Set II plus quality of water service and quality of sewage disposal service.

The results of the calibration tests of the model were quite acceptable, and the model may be considered as an acceptable, operational, forecasting technique, subject to the restrictions mentioned above.

The Harris retail trade model (1964) was developed as a component of the Penn Jersey Activities Allocation Model. Derived from selected portions of central place theory, the model attempts to describe the equilibrium which results from the supply of goods and services, the demand for them, and the pattern(s) of interaction between them. The model

> works in the following way. Given the location of demand and any set of interaction parameters, a computer program is provided with a preliminary estimate of the location of supply. The interaction formulas are then applied and result in a hypothetical set of "arrivals" at the supply points.
>
> . . . the result of this first estimate of "arrivals" is substituted for the original hypothesis as to the distribution of supply, and the process is reiterated. [Harris, 1964, p. 6.]

Although acceptable forecasting results have been obtained from this model, the estimation of its parameters (which may be highly collinear) is more difficult than usual. Therefore, the model cannot be considered to be a truly practical forecasting device. Its explicit consideration of supply and demand equilibrium is, however, unique and should be considered for future development.

The Lowry model (1964), as well as its derivatives (for example, Crecine, 1964), considers the location of market-sensitive employment simultaneously with the location of population. This model is also a market potential model, but the formulation of the concept is somewhat different. Lowry describes it as follows:

> The distribution of this retail employment among the . . . tracts depends on the strength of the market at each location. Assuming that shopping trips originate either from homes or from workplaces, the market potential of any given location can be defined as a weighted index of the numbers of households in the surrounding areas, and the number of persons employed nearby. [1964, p. 10.]

This model, too, is capable of producing acceptable forecasts, and its parameters have been successfully estimated for several different areas. The model is an operational forecasting technique.

The Bay Area Simulation Study developed several continuous function allocation models to deal with market-sensitive employment of different types (Center for Real Estate and Economics, 1968, pp. 179-234). In determining allocation of retail trade employment, an estimate is made of site suitability as a function of site attributes (including accessibilities). Expected retail trade demand is developed from a simple market potential expression. The expected demand is compared to the actual demand (existing employment) to

estimate potential "new" demand. Finally, a relative attractiveness index based on a combination of site suitability and potential "new" demand is calculated. Retail trade employment is then allocated in proportion to each sub-area's relative attractiveness. Services employment is allocated in proportion to a relative attractiveness index and the population percentage in each sub-area. Finance and government employments are allocated by a function of past employment and population. Contract construction is allocated in proportion to new houses and new employment in all other sectors.

These BASS employment submodels represent operational but highly empirical continuous function allocation procedures for forecasting the location of market-sensitive employment. It would be desirable to try to develop a more general, less empirical, type of procedure to make these forecasts.

Thus, there are continuous function allocation models which produce quite acceptable forecasts of market-sensitive location. All but two of these, however, omit consideration of scale economies. The Lowry model considers them in a rather crude fashion, and the Harris model considers them explicitly. Although the omission of scale economies is not a critically important failing, such economies should be considered in future model-building efforts. A more important problem is that many of these models include non-market-sensitive employment activities, such as office employment, among market-sensitive employment activities. This problem reflects back to the classification question posed earlier. For certain employment classes these models appear to be quite adequate. For other classes, other models are necessary. Thus a model for forecasting all employment types will have to be capable of both continuous function allocation and discrete facility location. It will also have to be able to determine which procedure is appropriate under what circumstances.

A Suggested Intraurban Employment Model

Based upon the work which has been done to date (much of which has been reviewed above), it is possible to describe the general structure of a model which should be capable of producing reliable and accurate forecasts of virtually all intraurban employment distribution with the principal exception of "unique locators." These are facilities whose location is largely determined by: (a) public policy (for example, large government facilities); (b) private policy having more to do with internal economies which are only vaguely determined by external factors; (c) special requirements such as rivers or land, (for example, airports). Any facility whose planned location is known and which will be located irrespective of the model's forecasts is also a "unique locator." Any model of intraurban employment must be capable of accepting prespecified unique locators as part of its input.

It is first necessary to define what is meant by facilities and firms. The facility life cycle concept is introduced in the Rose model. No distinction is made there between a firm's life cycle based on its profitability, and a facility (that is, productive facility) life cycle based on aging and changes in the relevant production technology. The Rose model implicitly assumes, therefore, that the life cycles of viable firms and usable facilities exactly coincide. In the following discussion, firms are defined as groups of employees. Such groups may be independent business entities or divisions (branch offices) of business entities. The structures which they occupy while doing their jobs are facilities. The model construct proposed below attempts to deal with this problem.

The model construct would begin by separating market-sensitive and non-market-sensitive employment sectors as discussed above. For both of these, the model would be further divided into an allocative section for growth or decline of existing firms and a locating section for birth or death (including relocation) of discrete facilities.

GROWTH OR DECLINE OF EXISTING FIRMS

One of the determinants of growth or decline of an existing firm in a particular area is the relative advantage or disadvantage in procurement that the area offers the firm. The measurement of procurement advantage presents the problem of determining the source (geographic as well as sectoral) of the firm's required inputs. Determination of the sectoral source of inputs, while difficult, may often be made: (1) by assuming that national or regional input-output coefficients, if available, hold in the specific urban area; (2) by applying or adapting local input-output studies, where available, or (3) by conducting special surveys. Determination of the geographic source is a different matter in that the assumption of national data (if any) is not relevant. It must first be determined whether inputs come from within or outside of the urban area. This would have to be done either by special survey or by rather gross assumptions. In view of the expense of special surveys, it is likely that this problem will be resolved by assumption. An assumption that is too general can be avoided in the following manner. Because the question is one of *relative* advantage of alternative intraurban sub-areas, it is not difficult to develop any of several types of accessibility measures to intraurban sources of inputs. External sources can be dealt with under any of the following assumptions: (1) that no urban sub-area has any advantage vis-a-vis any other for obtaining inputs from sources geographically external to the whole urban area (the least satisfactory assumption); (2) that most such externally procured inputs enter the urban area through a small number of points (for example, highway interchanges or transportation terminals) and consequently a measure of intraurban advantage (accessibility) can be constructed using only these points to represent all external sources, or (3) that major external sources may be identifiable and specific access measures from these sources to intraurban sub-areas can be constructed. It is likely that some combination of the second and third alternatives would be more than adequate for modeling purposes and would therefore allow the development of measures of procurement advantage or disadvantage.

Next it is necessary to develop measures of processing advantage or disadvantage. Both physical factors (plant and equipment), as well as labor considerations, are relevant here. One nonlabor component of processing cost is tax expense resulting from land and building values. This component might be represented by a measure of assessed value (as a surrogate for the usually unavailable sales value) of plant and property per employee, although even within given urban areas there are likely to be anomalous variations in assessment practices. One other possibility is an investigation of equalized tax rates as a surrogate variable. This same variable could also be a surrogate for land cost. Another measure, representative of certain attributes of older (obsolete) plant and equipment, should be developed. (Such attributes may or may not be desirable, depending on the type of employment.) For example, an age-of-structure index can be calculated for each sub-area and might be indicative of such costs. Alternatively, the same phenomenon might be represented in terms of structure-area to site-area ratios which would be indicative of age of structures, that is, of older, perhaps multiple-story or crowded buildings. It is unlikely that variables which directly measure the phenomena of interest here will be readily available. Indices or measures of relative advantage of one sub-area vis-a-vis the others are what is desired, however. Therefore, the use of various surrogate measures is a reasonable approach.

The availability of the appropriate labor force is the most important consideration in measuring processing advantage with regard to labor. First it is necessary to know what types of labor force are appropriate. Given this knowledge, additional information as to the magnitude and spatial distribution of the labor force will allow development of measures of accessibility. If the definition of the urban area being modeled excludes large supplies of labor (for example, suburbs), then these supplies must somehow be brought into the model, perhaps by artificial network links or similar means. (It is assumed that sub-area to sub-area wage rate differences are not large enough to have a significant effect on labor availability.)

A final measure, perhaps the most important for many employment sectors, is that of relative distribution (market) advantage. Here, as with the procurement measure, there are the dual questions of sectoral and geographical location of markets. These factors can be represented with one or more market potential measures, that is, with indices of access to markets which include the effects of competitors. Measures of relative distribution advantage require the same kind of assumptions as the procurement measures described above, specifically, assumptions about the possible importance to intraurban location or to differential growth of firms of markets external to the urban area. It is likely that intraurban locational sensitivity to external markets will be significant only when there are few specific intraurban points from which interurban connections are made and when there are significant differences in intraurban accessibility to those points. On the other hand, differences in intraurban access to intraurban markets will be of major importance to many employment types.

The measures of procurement, processing, and distribution advantage can now be used as independent variables to estimate differential growth of intraurban employment types in a continuous function allocation procedure.

"BIRTH" OR "DEATH" OF DISCRETE FACILITIES

In forecasting growth and decline of existing firms, the model will encounter situations where it is appropriate to begin generation or deletion of discrete firms or facilities. The obvious difficulties lie in determining precisely when this transition will take place and whether a firm or a facility is to be the unit allocated. Briefly setting those questions aside, two principal methods of approaching the problem suggest themselves. In the first approach, the discrete firm or facility changes are determined first, with residual growth and decline being dealt with in the continuous function allocation portion of the model. The second possibility is to have facility capacity or percentage change limits intervene in the continuous function allocation growth or decline processes, in order to generate new facilities or to have facility shutdowns, triggered by percentage decline limits and other measures that indicate shutdown or relocation likelihoods. The first alternative leads to the modeling of facility lifecycles, entailing the associated difficulties of parameter estimation, as in the Rose model. The second alternative leads to a model that is sensitive to difficult-to-estimate parameters, as in the Putman model. The most substantively satisfactory solution would be an integration of life cycle modeling procedures and continuous function allocation procedures. This solution is discussed below.

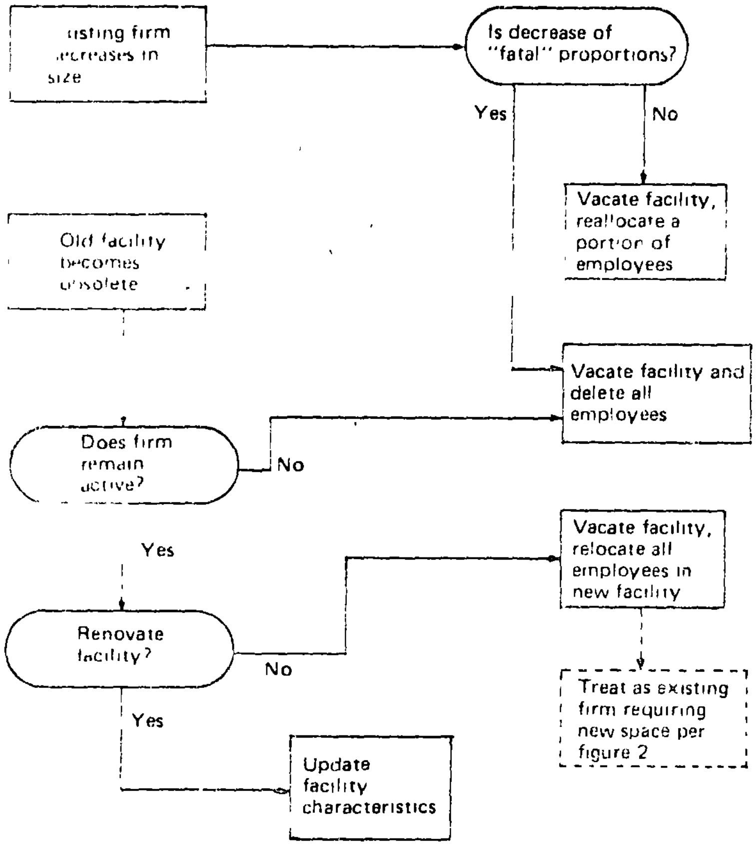


FIGURE 1 *"Decline" Procedure: Firm Decline or Facility Obsolescence*

Determining the consequences for employment location if an existing firm or facility declines is a fairly complex problem. If a firm outgrows its existing facility, it may choose to renew the facility, or to vacate it and relocate all employees. The firm might also choose to take that opportunity to cease operations, thereby vacating the facility and terminating the jobs of its employees. Certain alternatives intermediate to these may also be possible, but they are of less significance to the model structure. If, on the other hand, an existing firm contracts (shrinks) in size, the consequences may be different. If the contraction is complete, that is, if operations cease, then the results are as mentioned above. If the contraction is partial, but great enough to vacate a facility, the remaining employees are allocated elsewhere. Smaller contractions are handled in the continuous function allocation portion of the model. This structure is shown in figure 1.

The determination of the employment location effects of new firms or facilities is somewhat more complex. If an existing firm, due to expansion, requires additional space (that is, facility), the first decision is whether or not the existing facility will be vacated. If the facility is not to be vacated, the problem is similar to that of locating a new firm and will be discussed below. If the original facility is to be vacated, then the next decision is whether the firm will relocate to an existing vacant facility or to a newly constructed facility. If the move is to be to an existing vacant facility, then a search must be made for such a facility. If the move is to be to a new facility, then land requirements for the new facility must be calculated and a search instituted for a suitable site.

The problem of locating a new firm presents a similar "decision tree" to analyze. First, will the firm locate in an existing vacant facility, or does it require a new facility? If it is to locate in an existing vacant facility then a search for one must be made. If it is to locate in a new facility, the requirements must be calculated and, again, a suitable site searched for. Regardless of which alternative has been selected, suitable site area or vacant building space may not be available in the sub-areas most desirable to the locating firm. This raises the question of trade-offs between "site" attributes and "site" availability. So little is currently known about this phenomenon that it is not possible to predict what headway might be made here. In fact, all of the processes discussed above are rather poorly understood and will form a large unknown component of future research efforts on this topic. Some data on these behaviors do exist, and their importance to the forecasting of intra-urban employment location demands their inclusion in future modeling efforts. The decision processes described above are shown diagrammatically in figure 2.

LOCATION OF NEW AND RELOCATING FIRMS AND FACILITIES

The model should be able to determine which sub-area is the most desirable for a particular locating firm or facility. There are several problems which need to be resolved here. The first is that of measuring an area's attractiveness to different employment sectors. This may be done by calculating indices based upon the measures defined for use in the continuous function allocation portion of the model. It may be found from studies of past locational behavior that other, noneconomic, factors must be determined in order to estimate locations. It is possible that factors not included in the procurement, processing, and distribution tri-

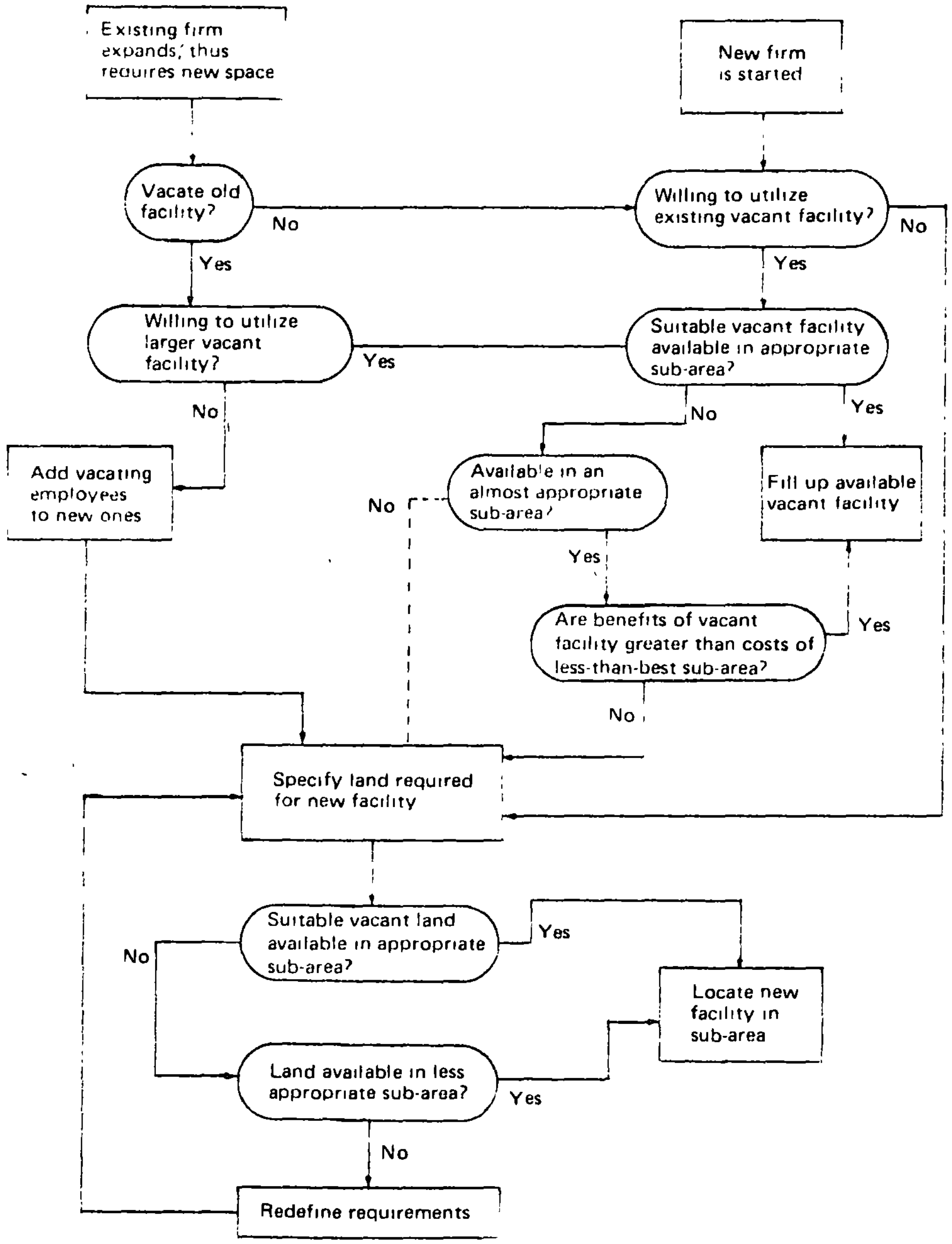


FIGURE 2 *"Growth" Procedure: New, Relocating, or Expanding Firms* chotomy may also be found to be of some importance in location decisions For example, locations offering prestige or "visibility" might account for some portion of a location decision, though data and research findings on these factors are conspicuously absent from the existing literature.

The first studies mentioned above in the review of models of non-market-sensitive location are good sources of factors to be investigated. A critical question also arises regarding the availability of the data necessary to develop the model's hypotheses It would be foolish to deny the difficulties likely to be encountered. On the other hand, it would be at least equally foolish to do nothing about these problems pending the availability of the perfect data set. It is possible to conceive of a number of smaller studies which could be undertaken to examine portions of the overall model being described here.

Another substantive problem is simultaneity of location. All of the present models of firm or facility location manage to avoid this problem. Nevertheless, simultaneity of location may be critical in two ways. First, there is, in all likelihood, competition among locating firms or facilities for particular sites. The importance of this competition for forecasting purposes is presently unknown, and such competition is not dealt with in nonsimultaneous location procedures. Second, sequential location has the effect not only of occasionally precluding location on a "used" site, but also of altering, subsequent to that location, the characteristics of the sub-area in which the location has taken place. These altered characteristics, in turn, influence the subsequent location (particularly, though not exclusively, in that sub area) of other facilities in the same and other sectors. These points notwithstanding, the imposition of a simultaneous location requirement on the model is not a trivial matter. Although such a requirement is substantively desirable, it may not be feasible. If continuous function allocation of market-sensitive employment involves knowing the current spatial distribution of population (and personal income) in order to make forecasts, then population location will have to be simultaneous with at least some market-sensitive employment location. Further, if there is some degree of competition for virtually all land in the urban area, then it is difficult to decide a priori that some activities will be located before or after others. There is no solution to that problem in this article; as the nature of the model becomes more specific it will be possible to determine how it may be resolved and to test the consequences of that resolution.

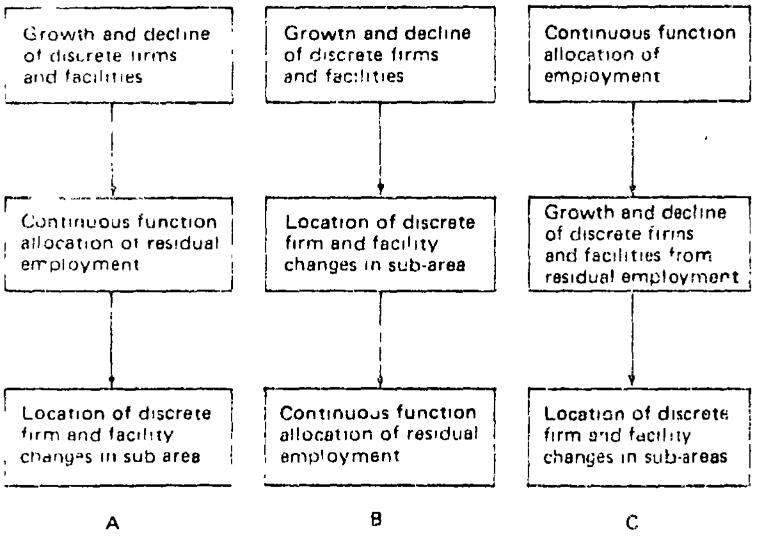


FIGURE 3 *Three Principal Forms of the Proposed Model Construct*

Thus there are several principal forms into which the overall structure of this new model can develop. Figure 3, A and B, shows two versions of the form where discrete firm and facility changes are forecast first and the residual changes are dealt with via continuous function allocation. In figure 3, A, the residual employment is allocated before the discrete firms or facilities are located. In figure 3, B, the discrete facilities are located first, then the residual employment is allocated. In either case, the model would first forecast changes in the numbers of firms and facilities. This might well be done by a modification of the life cycle technique proposed by Rose. The relation of these "births" and "deaths" to the areawide net changes in employment by sector would be an important consideration. In most situations only a portion of the citywide employment change for a given sector is due to the birth or death of new firms or facilities. It will be difficult to determine, in the model, what size this portion should be, since two processes are subsumed here. First, in a given sector there are new firms appearing and old firms disappearing somewhat independently (that is, due to life cycle changes) of the urban area's overall growth or decline in that sector. Second, new firms are appearing and old ones disappearing almost solely due to the area's overall growth or decline in the given sector. Finding a way to discriminate between these two processes is the key problem here. If facilities were generated or deleted as in Putman's INIMP model, the computational procedures would be somewhat simpler but slightly more arbitrary.

Because the accuracy with which the model determines how much of the net citywide change of a particular employment type is due to discrete facility generation or deletion is both difficult and critical to the models shown in figure 3, A and B, an alternative means for resolving the issue is worth considering. An alternative form for the new model is shown in figure 3, C. In this case the continuous function allocation is performed first, with the discrete firm and facility generation, deletion, and location taking place afterward. Here, too, it is necessary to know when to terminate the continuous function allocation and begin discrete facility forecasts. It is assumed that the continuous function allocation is more stable, and probably a more accurate estimator, than the discrete facility portion of the model. Consequently, the form of model shown in figure 3, C, might produce better forecasts than the forms in figure 3, A and B, given the present level of knowledge. It is probable that the *most* accurate form of the model would involve generation of firms and facilities concurrently with the continuous function allocation. Whether this is a feasible structure for the model cannot be determined prior to further research.

Summary and Conclusion

The procedure for the classification of employment is critical to the success of any attempt to model employment location. Previous attempts at such classification have been inconsistent with a comprehensive model of intraurban employment location. It is necessary, therefore, to develop new classification procedures in order to be more confident that each employment type is being forecast by an appropriate portion of the model construct. One possible procedure has been described in this article.

Intraurban employment location is the second step in a two-step process. The first step occurs when a firm decides to locate in, or leave, a particular urban region. The second step, with which this article is concerned, is the determination of where in the chosen urban region the location or departure will take place.

The intraurban location of some types of employment is partly determined by the spatial distribution of markets in the urban area, while the location of other types is not. There is a distinction between the growth or decline of an existing employment establishment (either firm or facility) and the location or deletion of a specific lump of employees either as a firm or a facility.

With these definitions and distinctions in mind, a review of most of the existing or proposed models of intraurban employment location was presented. At the conclusion of this review, a model was developed, integrating features of the reviewed models in an attempt to develop a comprehensive framework for intraurban employment forecasting.

In conclusion, although preliminary work has already been done on many of the components of a complete intraurban employment model, large gaps in our knowledge of these phenomena remain. The conspicuous absence of complete employment forecasts from urban simulation modeling today indicates that work needs to be done in this area. It has been the purpose of this article to review the work done to date and to suggest directions for future efforts to solve these problems. The development of such models and their subsequent integration with residential location and transportation network models would constitute a significant increase in our understanding of, and ability to forecast, urban phenomena.

Author's Note· *Work on this topic was partially supported by the US Department of Transportation, Federal Highway Administration, Bureau of Public Roads, under Contract No FH 11-7392 to the Institute for Environmental Studies, University of Pennsylvania. Portions of this manuscript were included in the final report, "The Widening Role of Urban Transportation Planning," prepared by B. Harris, S Putman, B. Allen, and staff under that contract*

NOTES

[1] Before proceeding with further discussion, a note regarding the modeling of employment location as opposed, for example, to sales or value-added location is in order. There are obvious disadvantages concomitant with the use of number of employees as a measure of activity levels Most of these disadvantages stem from the nonlinearities in each employment type's employee/output relationships (that is, inherently different production functions). Consequently, when these forecasts of economic activity, as represented by employment totals, are used in turn to generate other forecasts, certain inconsistencies are introduced into the results Despite this, employment is the measure most often used, for the single, overwhelming reason that it is the only measure for which data are usually available. If other data become available, then research into their use as a more appropriate means for modeling the intraurban location of economic activity will be justified. For the present and the immediate future, employment will have to continue to be the unit of measurement.

[2] These factors are· access measures to the following—labor supply, labor skills, executive participation, highways, railroads, mass transportation, waste disposal facilities, comparable firms, customers, and suppliers, slope conditions, land requirements, land costs; and taxes on land and improvements. Floor area ratios,

employee density, multiple land use characteristics, and expenditures per municipality have also been suggested as possible location factors

[3] A Markov process is a systems concept involving various "states" of being and various probabilities of "transition" from one state to another

REFERENCES

Bay Area Transportation Study Commission (1968) "BATSC Locational Model System," BTM 217/00/00 (Berkeley)

Carrier, R E and W R Schriver (1968) "Location Theory An Empirical Model and Selected Findings," *Land Economics*, 44(4) 450-460

Center for Real Estate and Urban Economics (1968) *Jobs, People and Land Bay Area Simulation Study (BASS)* (Berkeley Institute of Urban and Regional Development, University of California)

Cook, R W (1965) "Community Economic Profiles and Industrial Location in Delaware," mimeo , University of Delaware, Division of Urban Affairs

Cooper, L (1963) "Location–Allocation Problems," *Operations Research*, 11 (May–June) 331–343.

Cordey-Hayes, M (1968) "Retail Location Models," CES WP 16 (London Centre for Environmental Studies)

Creamer, D (1963) *Changing Location of Manufacturing Employment Part I Changes by Type of Location, 1947–1961* (New York National Industrial Conference Board Inc.).

Crecine, J.P (1964) "A Time-Oriented Metropolitan Model for Spatial Location," mimeo , Pittsburgh. Department of City Planning

Fuchs, V R. (1962) *Changes in the Location of Manufacturing in the United States since 1929* (New Haven· Yale University Press)

Goldberg, M. (1967) "Industrial Location Model for the San Francisco Bay Area," paper presented at the Western Regional Science Association, Las Vegas, Nevada.

Goldman, T A (1958) "Efficient Transportation and Industrial Location," *Papers and Proceedings of the Regional Science Association*, 4 91--104.

Hamburg, J R and G. T Lathrop (1964) "An Opportunity-Accessibility Model for Allocating Regional Growth," mimeo , Albany Upstate New York Transportation Studies

Hamilton, H R , S E Goldstone, F. J Cesario, D. C. Sweet, D E Boyce, and A L Pugh III (1966) "A Dynamic Model of the Economy of the Susquehanna River Basin," mimeo , Columbus, Ohio Battelle Memorial Institute.

Harris, B (1964) "A Model of Locational Equilibrium for Retail Trade," mimeo , Philadelphia Institute for Environmental Studies, University of Pennsylvania

Hill, D M (1965) "A Growth Allocation Model for the Boston Region," *Journal of the American Institute of Planners*, 31 (May) 111 -120

Hoover, E M (1948) *The Location of Economic Activity* (New York McGraw Hill Book Co)

Hotelling, H. (1952) "Stability in Competition " Pp 467--484 in Stigler and Boulding, eds , *A E A Readings in Price Theory* (Chicago Richard D Irwin)

Huff, D L (1962) "Determination of Intraurban Retail Trade Areas," mimeo , Los Angeles Real Estate Research Program, University of California.

Lakshmanan, T R and W G. Hansen (1965) "A Retail Market Potential Model," *Journal of the American Institute of Planners*, 31 (May) 134 143

Lakshmanan, T R (1965) "A Theoretical and Empirical Analysis of Intraurban Retail Location," Ph D dissertation, Ohio State University

Lefeber, L (1958) "General Equilibrium Analysis of Production, Transportation, and the Choice of Industrial Location," *Papers and Proceedings of the Regional Science Association*, 4 77 86

Loewenstein, L K and D Bradwell (1966) "What Makes Desirable Industrial Property," *The Appraisal Journal*, April 263 267

Loewenstein, L K (1967) "A Proposed Manufacturing Location Model," paper presented to the Western Regional Science Association, Las Vegas, Nevada.

Losch, A (1967) *The Economics of Location* (New York John Wiley & Sons)

Lowry, I. S (1964) *A Model of Metropolis* (Santa Monica California RAND Corporation)

Luttrell, W F. (1962) *Factor Location and Industrial Movement A Study of Recent Experience in Great Britain* (London Th National Institute of Economic and Social Research).

Massey, D B (1969) "Some Simple Models for Distributing Changes in Employment Within Regions," mimeo , CES–WP 24, London· Centre for Environmental Studies

McMillan, T E , Jr. (1965) "Why Manufacturers Choose Plant Locations vs Determinants of Plant Locations," *Land Economics* 41(3) 239-246

Putman, S H (1967) "Intra-Urban Industrial Location Model Design and Implementation," *Papers and Proceedings of the Regional Science Association*, 19 199–214.

Putman, S H (1969) "An Economic Model for Forecasting Employment by Industrial Sector in Small Sub Areas of a Large Region," Ph.D dissertation, University of Pittsburgh (Ann Arbor University Microfilms)

Putman, S H. (1970) "Developing and Testing an Intra Regional Model," *Regional Studies*, 4(4) 473--490

Rose, J (1967) "Intraregional Industrial Location Theory and Methodology," paper drafted for the Regional Planning Council, Baltimore

Rose, J (1968) "Interim Industrial Allocation Model Structure and Analysis," mimeo., Regional Planning Council, Baltimore.

Seidman, D. R (n.d) "A Linear Interaction Model for Manufacturing Location Part I," mimeo , Penn Jersey Transportation Study

Seidman, D. R (1966) "A Decision-Oriented Model of Urban Growth," paper presented to Fourth Annual Conference of International Federation of Operations Research Societies

Seidman, D R. (1969) "The Construction of an Urban Growth Model," report no 1, vol A, Philadelphia Delaware Valley Regional Planning Commission.

Smith, W. and Associates (1966) "Maryland Capital Beltway Impact Study, Impact on Industrial Areas," mimeo , Maryland State Roads Commission

Stefaniak, N J (1962) "Industrial Location within the Urban Area," *Wisconsin Commerce Reports*, 6(5).

Stevens B H and R. E. Coughlin (1959) "A Note on Inter-Areal Linear Programming for a Metropolitan Region," *Journal of Regional Science*, 1(2) 75–83.

Stevens, B H and W. Wheaton Jr (1970) "A Review of Available Land Use Activity Distribution Models," TM–4524 (Santa Monica, California System Development Corporation)

Townroe, P. M (1969) "Locational Choice and the Individual Firm," *Regional Studies*, 3(1) 15-24

"Urban Firm Location and the Structure of Urban Land Values," mimeo , author unknown

S. H. PUTMAN

Transportation Planning and Technology
1976, Vol 3, pp 165-173

FURTHER RESULTS FROM THE INTEGRATED TRANSPORTATION AND LAND USE MODEL PACKAGE (ITLUP)

STEPHEN H. PUTMAN

Associate Professor of City and Regional Planning, University of Pennsylvania, Philadelphia, Pa 19174

(*Received July 1, 1975 in final form November 15, 1975*)

This paper has three main sections. The first section is a brief review of the initial development of the Integrated Transportation and Land Use Model Package (ITLUP). This development was the first successful attempt at a full integration of land use and transportation models, incorporating the principal reciprocal relationships between these phenomena. The second section discusses some results of work with ITLUP done since the completion of the original project. The third section suggests several areas in which future development of the model package is contemplated.

I. HISTORY AND DESCRIPTION OF ITLUP

The Integrated Transportation and Land Use Model Package (ITLUP) was developed to investigate the feasibility of balanced development of land use and transportation facilities.[1,2] The essence of this question is the problem of manipulating (i.e. altering, directing, and modifying) the spatial organization of the metropolis. In particular the project focus was on spatial expansion, and the feasibility of attaining a balanced development of land use and transportation facilities, and the means (policies) relevant to achieving that balance.

An Integrated Model Package

The inherent complexity of a comprehensive analysis of both transportation and land use, coupled with the requirement for a self-consistent procedure for testing the sensitivity and response of transportation and land use to policies for their manipulation, determined the need to use simulation models as the method of approach.

Unfortunately, the state-of-art in simulation modelling of urban and regional phenomena was such that, in general, those models which had operational success were rather too highly aggregated both in spatial and sectoral detail as well as in concept. On the other hand, those models which were perhaps more conceptually complete, and which might have provided highly detailed sectoral forecasts, were not operational at a sufficient level of areal detail. Examples of all these types of model are reviewed elsewhere.[3-]

At the same time, it was clear from the outset, that an appropriate model package would need to be even more complete (and therefore complex) than any previously developed. It was required that a complete system, i.e. population and employment sectors, land use model and a complete, i.e. public and private, transportation network model, be assembled. The overall system of models had to be highly integrated with respect to model inputs, and feedbacks. With this integration the model package would attempt to eliminate the principal failing of contemporary land-use or transportation simulation studies by explicitly including the appropriate feedback loops.

Typically, a transportation study assumes a future land use pattern as given, and designs a transportation system to cope with it. This procedure ignores the redistributive effects which are produced by the construction of the system. In other words, transportation systems do not suddenly appear on the ground, but are constructed in stages with consequent redistribution of activities all during the period of construction. In contradistinction, the typical land use study accepts a transportation system as given, and then estimates the consequent distribution of activities. This procedure thus ignores the new travel demand and subsequent congestive effects of the activity distribution on the network. The ITLUP package attempts to remedy these omissions and to capture the interrelatedness of the transportation system and the distribution of activities.

The general structure of the model package first uses the base-year data on the obtaining spatial distributions of activity, along with data on the characteristics of the unloaded base year transportation network. These data are then used to generate a preliminary, and probably inflated, estimate of base year trips in the metropolitan area. This preliminary estimate of trips are then loaded on the future (projection year) network so that its travel characteristics, i.e. time and cost, would, at least, reflect the traffic volumes on the network if there were no change in the spatial distribution of activities from the base year. These network characteristics, along with the base year data and the projection year control totals are then used to generate a spatial distribution of activities for the projection year. From this spatial distribution a new estimate of metropolitan trips is produced. These trips were then loaded on the projection year transportation network. The modified characteristics of the transportation networks are then used to reallocate the projection year spatial distribution of activities. This distribution of activities was then compared to the first estimate thereof. If there were no significant differences an equilibrium would have been reached and the model run ended. If there were significant differences, new trips would be generated and loaded on the networks and further iteration run.

The Operational ITLUP Package

It is very often the case in complex simulation studies that the operational version of the model system is somewhat less comprehensive than the system originally proposed. While some simplifications were necessary in the ITLUP study too, the principal goal of an operational integrated transportation and land use model was attained. The land use model used in the model package was a modified form of the IPLUM model. This model accepts the spatial distribution of "basic" employment as input and then solves for the interrelated spatial distributions of population and "non-basic" or "local-serving" employment. Regional levels of activity are exogenously determined, as is the case with virtually all urban land use models. The "basic" employment location was forecast exogenously by a separate model, BEMOD, which is not included in the model package. Consequently in this preliminary version of ITLUP the spatial distributions of "basic" employment are independent of the remainder of the activities in the model package. This is clearly not a valid assumption over any substantial projection period and is an aspect of the model package which is currently being improved. One possibility would be to integrate the modelling of all employment locations, link this to the residential model, and link this overall model to the network model.

As described above, the principal determinants of spatial distributions of activities in IPLUM are access to workplaces for population, and access to customers for non-basic employment. The access measures are non-linear functions of travel-times weighted by attractiveness measures of the "trip destinations." The actual allocation functions used in the IPLUM model had never been subjected to thorough statistical examination, and early experience with model runs for the ITLUP project indicated that there was too heavy a reliance on somewhat arbitrary "correction-factors." Consequently, a thorough investigation of the model algorithms was made, and they have since been modified. The version of ITLUP whose results are described here used a modified form of IPLUM. Future work with ITLUP will use the entirely new DRAM model.[4] These modifications appear to have improved substantially the model's allocations, though attempts to evaluate these improvements have not yet been completed. Of course, as is the case with all other urban land use models, IPLUM's reliability for specific small area forecasts is less than its reliability for more aggregated regional patterns.

The transportation model package selected for use in the ITLUP package was developed by the Planning Sciences Group of the University of Pennsylvania. It contains programs for network coding, tree tracing, and several alternative capacity restrained traffic assignment procedures. The principal connection going from IPLUM to the Network model is via trip generation. That is, the main phenomenon which relates the spatial distribution of activities back to the transportation facilities is specified in terms of the trips between activities which make use of the transportation system, and which, in the form of congestion, alter that system's operating characteristics.

Two types of trips are considered in ITLUP, worktrips and non-worktrips. In many transportation studies where land-use projection models are being utilized, the worktrips are generated by procedures which are separate from the land use models. These trip patterns, being independently calculated, may not be consistent with the worktrips implicit in the land use model's patterns of residence and workplace. In IPLUM the location of residences is generated by worktrip allocation functions based on transportation impedances and the locations of workplaces. In so doing, IPLUM implicitly generates a matrix of work-to-home trips which, in the ITLUP package, are extracted and used as the worktrips estimates.

Similarly, the location of "non-basic" or "local-serving" employment in the IPLUM model is generated by home-to-shop and work-to-shop allocation functions based on transportation impedance and the location of the activities. This procedure generates an implicit matrix of shopping trips which are extracted, and by a somewhat different procedure involving a gravity model type algorithm for balancing the origins and destination of the trips, are used as the packages's non-worktrips. In the present version of the model package, peak hour worktrips and shopping trips are loaded on the networks. A problem to be resolved for later versions of the model package is that of other non-worktrips besides the shopping trips. Currently shopping trips are "inflated" to equal all non-worktrips.

The first tests of the model package were done for a small prototype urban region of eight areas, while later full-scale runs used San Francisco Bay area data. The current form of the model system utilizes two versions of the PLUM model. The first is essentially the PLUM "state-variables" model where the forecasts are made in terms of the value of variables at time t. The second version used is a modified form of incremental variable IPLUM model, where the forecasts are made in terms of changes in values of variables from time t-1 to time t. These changes are then added to the base

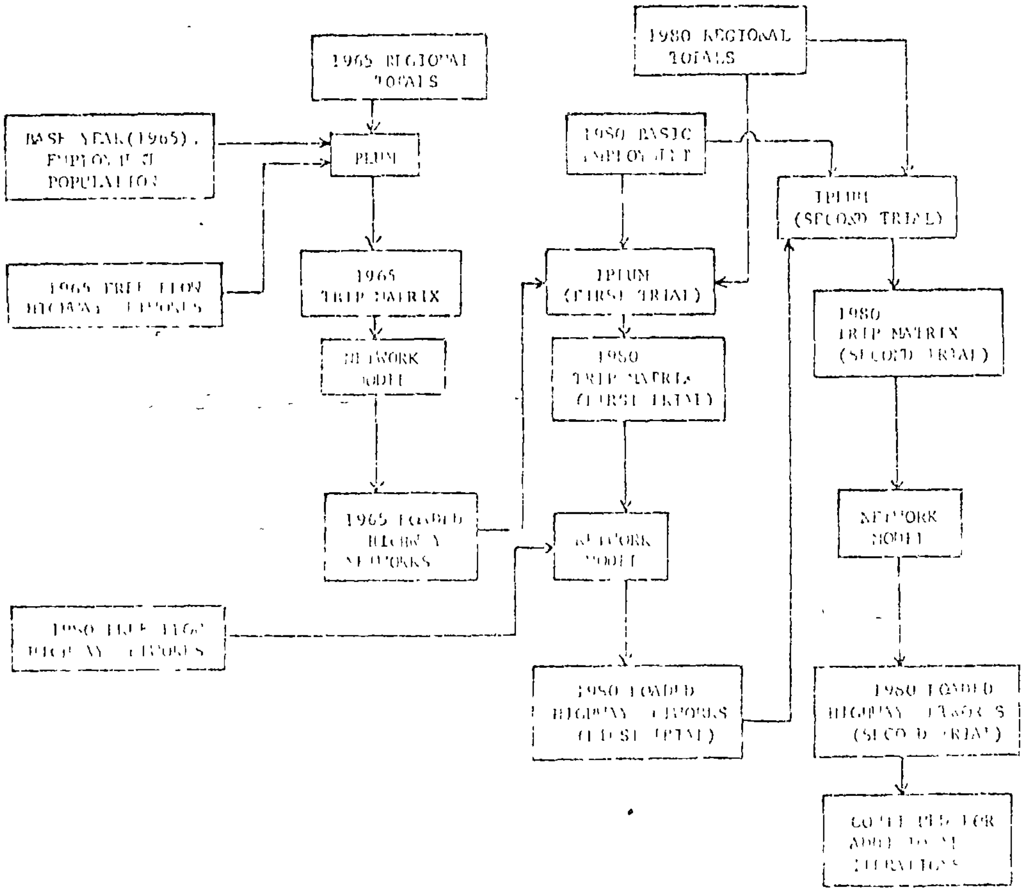


FIGURE 1 Schematic of experimental model runs

values (i e those at time t- 1) to yield the forecast of time t variables

Basically the procedure involves a run of PLUM to generate base-year trip matrices for input, after loading onto a network, to the IPLUM model After this successive iterations of IPLUM and the network model are run The assemblage of models used for these runs is shown in Figure 1

II RESULTS OF RECENT WORK WITH ITLUP

The New Process Runs

During the test runs of ITLUP problems of stability arose that proved difficult to resolve These test runs used data from the San Francisco region and went from a base year of 1965 to a forecast year of 1980 The lack of stability resulted in the model package reaching a condition of relatively stable oscillation after three to five complete iterations of the land use and transportation models After investigating these runs and finding no faults in the operation of the component models of the package, a lagged averaging procedure was developed which resulted in a damping of the oscillations With this modification the model package was able to achieve stable solutions in four to six full iterations Nevertheless it seemed that an alternative solution to this procedure was desirable

It is a classic example of "psychological set" of researchers that it was not until this point in the project, very near its end, that the true cause of the model package's oscillations and their immediately obvious cure, became apparent The problem was the size of the projection interval Attempting to forecast 1980 from a 1965 base resulted in change in spatial distributions and network loadings which were too large to be accommodated smoothly by the integrated models. The ideal test of this hypothesis would have been to make the forecasts in several successive steps of various sizes Time and resources prohibited this systematic approach, so an approximate five year interval was selected The model package would then estimate 1965 1970 1975-1980 Exact regional control totals for 1970 and 1975 were unavailable, therefore the interval from 1965-1980 was simply split into three subintervals The model package was then re-run in three steps, 1965–1970 (est.), 1970 (est) 1975(est), and 1975(est)–1980 (regional projection)

Running the model package in this form, as opposed to the procedure used previously necessitated minor changes in some of the component

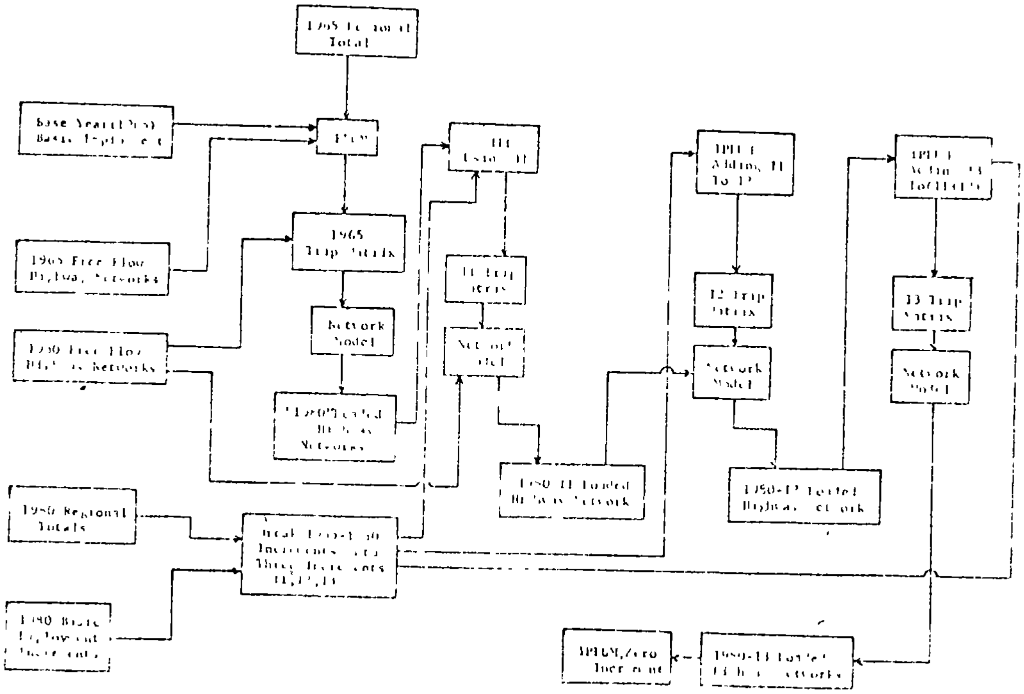


FIGURE 2 Schematic of new process for model package

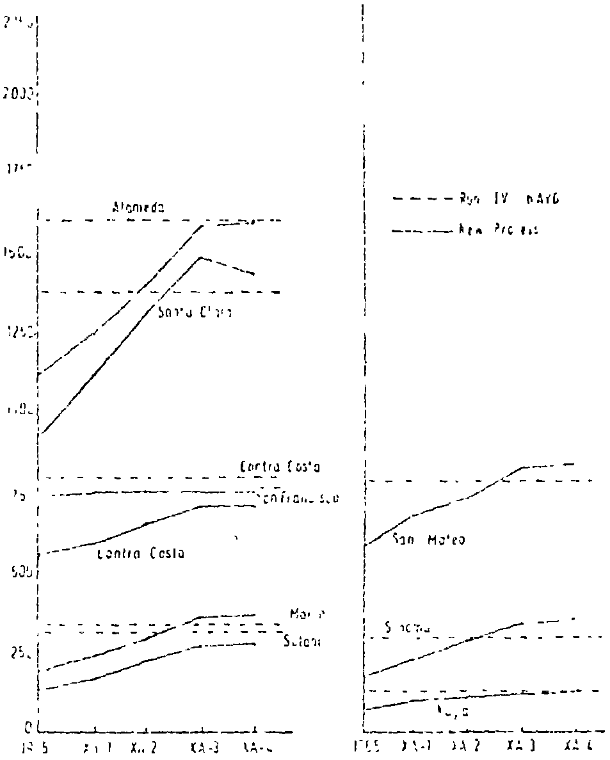


FIGURE 3 New Process, Population Projection for 1980

models and in the executive routines which controlled them. This new arrangement is called the "New Process" and is shown in Figure 2, and described below.

1) Assume that the base year is 1965 and the projection year is 1980 and that there are two equally spaced intermediate time points of 1970 (est.) and 1975 (est.).

2) The base year (1965) trips are loaded on the 1980-network and zone-to-zone impedances are calculated.

3) With these impedances, one-third of the regional control totals, and one-third of the 1965→1980 basic employment increments, as inputs, IPLUM is run and new trip matrices generated for 1970 (est.).

4) The 1965 trips are subtracted from the new trip matrices and the remaining new trips are loaded on the already partly loaded 1980-network. New zone-to zone impedances are calculated.

5) With these impedances, two-thirds of the regional control totals, and an additional one-third of the basic employment increments, as inputs, IPLUM is run and new trip matrices are generated for 1975 (est.).

6) The trip matrices from Step 4 are subtracted from these new trip matrices and the remaining new trips are loaded on the partially loaded 1980-network. New zone-to-zone impedances are calculated.

7) With these impedances, the full regional control totals, and the final one-third of the basic employment increments, as inputs, IPLUM is run and new trip matrices are generated for 1980 (est.).

8) The trip matrices from Step 6 are subtracted from these new trip matrices and the remaining new trips are loaded on the partially loaded 1980-network. New zone-to zone impedances are calculated.

9) With these impedances, the full regional control totals, and zero basic employment increments, IPLUM is run and should then produce an equilibrium solution for the 1980 forecasts.

This new process was run through Step 8. The resulting population projections, where XA−1 = Step 3, XA−2 = Step 5, and XA−3 = Step 7, are shown in Figure 3. The dotted lines in Figure 3 represent the projections from Run IV 6AVG, the equilibrium solution of the original version of the model package. It can be seen that the new process comes rather close after three recursions to where the old process was after six iterations. The fourth recursion of the new process was not run as described above, because of an unexpected operational deficiency in IPLUM. The model is theoretically capable of projecting the response in terms of spatial redistribution of local-serving employment and population, to changes in basic employment and/or transportation facilities. However, it cannot operate with a zero change in basic employment. Since there were not sufficient resources (or time) on this project to correct this deficiency in IPLUM, it was not possible to run the fourth recursion (Step 9) of the new process as originally intended.

As an alternative test, the impedances from Step 8 were used and IPLUM was rerun as in Step 7. This means that the last growth increment was added to the previous two-thirds, but using the impedances derived from the prior estimate of the last growth increment. The results of these two estimates of the spatial distribution of the last growth increment (one based on a two-thirds full network and the other on a three-thirds full network) should, if the new process had found an equilibrium in the first three steps, be nearly identical. In fact, the largest difference, at the county level, was less than five percent.

The results of the population projection made with the "New Process" are shown in Table I. Of particular interest here is the degree to which the results of the third and fourth steps in the process suggest a rather stable estimate of 1980 population. These particular runs were made with the input impedance to the first step being derived from an estimate of 1965 trips loaded on the 1980 network. In order to test the sensitivity of the process to this base point, an alternative set of runs was produced where the input impedance to the first step was derived from the XE 1 trips loaded on the 1980 network. The second step used the trips from the previously run XE 2, etc. These new runs were called XE 1A, XE 2A, and XE 3A. The population increments obtained from these runs are compared to the population increments from the previous run series in Table 2. It will be noticed that there are substantial differences to be found, depending upon the loads on the networks used as input to the land use model.

Various Tests Against the 1970 Census

Another set of tests was done comparing various estimates of 1970 population, derived from alternative model runs, with the 1970 census of population. The first 1970 estimates tested were those produced by the XE 1A run, which were expected to be the best set. The 1970 census data were supplied by the San Francisco Metropolitan Transportation Commission (MTC). The variable used was total population by zone (the 290 zone data file supplied was modified by splitting two zones, to return to the old 291 zone system formerly used by MTC and still used in ITLUP).

The impedances used to produce the XE 1A land use estimates resulted from the loading of estimated 1965 trips plus the estimated 1970 trip increment loaded on the capacity restrained 1980 transport network—1970 network data were not available. The XE 1A estimates, as well as all other IPLUM estimates, include somewhat arbitrary estimates of group quarters population

TABLE I

Results of "New Process" Trial Run—Population Projections

	BASE 1965	XE 1 (1970)	XE 2 (1975)	XE 3 (1980-1)	XE 4 (1980-2)	% Change XE 3→XE 4
San Francisco	754,754	708,981	711,829	712,669	712,790	+0.000
Marin	192,523	240,118	315,297	407,679	406,410	-0.003
San Mateo	557,271	581,884	674,016	720,141	726,317	+0.009
Sonoma	167,497	206,051	258,009	307,402	308,027	+0.002
Napa	69,378	89,853	130,037	185,853	178,564	-0.039
Solano	149,386	204,117	253,412	285,862	291,988	+0.021
Contra Costa	534,056	580,502	695,952	779,948	785,573	0.011
Santa Clara	922,357	1,095,546	1,266,400	1,484,735	1,498,523	+0.009
Alameda	1,095,236	1,121,639	1,287,561	1,472,248	1,445,358	-0.018
Total	4,138,358	4,828,605	5,592,530	6,356,172	6,356,470	0.000

TABLE II

Comparison of Population Increments, XE and XE A Series

	XE 1	XE 1A	%Δ	XE 2	XE 2A	%Δ	XE 3	XE 3A	%Δ
San Francisco	14,909	13,980	0.062	2,849	3,028	-.062	839	950	-.132
Marin	67,388	59,737	0.114	75,178	64,876	0.137	92,383	91,114	0.013
San Mateo	86,858	84,733	0.024	92,132	95,127	-.035	46,125	52,301	-.133
Sonoma	36,619	39,523	-.079	52,061	41,589	0.201	49,302	48,927	-.012
Napa	21,963	23,440	-.067	40,184	39,837	0.005	55,816	48,526	0.130
Solano	57,153	57,150	0.000	49,295	46,293	0.060	32,450	38,577	-.188
Contra Costa	121,217	113,080	0.067	115,150	110,150	0.045	83,996	92,621	-.102
Santa Clara	208,472	221,978	-.065	170,856	190,991	-.117	218,311	232,128	-.063
Alameda	149,238	150,195	-.006	165,921	171,736	-.035	181,685	157,794	0.131
Total	763,891	763,818	--	763,926	763,925	--	763,936	763,936	-

Further, the estimates of 1970 basic employment which were input to all of these runs of IPLUM were produced by BEMOD (the existing basic employment model used by MTC whose outputs were used throughout the ITLUP project). Further, actual 1970 BEMOD estimates do not exist, but were produced by simply taking one third of the estimated 1965 to 1980 incremental employment estimates. Finally, the 1970 regional population control totals for IPLUM did not exactly match the 1970 census totals because the IPLUM runs were done before the Census data were available. It was decided not to spend the computer time to rerun the model runs. Instead, the model runs were simply scaled afterwards to the Census totals (called XE 1s, and XE 1As).

The coefficient of determination r^2 between the 1970 Census population data and the scaled XE 1A estimates of 1970 population was 0.8788. The population distributions, summarized to the county level, were as shown in Table III.

The difference between XE 1 and XE 1A was that the impedances used in XE 1 resulted from only estimated 1965 trips loaded on the capacity restrained 1980 network. There was only a minor difference in the coefficient of determination (r^2) between XE 1 and 1970 Census and the r^2 between XE 1A and 1970 Census. This small difference would undoubtably have been larger if 1970 capacitated networks had been available for use instead of the 1980 networks. This r^2 was 0.8774, and the comparable county level statistics are shown in Figure IV. At the county level the differences between the two runs, as compared to the Census, are not very great.

A third comparison run was made, XE 1C, using the estimates of 1965 "peak" times prepared by MTC as the inputs to the IPLUM 1970 population estimates. In this case the correlation (r^2) between the IPLUM estimates and the 1970 Census dropped to 0.8601, the worst of the lot.

The IPLUM model was modified during the ITLUP project to include a correction factor for the effect of available vacant land on residential location. This correction factor was intended as a surrogate measure of "developability" of land in a given zone. All of the above runs of IPLUM were made with this correction included. A run was

TABLE III

Population distributions summarized to county level for XE 1A

	XE 1A	1970 Cens	XE 1As	%Δ −XE 1As
San Francisco	708,052	715,674	678,667	−.052
Marin	232,368	206,038	222,724	0.081
San Mateo	579,759	556,231	555,699	−.001
Sonoma	208,943	204,885	200,271	−.023
Napa	91,330	79,140	87,539	0.106
Solano	203,114	169,941	195,643	0.151
Contra Costa	572,335	558,389	548,583	−.018
Santa Clara	1,109,102	1,061,714	1,063,074	−.002
Alameda	1,122,609	1,073,184	1,076,012	0.003
Total	4,828,525	4,628,184	4,628,212	—

TABLE IV

Population distributions summarized to county level for XE.1A

	XE 1	1970 Cens	XE 1s	%Δ −XE 1s
San Francisco	708,981	715,674	679,553	−.050
Marin	240,118	206,038	230,151	0.117
San Mateo	581,884	556,231	557,732	0.003
Sonoma	206,039	204,885	197,497	−.036
Napa	89,853	79,140	86,124	0.088
Solano	203,117	169,941	195,635	0.151
Contra Costa	580,502	558,9[illegible]	556,407	−.005
Santa Clara	1,095,5[illegible]	1,061,714	1,050,073	−.013
Alameda	1,121,639	1,073,184	1,075,083	0.002
Total	4,828,663	4,628,184	4,628,265	

made, analogous to XI.1C, but with the correction factor deleted. For this run, called XI.1D the r^2 with 1970 Census dropped even further to 0.8176.

Several preliminary conclusions may be drawn from these runs. First, the use of a correction factor which, even in a crude way, takes some account of the apparent developability of vacant land, produces considerable improvement in the model's estimates. Second, the use of networks congested by the model's estimates of trips produces better land use projections than were obtained by the use of networks whose characteristics were estimated by XI.1C independent of the land use estimates. Third, at the county level the differences between partial loadings of the networks by the model packages do not show up very dramatically. There are, however, sharp differences at the zone level which suggest that careful further investigation of these interrelationships is desirable.

The Development of DRAM:

A most important recent development with the ITLUP package has been the construction and preliminary testing of a Disaggregated Residential Allocation Model—DRAM. This model was a by-product of an attempt to find a means of doing a true statistical calibration of IPLUM.[4]

During this attempt it was discovered that, with one exception, no Lowry derivative model had ever been statistically calibrated in U.S. practice. In contradistinction it was also discovered that by recasting these models in the Wilson entropy-maximizing formulation, the British had made their proper calibration a matter of course in modelling projects.[5] Consequently IPLUM was so recast, and at the same time modified to a somewhat different form and its single household type disaggregated to four household types. This reformulated model has been named DRAM and calibrated on both San Francisco and Minneapolis data sets. The results of this calibration work have yielded results which suggests that each of the four household types (defined by income class) is more accurately forecast than was total population in IPLUM.

Conclusions from these Further Efforts with ITLUP:

Two main conclusions are to be drawn from this work. First, it is both possible and necessary to do land use and transportation forecasting with an integrated set of models, as the feedback relationships which are otherwise ignored can have significant effects. Second, having assembled such a package, despite both known and unknown problems with its component models, the need for further research efforts becomes more clearly directed and further, it can be conducted while the previous version(s) of the package remain available for policy analysis. The next section of this paper will outline a number of areas in which further research is currently in progress.

III. NEXT STEPS WITH ITLUP

In the Summer of 1975 work began on further modification and improvements to ITLUP, sponsored by a grant from the U.S. National Science Foundation. The research portions of this work will focus on 1) the integration of existing models into ITLUP for the purpose of extending and or improving its policy analysis capabilities, and 2) the development of models to fill gaps in the anticipated, more sophisticated, new version of ITLUP.

One of the key modifications to be made to the existing package is the substitution of the newly developed DRAM residential model for the modified IPLUM currently being used. Another important modification will be to add in a procedure for estimating basic employment. The third major modification will be to add in a mode split capability. Finally, it is hoped that it will be possible to include a modest set of procedures for investigating some of the environmental impacts and energy consequences of alternative urban development patterns.

The ITLUP package, even in the preliminary form which produced the results described here, offers a substantial increase in policy analysis capability over conventional land use or transportation simulation models. Another portion of the work to be done with ITLUP involves its preparation in a form which may be readily distributed to any organization desiring it.

Prior to the development of ITLUP, the land use models and the transportation models were always run separately. This practice, by definition, produced only partial answers to any particular policy investigation. It was precisely this fact that led to the initiation of the development of the Integrated Package several years ago. That this was a timely decision becomes more clear with each increase in energy conservation, emission controls, land use controls, center city vehicular traffic restrictions, and other policies and implications of alternative urban system configurations.

Integrated transportation and use planning has long been discussed as a desirable procedure, but the technology for its application in practice did not exist. Initial work with the preliminary version of ITLUP clearly demonstrated that systematic errors would be made in estimating land use and transportation consequences if only partial analyses were conducted. The further work described here suggests that integrated models will produce better forecasts than partial models. The availability of the improved version of ITLUP proposed here will greatly enhance planners' ability to perform more complete analysis. Interest of planners in ITLUP, even in its preliminary form, has been considerable. Further improvements, along with a considered effort to make its use less difficult should result in its rather widespread use. If so, the benefits to be had from the improved planning analyses which it would make possible would be considerable.

REFERENCES

1. S. H. Putman, *et al.* (1973) "The interrelationships of transportation development and land development," Report on Contract DOT-FH-11-7813, U.S. Department of Transportation, Federal Highway Administration, Urban Planning Division, Urban Development Branch.
2. S. H. Putman, (1974) "Preliminary results from an integrated transportation and land use models package," *Transportation*, Vol. 3, pp. 193-224.
3. S. H. Putman, (1975a) "Urban land use and transportation models: a state of the art summary," *Transportation Research*, Vol. 9, pp. 187-202.
4. S. H. Putman, (1975b) "Calibrating a disaggregated residential allocation model DRAM: the residential submodel of the integrated transportation and land use package - ITLUP." Presented to the British Regional Science Association, London, September 4-5, 1975.
5. M. Batty and S. Mackie (1972) "The calibration of gravity, entropy and related models of spatial interaction," *Environment and Planning*, Vol. 4, pp. 205-233.

Transportation 3 [illegible]) 193 - 224
© Elsevier Scientific Publishing Company, Amsterdam – Printed in the Netherlands

PRELIMINARY RESULTS FROM AN INTEGRATED TRANSPORTATION AND LAND USE MODELS PACKAGE

STEPHEN H. PUTMAN

Associate Professor, Department of City and Regional Planning, University of Pennsylvania

ABSTRACT

This paper describes the results, to date, of an effort to integrate a land use model with a transportation network model for the purpose of analyzing the interrelationships of transportation facility development and land development. In the system which has been developed each model provides input to, and receives feedbacks from, each other model. To the author's knowledge, the effort described here represents the first successful attempt to develop and test an integrated model package involving these reciprocal relationships. The results obtained from preliminary runs of this package should be of considerable interest to both transportation planners and land use planners. With this integrated system it has been possible to observe the interrelationships, and in particular the feedbacks, between land use and levels of traffic on the networks. Preliminary results indicate that congested networks produce tendencies toward metropolitan centralization. Attempts to relieve congestion seem to produce metropolitan decentralization and increased travel which lead, in turn, to metropolitan sprawl and increased spread of congestion.

Introduction

It is common knowledge that in the past two decades there has been enormous public investment in highway construction. This same period has also witnessed large scale shifts in population from rural areas to urban areas and within the urban area, from center-city to suburb. Even without attempting here to define the explicit causal relationships between these two developments, it is possible to describe a related phenomenon which is the focus of the research described here. This phenomenon is observed, not infrequently, when the construction of a new section of roadway is followed, all too soon, by very heavy usage and subsequent congestion. Specifically, the nature of this process seems to be that: a) Due to the inadequacy of existing facilities a decision is taken in a metropolitan area to improve transport facilities (usually by road building) for a particular part of the area. Then b) assuming that the decision is approved, in anticipation of the new roadway, land developers and/or speculators become involved with properties near the proposed route. As construction of the facility begins, some homeowners and businesses may consider and even act upon location or relocation decisions. (We do not refer here to relocation forced by the facility construction.) Upon completion of the facility, additional location decisions are made. Finally c) a relatively short time after completion of the facility, the demand for its use greatly exceeds the demand which existed prior to the decision to construct it. Consequently the design capacities are soon reached and often exceeded, resulting in congestion and premature obsolescence of the facility. While this is a rather generalized description of a very complex process, it is reasonably accurate. The principal question addressed by this research is whether it is feasible, via integrated transportation planning and land use planning, to avoid or ameliorate the occurrence of this particular phenomenon in the future. Further it was the intention of the research described herein to analyze this process and to determine whether a) balanced development of transportation facilities and land use is feasible, and b) if it is feasible, what means are available to accomplish it.

There is considerable evidence indicating that the demand for highway travel is rather sensitive to changes in highway capacity. This sensitivity, as described above, frequently results in heavy utilization and congestion of new facilities soon after their construction. It might be argued that the solution to this problem is to construct more facilities. It is possible that at some point, if this policy were followed, an equilibrium situation would indeed be reached. This conclusion can be supported by asserting that the elasticity of demand for highway travel is finite. However, if the "population" generating that demand continually increases at the same time, it is not clear that the total demand can easily be satisfied in this manner. The limit of this strategy, at the extreme, would be reached when so much land is converted to roads that land development for other purposes is restricted, with a consequent limit on trip generation and road use. It is obvious that this "equilibrium" is not the desired solution and is hopefully not feasible in any case. It will therefore be necessary to analyze possible "intermediate" solutions.

Any potentially feasible solutions discovered by these analyses will require evaluation. This evaluation, leading to a definition of feasibility, would undoubtedly involve measures of private and social costs, and measures of the disparities between them. Such costs could include user's costs (e.g. operating costs, taxes, or tolls), pollution (of all types), costs of relocation of activity, societal costs of activity dispersion, and costs of disruptions caused by the construction process, all of which are associated with the various levels of transportation facility development Further costs, associated with land development include environmental pollution, loss of open space, and congestion of various public facilities (which results either in deterioration of service levels of costs of expansion), in addition to the congestion of the transport facilities, as described above. Such evaluations were not undertaken as part of this project.

A balance between transportation facility development and land development implies a market equilibrium of the demand for use of the transportation facility and its supply, i e its speed and capacity characteristics There are two basic alternatives available to the planner who wishes to modify an existing transportation and land use situation sufficiently to achieve such a balance, though it is likely that the best strategies will be mixtures of the two alternatives The first alternative is to *allow demand for transportation to fluctuate freely,* with no interference, and attempt to cope with it. This would be accomplished, as in the past, by new highway construction or by implementation of mass transportation systems. The second alternative is to *attempt to restrict the demand for transportation* so that facilities do not become overloaded. This could be accomplished in three ways: 1) the existing transportation facility could be made more costly to use, e g. by the imposition of tolls or by allowing congestion to develop, thus imposing a time penalty on users; 2) land development controls could be imposed, thus reducing (or slowing the growth of) trip generating activities, and 3) a mixture of these two actions could be implemented. Finally, a mixture of the two basic alternatives could be attempted, where an attempt would be made to cope with a certain amount of transportation demand (by improving transportation facilities) at the same time that an attempt was being made to control its increase.

To summarize, the essence of the problem which this research addressed is that of controlling (i e altering, directing, and modifying) the spatial organization of the metropolis In particular, the concern was with its spatial expansion and with the feasibility of balanced development of land use and transportation facilities. Finally, when such balanced development appears feasible, potential methods for testing the policies which may be relevant to achieving that balance are explored.

Method of Approach

Two aspects of the problem were the principal determinants of the method of approach. The first was the inherent complexity of a comprehensive analysis of both transportation and land use. The second was the requirement for a self-consistent procedure for testing the sensitivity and response of transportation and land use to policies for their manipulation. The only analysis method with any chance of meeting these requirements was that of computer simulation modelling The remainder of this paper is devoted to describing the package of models which has been developed and tested as the central focus of this research.

The state-of the-art in simulation modelling of urban and regional phenomena poses a problem for projects of this sort In general, those models which have had operational success are, at present, rather too highly aggregated both in areal and sectoral detail as well as in concept (Putman, 1974) Two such models are EMPIRIC (Hill, 1965) and PLUM (Goldner, 1968). On the other hand, those models which are more conceptually complete and which could provide highly detailed forecasts. are not yet operational. An example of one such model is the Herbert-Stevens model as developed for the Penn-Jersey Transportation Study and subsequently modified at the University of Pennsylvania (Harris, et al , 1966). After some consideration of the project goals, it was decided to begin with an existing operational land use model and attempt to modify and improve it concurrently with efforts to integrate it with a suitable transportation model. The project requirements suggested, from the outset, that an appropriate model package would need to be far more complete (and therefore complex) than any now extant. What was ulti-

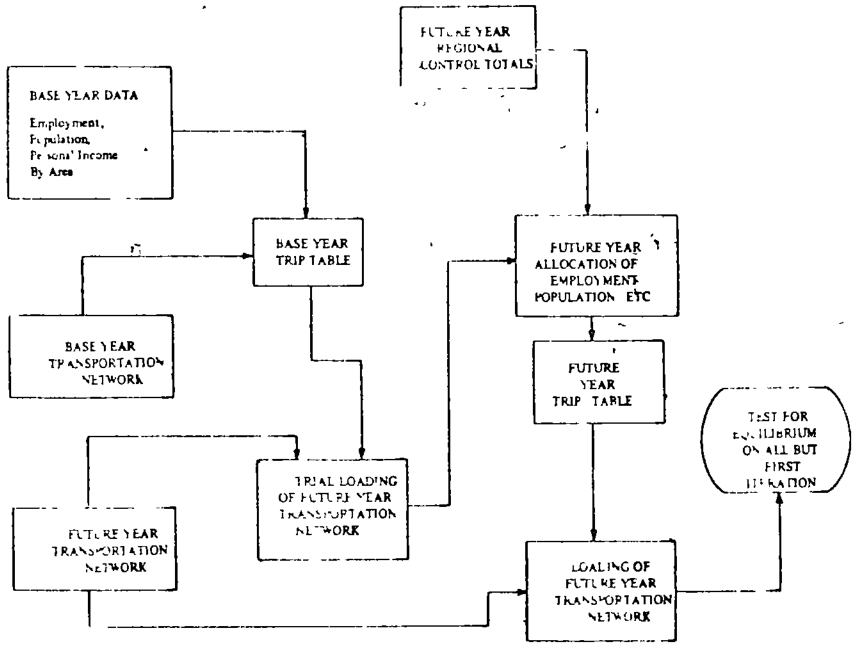


Fig 1. General Scheme of Integrated Model Package.

q[illegible] were a complete – i e population and employment sectors – land use model and a con[illegible]ete – i e public and private – transportation network model. Further, the overall system of models would need to be highly integrated with respect to model inputs, outputs, and feedback loops. In Figure 1 a block diagram of the desired system is presented.

As shown in Figure 1, the process would begin with the various base-year inputs dealing with spatial distributions of activity, along with data on the characteristics of the unloaded base year transportation network. These data would then be used to generate a preliminary, and probably inflated, estimate of trips taking place in the metropolitan area. Given this preliminary estimate of metropolitan area trips, it would become possible to load the future (projection year) network so that its travel characteristics i e time and cost, reflect the traffic volumes which would be on the network if there were no change in the spatial distribution of activities from the base year. These network characteristics, along with the base year data and the projection year control totals, would then be used to generate a spatial distribution of activities for the projection year. From this spatial distribution a new estimate of metropolitan trips would be produced. These estimated trips would, in turn, be loaded on the projection year transportation network. The modified characteristics of the transportation networks would be used to reallocate the projection year spatial distribution of activities. This distribution of activities would then be compared to the first estimate thereof. If there were no significant differences, an equilibrium would have been reached and the model run would end. If there were significant differences, new trips would be generated and loaded on the networks and further iterations made.

This proposed package of models attempts to eliminate the principal failing of contemporary land use or transportation studies, by explicitly including feedback loops. Typically, a transportation study assumes a future land use pattern as given, and designs a transportation system to cope with it. This procedure ignores the redistributive effects which are produced by the construction of the system. Transportation systems obviously do not just suddenly appear but are constructed in stages with consequent redistribution of activities all during the period of construction. The typical land use study accepts a transportation system as given, and then estimates the consequent distribution of activities, while ignoring the congestive effects of that distribution on the network. The proposed model package, as described above, attempts to capture the interrelatedness of the transportation system and the distribution of activities. This comprehensiveness is clearly necessary to meet the research objectives of the project regarding the feasibility of balance between transportation facilities and patterns of land use, as well as the testing of policies to achieve such a balance.

The Operational Model Package

It is very often the case in complex simulation studies that the operational version of the model system is somewhat less comprehensive than the system originally proposed. While some simplifications had to be accepted in this study, the principal goal of an integrated transportation model and land use model was attained.

The land use model selected for this first operational model package was, despite certain apparent (and some less apparent) shortcomings, the Projective Land Use Model – PLUM (Goldner, 1968). The PLUM model (actually there are several versions of the model, two of which are used, with modifications, in this model) is in essence a derivative of the LOWRY model (Lowry, 1964). This class of land use models is macro-descriptive in construct, asserting that the spatial distributions of residences and resident-serving employment are functions of the location of "basic" employment, transport facility characteristics, and the trip taking propensities of residents. This means that the model accepts as given the spatial distribution of "basic" employment and then solves for the interrelated spatial distributions of population and "non-basic" or "local-serving" employment. The principal determinants of these spatial distributions are access to workplaces for population, and access to customers for non basic employment. The access measures are non-linear functions of travel-times weighted by attractiveness measures of the "trip destinations."

As a result of this decision, regional levels of activity must be exogenously determined. This is not a particularly unusual situation, as virtually all urban land-use models have this property. Further, the selection of this model implies that "basic" employment location will be forecast exogenously. These forecasts do, in fact, come from a separate model, BEMOD, a basic employment model developed as part of the Bay Area Transportation Study (Nathanson, 1970). An implicit assumption of this procedure is that the spatial distributions of these activities are independent of the remainder of the activities in the model package, i.e. population, "non-basic" employment, and transportation networks. This assumption is clearly not valid over any substantial projection period and is an aspect of the model package which should be improved as soon as possible. One possibility would be to integrate the modelling of all employment locations, link this to the residential model, and link this overall model to the network model, Finally, a further consequence of the selection of the PLUM model is that the level of sectoral aggregation in the PLUM model is rather more gross than is desirable Fortunately this was tolerable in the preliminary development of this package.

It should also be noted that the actual allocation functions used in the PLUM models have never been subjected to thorough statistical examination, and this project's early experiences with runs of this model indicated that there was too heavy a reliance on somewhat arbitrary "correction-factors." Consequently, a thorough investigation of the model algorithms was made. Certain anomalies were found in the procedures used for activity allocation, which have since been modified These modifications appear to have substantially improved the model's allocations, though attempts to evaluate these improvements have not yet been completed.

On the positive side, the PLUM model does seem to produce reasonable estimates of the general distribution of activities in a large urban region. Of course, as is the case with all other urban land use models, its reliability for specific small area forecasts is considerably less than that for general regional patterns. Finally, and not an unimportant part of the decision, was the fact that the PLUM model was available and operational in the "Models Library" of the Planning Sciences Simulation Laboratory at the University of Pennsylvania prior to the initiation of the research.

The transportation network model package selected for use in this project was developed by the Planning Sciences Group at the University of Pennsylvania (Kuner, 1973a). This package contains programs for network coding, tree tracing, and several alternative capacitated traffic assignment procedures This network package, as compared to others which were available such as the FHWA (formerly known as the BPR) package (U S. D O T., 1972), lacks the ability to handle highly detailed network specifications such as turn-lane penalties, etc. This deficiency was not relevant to the needs of the project, in that there was no intention to conduct any investigations at that level of detail. On the other hand, the computer operation of the Planning Sciences network package is significantly more efficient than that of the FHWA package.

In many transportation studies where land use projection models are being utilized, the work-trips are generated by separate procedures These trip matrices, independently calculated, may not be consistent with the work-trips implicit in the land use model's patterns of residence and workplace. In the PLUM model, the location of residences is generated by work-trip allocation functions based on transportation impedances (derived from the transportation network characteristics) and the locations of workplaces In so doing, PLUM implicitly generates a matrix of work-to-home trips The work-trips used in this project's model package are extracted directly from PLUM's implicit work-trip matrix and thus avoid the above mentioned inconsistency It was, of course, necessary to make certain adjustments to this set of work-trips in order to assure that their total equaled the totals indicated for the region These adjustments include ratios of auto-to-transit person trips, auto occupancy rates, percentages of all work trips which occur during peak hours, etc.

The location of "non-basic" or "local-serving" employment in the PLUM model is generated by home-to-shop allocation functions based on transportation impedances and the locations of the activities As was the case for work-trips, this procedure generates an implicit matrix of shopping trips. These trips too, are extracted from PLUM (by a somewhat different procedure involving the use of a gravity model type algorithm for balancing the origins and destinations of the trips). Again, adjustments are necessary to insure that the projected trip totals equal the totals for the region

The principal connection going from the PLUM model to the Network model is via trip generation. That is, the main phenomenon which relates the spatial distribution of activities to the transportation facilities is specified in terms of the trips between activities, which make use of the transportation system. In the present version of the model package, peak-hour work trips and shopping trips are loaded on the networks A problem to be resolved for later versions of the model package is that of other non-work-trips besides the shopping trips Additionally, the problem of mass transportation facilities has, for the present, been ignored There is, however, no reason to exclude such trips from the model package in future efforts. Their inclusion would involve construction of a modal-split model and a procedure for dealing with multi-modal networks in the PLUM allocation functions Neither of these appears to present an insurmountable problem for future efforts.

Several alternative procedures for loading the trips on the network have been tested and are described elsewhere (Kuner, 1973b) The procedure used, called "incremental tree-by-tree loading method," is a modification of one developed at the Chicago Area Transportation Study. Beginning with a free-flow network, a single tree is traced (i e. all the minimum paths from that origin to all destinations). Onto these paths a portion of the trips from that origin to the various destinations is loaded By use of a volume-capacity relationship, the new volumes on the links of the tree are used to calculate new travel times. Then a new origin is selected, a new tree is traced, and the process is repeated When a tree has been traced for each origin, the network has been loaded with a portion of all the trips. This overall procedure is then repeated. Currently, the networks are loaded in three passes of 40%, 40%, and 20% of the trips respectively. After the network loading procedures have been completed, a matrix of travel times from each zone to each other zone, via the minimum path on the loaded network, is calculated for input to the land use model.

Preliminary Tests of the Model Package Prototypes, Rounds I and II

The first tests of the model package were done for a small prototype urban region of eight areas. After ironing out several minor problems, acceptable results were obtained and preparation began for full-scale model runs using San Francisco Bay area data. The San Francisco area was selected as a test case for two reasons. First, this is the area for which the PLUM model was originally developed and tested. Second, substantial data files for this region were readily available to the project. This was by no means a small undertaking; however, the results currently being obtained indicate that it is probably worthwhile. These results as well as the precise form of the model system now in use are described below.

It must first be noted that there are two forms of the PLUM model. The first of these is essentially a "state-variables" model in which the forecasts are made in terms of the values of variables at time t (Goldner, 1968).

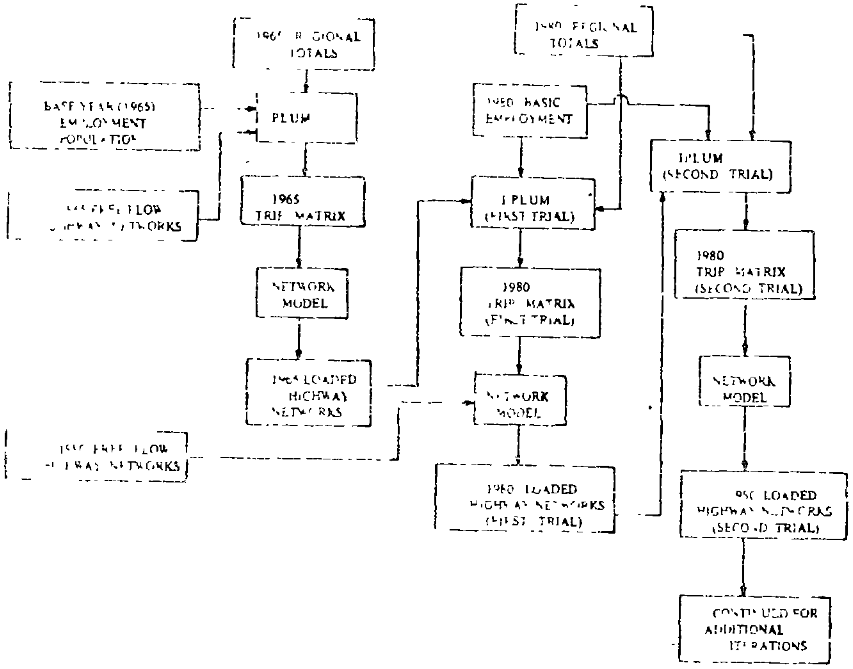


Fig. 2. Schematic of Experimental Model Runs.

The second form of the model (referred to hereafter as IPLUM) is an incremental variable version in which the forecasts are made in terms of changes in values of variables from time $t-1$ to time t (Goldner, 1972). These changes are then added to the base values (i.e. those at time $t-1$) to yield the forecasts of time t variables.

The assemblage of models used for these runs is shown in Figure 2. Basically the procedure involves a run of PLUM to generate base-year trip matrices for input, after loading onto a network, to the IPLUM model. After this, successive iterations of IPLUM and the network model are run.

Before presenting the results of these runs, it is appropriate to consider what expectations one would have as to the results of such a test run. If, as suggested earlier in this paper, important interrelationships are ignored by the typical transportation study or typical land use study, both of which assume fixed values for the other, then to the extent that these models include those interrelationships, they should indicate significant and systematic differences in output from one iteration to the next. In particular, in terms of the degree to which accessibility shapes metropolitan areas, one would hypothesize that:

a) If land use forecasts are based on a free-flow network or a network where flows did not affect travel time, congestion and therefore travel time are substantially underestimated, consequently the overall spatial distribution of activities will be more dispersed than is realistic.

b) The consequent excessively dispersed pattern of activities will, in turn, produce overestimates of tripmaking and, if loaded on a capacity restrained network, of network congestion.

c) If land use forecasts are based on overcongested networks, with consequent overestimates of travel time, the overall spatial distribution of activities will be excessively concentrated.

d) An excessively concentrated pattern of activities will produce underestimates of tripmaking, the consequent travel time and network congestion.

Given these expectations, it is to be anticipated that a linked land use and transportation model would oscillate between spatial decentralization and centralization and consequent network overcongestion and "undercongestion." The results of the model runs bore out these expectations. The policy implications of these results are quite important and are discussed, after the discussion of the model runs, in the concluding section of this paper.

The first "Round" of full model runs was done with a version of IPLUM not incorporating our above mentioned modified allocation function, and using only "work-trips" for network loading. The population

TABLE I

Round I Base Population, Projected Total Population for 1980, and Percent Change 1965–1980

	1965	1980–Run I.1		1980–Run I.2		1980–Run I.3		1980–Run I.4	
San Francisco	754,754	764,445	1.3%	763,414	1.1%	765,068	1.4%	765,067	1.4%
Marin	192,523	469,669	243.9	319,634	66.0	301,379	56.5	259,124	36.6
San Mateo	523,271	801,723	53.2	756,041	44.5	835,885	59.7	938,460	79.3
Sonoma	167,497	368,057	119.7	591,781	253.3	444,212	165.2	377,406	125.3
Napa	69,378	135,910	95.9	183,943	165.1	156,245	125.2	130,538	88.2
Solano	149,386	363,338	143.2	437,717	193.0	316,304	111.7	276,112	84.8
Contra Costa	534,056	666,716	24.8	720,394	34.9	778,010	45.7	672,925	26.0
Santa Clara	922,357	1,308,470	41.9	1,238,821	34.3	1,341,138	45.4	1,499,215	62.5
Alameda	1,095,236	1,528,722	39.6	1,395,805	27.4	1,469,301	34.2	1,488,690	35.9
TOTAL	4,438,458	6,407,473	44.4	6,407,471	44.4	6,407,470	44.4	6,407,474	44.4

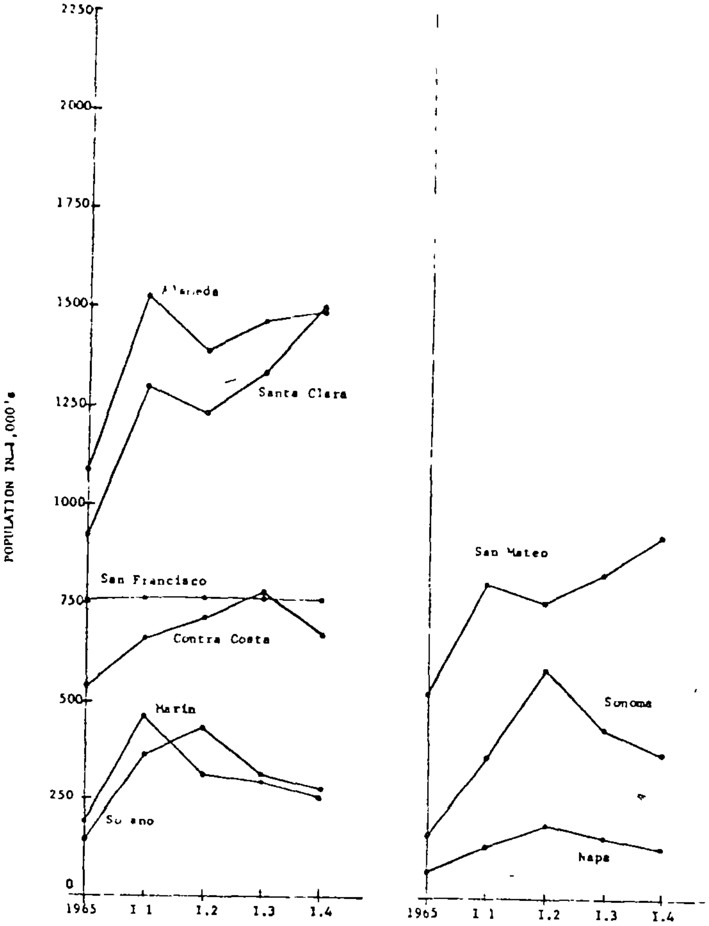
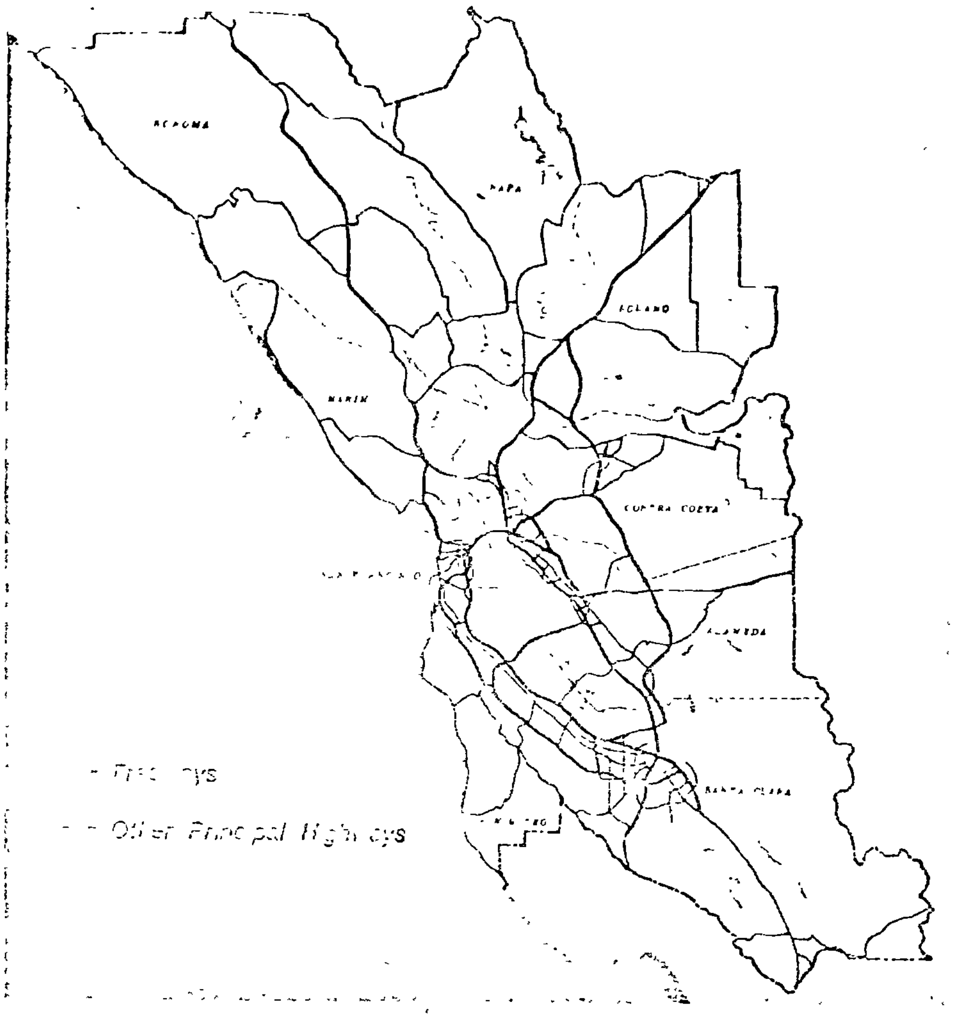


Fig 3 Round I – Population Projections by County.

estimates produced in this round are shown in Table I and Figure 3. An immediate and important conclusion which can be drawn from these results is that there is a substantial difference in the spatial distribution of population as a function of whether the impedance inputs are derived from loaded or empty (unloaded) networks. Further, it is clear that there are oscillations in the population allocations as a function of levels of congestion on the loaded networks. These conclusions raise serious questions as to the adequacy of either a transportation or a land use forecasting method which ignores the relationships between these two components of the problem.

The San Francisco data base used in these tests has two levels of spatial detail. The zone level consists of the 291 analysis zones defined by the Bay Area Transportation Study. These zones may be aggregated to the nine counties in the region. At the zonal level of detail we found the



Map. 1. The Nine County Region and Highway Network.

general, expected pattern of uncongested (underutilized?) networks producing spread-out population allocation, which produced congested networks, which then produced more centralized population allocation, which in turn produced less congested networks, etc. At the county level of detail, however, much of this behavior was obscured. See Map 1 for reference to the location of the various counties. The general performance of the model package was considered to have met our expectations. In view, however, of the obvious "holes" in the package, further, more detailed analysis of these first outputs was curtailed in favor of work on the needed improvements in the package.

There were three principal areas of further investigation which ultimately led to the second Round of model runs First, as mentioned above, preliminary investigations of the operation of IPLUM gave evidence of serious inconsistencies in the allocation functions The conclusion of these investigations was that the IPLUM allocation procedure required important modifications These were implemented and resulted in an apparent improvement in the model's performance (Putman, et al., 1973) Second, it was clear that an important proportion of total trips was being ignored by dealing only with work-trips. A procedure for making explicit the home-to-shop and work-to-shop trips implicit in IPLUM was developed, along with a gravity model for distributing these trips. The number of shopping trips was then scaled up to equal the total non-work-trips in the region These, plus the work-trips, it was felt, would be a reasonably good estimate of the total trips on the network. Third, and finally, a procedure was developed to effect a substantial savings in computer time for the network package In general this procedure involved the assumption that, after the first two iterations (runs) of IPLUM in a given Round, a large proportion of the consequent trips on the network would be relatively stable. Consequently a marginal loading of trips would be a possible substitute for the full network loading runs used in Round 1.

When these investigations and modifications were completed, it was possible to proceed with Round II of the model runs. The results of four iterations of the modified IPLUM, trip generation, and the modified network package incorporating the marginal loading procedure are shown in Table II and Figure 4. These runs were later given the suffix B and are referred to in that way. It is immediately obvious from Figure 4 that the problems of oscillation, from one iteration of IPLUM to the next, were still present. While this was not a particularly unlikely outcome, it was definitely not a welcome result.

A first attempt to isolate the source of these oscillations involved reverting to the full network loading procedure. Since the marginal procedure had been instituted only after run II.2 (to prepare input for run II 3B), it was not necessary to repeat runs II.1 and II 2 of IPLUM. Three complete iterations of the network package, IPLUM, and trip generation were run, yielding runs II 3A, II.4A, and II.5A. These results indicated that the oscillations were still present in the system and, consequently, that the marginal network loading procedure was not at fault.

A more detailed investigation of these outputs revealed results that had been obscured at the county level of detail. The oscillations observed were not really county level oscillations, but zonal oscillations In each county the bulk of the zones remained rather stable while a few of the zones underwent rather large scale oscillations. For example, in Alameda

TABLE II

Round IIB Base Population, Projected Total Population for 1980, and Percent Change 1965–1980

	1965	1980–Run II.1		1980–Run II 2		1980–Run II 3B		1980–Run II 4B	
San Francisco	754,754	759,821	0.7%	762,455	1.0%	758,553	0.5%	760,438	0.8%
Marin	192,523	313,982	63.1	251,919	30.9	341,856	77.6	223,377	16.0
San Mateo	553,271	840,265	51.9	719,364	30.0	787,382	42.3	700,500	26.6
Sonoma	167,497	440,337	162.9	255,273	52.4	708,799	323.2	244,780	46.1
Napa	69,378	118,582	70.9	162,268	133.9	190,562	174.7	93,982	35.5
Solano	149,386	223,074	49.3	351,312	135.2	288,630	93.2	229,551	53.7
Contra Costa	534,056	655,638	22.8	771,037	44.4	635,465	19.0	783,855	46.8
Santa Clara	922,357	1,412,839	53.2	1,348,370	46.2	1,239,020	34.3	1,429,129	54.9
Alameda	1,095,236	1,643,003	50.0	1,364,748	24.6	1,457,279	33.1	1,941,722	77.3
TOTAL	4,438,458	6,407,472	44.4	6,407,476	44.4	6,407,468	44.4	6,407,474	44.4

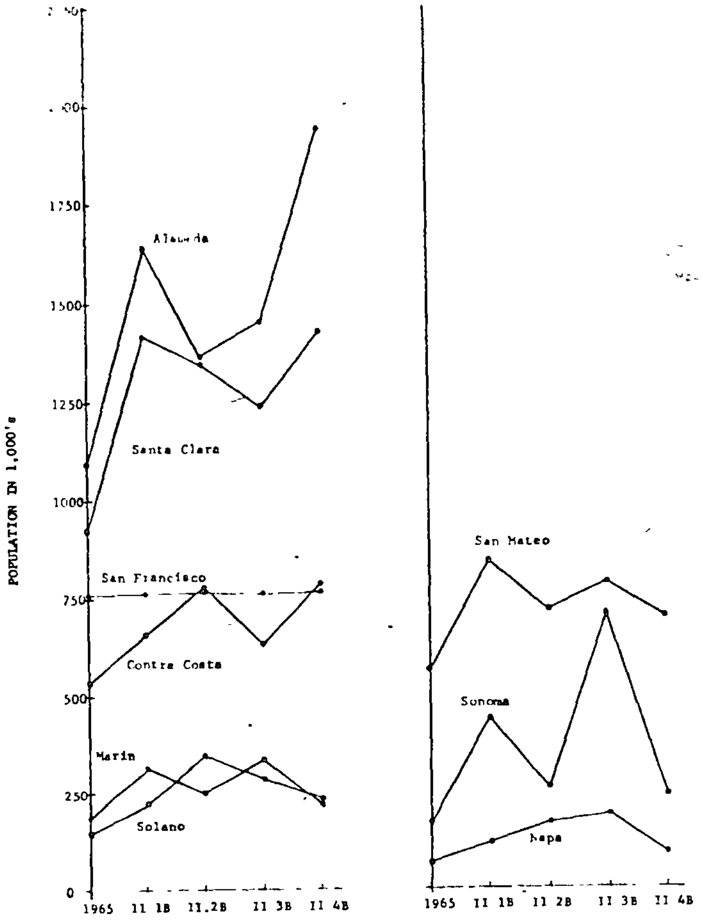


Fig 4 Round II. B – Population by County, 1980

county, with 69 zones, only 10–15 zones showed significant oscillations and of these, only 2 or 3 zones showed really large oscillations. In addition, it was determined that virtually all of the oscillating zones were on the fringes of the region and/or had high percentages of available vacant land in them.

Further Tests of the Model Package. Round IV, Network Measures

Several statements by Goldner in his report on work with IPLUM for the U S D O T. (Goldner, 1972, Vol II, pp. 57–58, 94) allude to the problems of largely vacant fringe zones In general, these problems stem from the form of the function which calculates the measure of attractiveness of zones for residential allocation In Goldner's IPLUM, as well as in the version of IPLUM used in the runs described above, this measure was essentially a function of the vacant land in the zone. Clearly, for large, essentially undeveloped (i.e. largely vacant) zones, the level of attractiveness would therefore be quite high. At some point in the development of IPLUM Goldner added a correction factor to the attractiveness measure to deal with this problem. Later, for reasons which were not stated, he deleted this correction factor and thus restored the model to its original form Based on the results obtained from Rounds IIA and IIB it was decided to introduce this correction factor to our version of IPLUM.

The theoretical basis for this correction factor is that the attractiveness of a zone is determined by developable vacant land, rather than vacant land *per se* This being the case, a measure of the developability of land may be constructed in terms of the proportion of developable land in a zone which has already been developed. It may then be asserted that this would be a reasonable surrogate measure of the available "infrastructure" e.g. sewer and water lines, roads, etc. in the zone. After this correction factor was introduced into the IPLUM household allocation function, a new Round of model runs, called Round IV, was undertaken. The results of this Round are shown in Table III and Figure 5. It can be seen in Figure 5 that compared to Round II, a substantial reduction in the system's oscillations has been achieved for several counties. Nonetheless, the oscillations are still present and still significant. This oscillation was also visible in the network loadings, with network links connecting to zones of population oscillation showing correspondingly large traffic volume oscillations.

Referring to Figure 5, it appears, with the exception of Sonoma County, that in going from Run IV.2 to Run IV 3 the system may have gone past or "overshot" an equilibrium point. Having done so, the subsequent attempt from Run IV.3 to IV.4 to return to the equilibrium point, appears to overshoot in the opposite direction Once started, this process appears to be capable of continuing indefinitely. Once having entered this "unstable" situation, the model system appears unable to move to a stable configuration. The next experiments were attempts to find procedures which would help the model system either to achieve a stable situation directly, or to move to one, from a prior unstable one. The trips generated from Run IV 3 and IV.4 were averaged together O D.-pair-by-O.D -pair. The resulting "averaged" trip matrix was loaded on the 1980-network and used as input to Run IV.5 AVG. The trips from Run IV.5 AVG were then averaged with those used as its inputs (i.e. the average of IV 3 and IV 4) and again loaded on the 1980-network The resultant impedances were used as input to IPLUM Run IV.6 AVG. This procedure, as one would expect, did virtually eliminate the model sys-

TABLE III

Round IV Base Population, Projected Total Population for 1980, and Percent Change 1965–1980

	1965	1980–Run IV.1		1980–Run IV.2		1980–Run IV.3		1980–Run IV.4		1980–Run IV.5	
San Francisco	754,754	759,821	0.6%	767,173	1.6%	763,390	1.1%	766,364	1.5%	763,276	1.1%
Marin	192,523	313,982	63.0	275,484	43.0	363,621	88.8	238,932	24.1	380,640	97.7
San Mateo	553,271	840,265	51.8	762,003	37.7	746,089	34.8	788,324	42.4	735,260	32.8
Sonoma	167,497	440,337	162.8	242,874	45.0	474,468	183.2	230,152	37.4	520,828	210.9
Napa	69,378	118,582	70.9	118,737	71.1	163,118	135.1	95,836	38.1	179,825	159.1
Solano	149,386	223,074	49.3	363,296	143.1	328,929	120.1	262,109	75.4	450,116	201.3
Contra Costa	534,056	655,638	22.7	829,304	55.2	738,900	38.3	787,239	47.4	703,485	31.7
Santa Clara	922,357	1,412,834	53.1	1,412,478	53.1	1,325,489	43.7	1,520,842	64.8	1,287,851	39.6
Alameda	1,095,236	1,634,003	49.1	1,636,193	49.3	1,503,543	37.2	1,717,740	56.8	1,386,261	26.5
TOTAL	4,438,358	6,407,472	44.4	6,407,542	44.4	6,407,547	44.4	6,407,568	44.4	6,407,542	44.4

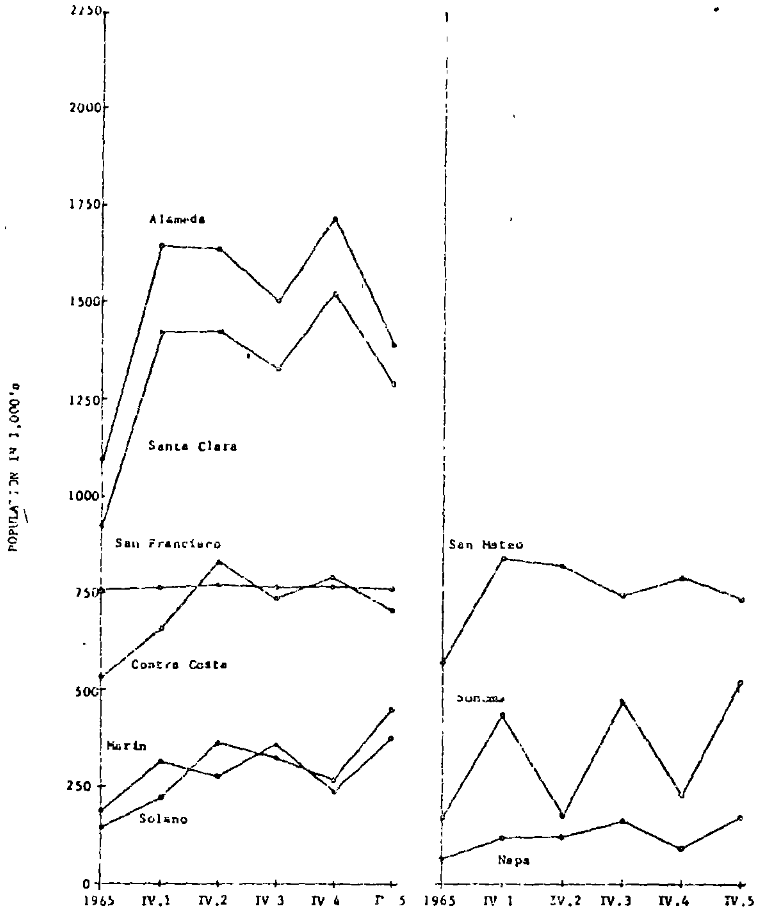


Fig 5. Round IV – Population by County, 1980

tem's oscillations. The results of these runs are shown in Table IV and Figure 6. It appears that if this procedure of trips averaging were imposed on the model system beginning with the inputs to Run IV.3, i e with average of Run IV.1 and Run IV.2 trips, the oscillations would be brought to within a tolerable range. Consequently, this trip averaging procedure, which is not a particularly "extreme" solution to the oscillation problem, was adopted as standard procedure in the model package.

As described earlier, two principal aims of this research effort were wholly to eliminate forecast errors due to the differences between free-flow and loaded network characteristics, and partially to eliminate the errors due to the incremental land development resulting from the changes in the transport networks. The first aim is achieved by using loaded networks as input to the land use model. The second aim is partly achieved by repeated iterations of the land use model and the network

TABLE IV

Round IV AVG Base Population, Projected Total Population for 1980, and Percent Change 1965–1980

	1965	1980–Run IV.2		1980–Run IV.3		1980–Run IV 4		1980–Run IV 5 AVG		1980–Run IV.6 AVG	
San Francisco	754,754	767,173	1 6%	763,390	1.1%	766,364	1.5%	765,234	1.3%	764,950	1 4%
Marin	192,523	275,484	43.0	363,621	88 8	238,932	24.1	334,808	73.9	320,451	66 4
San Mateo	553,271	762,003	37.7	746,089	34.8	788,324	42 4	788,168	42 4	778,837	40 8
Sonoma	167,497	242,874	45 0	474,468	183 2	230,152	34.4	294,499	75.8	288,333	72.1
Napa	69,378	118,737	71.1	163,118	135 1	95,836	38 1	126,219	81 9	130,566	88 2
Solano	149,386	363,296	143.1	328,929	120.1	262,109	75 4	319,934	114 1	307,813	106 1
Contra Costa	534,056	829,304	55 2	738,900	38.3	787,239	47.4	780,736	46 1	788,455	47 6
Santa Clara	922,357	1,412,478	53.1	1,325,489	43.7	1,520,842	64 8	1,382,363	49.8	1,401,803	52 0
Alameda	1,095,236	1,636,193	49.3	1,503,543	37.2	1,717,740	56 8	1,621,579	48.0	1,626,328	48.5
TOTAL	4,438,358	6,407,542	44.4	6,407,547	44.4	6,407,568	44 4	6,407,540	44 4	6,407,556	44.4

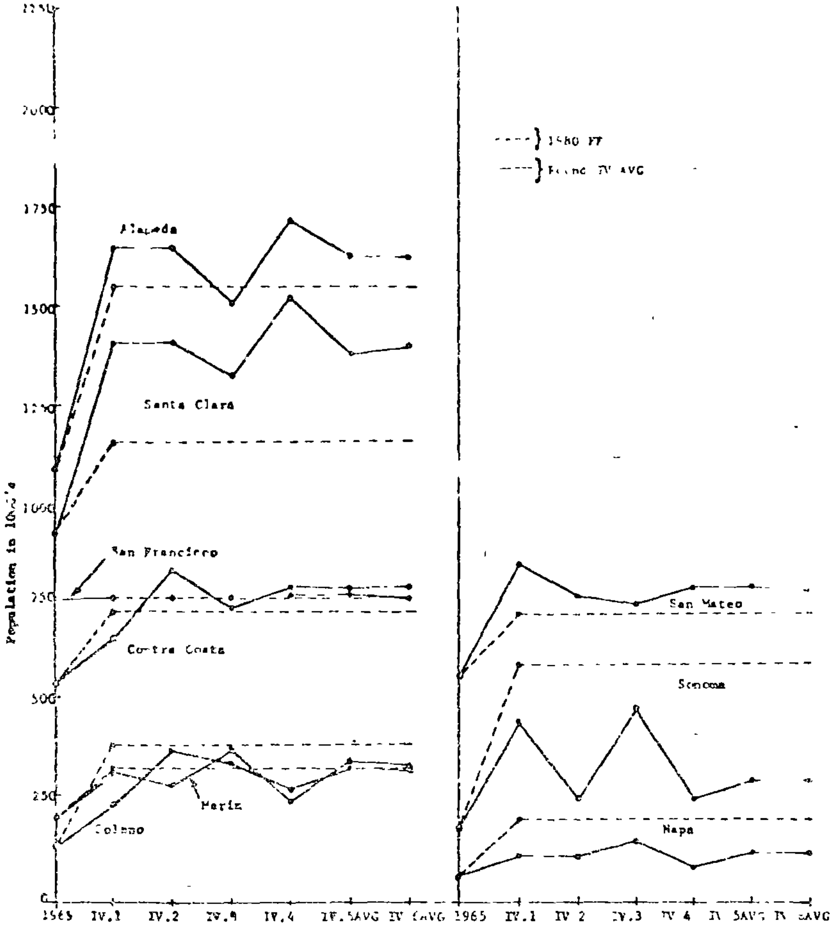


Fig 6. Round IV.AVG and Run 1980 FF

model. It should be noted that while the various intermediate steps in this iterative process do not purport to have any one-to-one correspondence to simulated points in time, the overall stream of iterations is probably similar to a simulated stream of time

An obvious point of comparison for this integrated forecasting procedure is with similar forecasts made in the traditional manner The modified version of IPLUM was run with the "empty" (lacking any trip loadings) 1980 network impedances as input This, of course, represents the standard procedure for unintegrated use of a land use model. This yielded a 1980 land use forecast based on the projected 1980-network without any consideration of the traffic which would develop as a result of the changed land use distribution. The county population forecasts from this run (called 1980-FF) are:

San Francisco	757,892	Solano	[illegible]74,101
Marin	319,089	Contra Costa	725,874
San Mateo	716,536	Santa Clara	1,162,194
Sonoma	589,129	Alameda	1,548,474
Napa	213,256		

This set of forecasts does not reflect any of the feedback from land use to transportation. If the actual links of the 1980 networks had unlimited capacity, under which circumstance the travel characteristics of the links would not deteriorate as the volume on the links increased, then theoretically these forecasts would be acceptable, though still not quite correct. The remaining theoretical error is due to the incremental development of changes in actual transport facilities which allows portions of them to come into public use and to have some effect on land use before the completion of other portions Any model which makes its forecasts by discrete jumps through time will inevitably miss some of these phenomena.

The 1980-FF forecasts are plotted in dashed lines on Figure 6 It is quite clear from the figure that these results differ substantially from the results indicated by the stream of iterations. While it has not yet been possible to make a statistical comparison of these results, it appears that the results from the integrated model system are much more likely to be correct, i e better forecasts Finally, it should be noted that in those areas where there has been substantial real growth in population (the southeastern counties); the integrated model system results are consistently greater than those of 1980-FF. This is an encouraging result and is suggestive of the potential advantages in forecasting accuracy to be had from the integrated model system.

Before concluding this paper, some further discussion of the network modelling is in order. It is clear that it is of importance to have ways of assessing the "situation" on the transport networks Since the origin-to-destination trip matrix will change, as described above, as the location of population and employment changes, the volume loadings on the network links will provide an indication of the traffic conditions on the network. One possible measure is of "total-trip-minutes" Unfortunately this measure does not accurately reflect the true traffic situation on the network since there is no way to distinguish whether the total-trip-minutes come from large numbers of short trips or fewer numbers of long trips. After further consideration, it was decided that the mean and standard deviation of 1) link volume/link capacity ratio, 2) origin to destination impedance, and 3) trip length would provide some useful information about the traffic situation on the network These data are shown for Round IV.AVG in Table Va

TABLE Va

Network Characteristics on Input to IPLUM Run

	IV 3	IV 4	IV 5AVG	IV 6AVG
Mean Vol./Cap.	0 798	0 883	0 827	0 811
S Dev Vol./Cap.	0 839	0 960	0 822	0 804
Mean Imped.	74.2	89 0	84 0	80.9
S Dev Imped	38 3	52 0	44.3	42 5
Mean Trip Lgth.	32 0	43 5	36 2	34 1
S Dev Trip Lgth	23.4	39 2	27 2	24.7

TABLE Vb

Congestion Measure on Input to IPLUM Run

	IV 3	IV.4	IV 5AVG	IV 6AVG
San Francisco	1 891	2 005	1 939	1 928
Marin	1.425	2 025	1 229	1 090
San Mateo	2 246	2 176	2 144	2 102
Sonoma	2 446	2.711	2 453	2 407
Napa	1 048	2 218	1 341	1 262
Solano	1.837	1 467	1 526	1 551
Contra Costa	1 558	2.311	1 567	1 566
Santa Clara	1.719	1 663	1 692	1.677
Alameda	1 810	1 922	1.777	1 766

While most of the oscillations in the county population forecasts can be seen to follow the oscillations of the network measures, these measures are not zone or county specific. Consequently it is not possible to compare zonal or county forecasts with the conditions on the relevant portions of the network. A crude zone-specific measure of congestion was therefore developed, beginning with the fact that in each zone there
network load-node All traffic which either originates or terminates i
zone must pass through this load-node We look first at all of the netw
links which originate at the load-node, the traffic on these links, and
volume/capacity ratio for the loaded links It is then possible to com
an average volume/capacity ratio weighted by the traffic experiencing
level of congestion implied by the volume/capacity ratio. While
measure does give an indication of zone-specific congestion, it has
fault of not distinguishing “through” traffic on the links passing thr
the load-node (i e traffic which neither originates nor terminates in
zone), from trips originating at the load-node. Consequently, the part
lar placement of the load-node in the zone affects the amount of thr
traffic and the results Nevertheless, this zonal measure can be extende
the county level which averages the effect of particular load-node pl
ments, and thus comprises a county level measure of transport netw
congestion. These measures were calculated for the runs of Ro
IV.AVG and are shown in Table Vb.

If the behavior of the congestion measures is compared to
population projections, some interesting observations may be develo
There are two general classes of behavior to be observed. The first is w
the population projections and congestion measures are one itera
“out-of-phase” with each other, e g. IV.3 congestion is “high” and I
population is “low,” while IV 4 congestion is “low” and IV 4 populat
is “high”. This situation occurs for five of the nine counties (Santa Cl
San Mateo, Sonoma, Napa, and Marin) This situation is what was
pected of the model system, being of the general form high populatio
high congestion → lowered population → lowered congestion → hig
population → etc.

Second, the remaining four counties (Alameda, Contra Costa,
Francisco, and Solano) exhibit a different form of behavior. In th
counties the population projection and congestion measures are
phase” with each other, e g IV.3 congestion is “high” and IV.3 pop
tion is “high,” while IV.4 congestion is “low” and IV.4 populatio
“low.” This situation is of the general form. high population → l
congestion → lowered population → higher congestion → higher populat
→ etc. This is a rather unexpected mode of behavior and is currently be
investigated. There are several possible explanations to be tested Fi
there is the possibility that the “through-traffic” ignored by the cong
tion measure is sufficiently important to be biasing the measure a
making it an inaccurate indication of the true area-specific levels of con
tion. A second possibility is that congestion is not properly measu
without considering numbers of originating trips and average trip lengt

e congestion in an area could increase if number of trips declined slightly and trip lengths decreased substantially The fact that there appears to be a systematic spatial distribution of the counties displaying these modes of behavior suggests that this problem can be resolved in a systematic fashion.

Two important positive points to be noted here will conclude the discussion of network measures First, the overall network measures given in Table Va do converge with the population projection convergence and are probably indicative of the desired equilibrium being reached Second, the congestion measures shown in Table Vb also seem to converge with the converging population projections. The anomolous iteration-to-iteration behavior of the congestion measures is more likely to shed additional ight on the nature of the process than it is to vitiate the results obtained. Finally, we note that in the development of a system of models and nterrelationships as complex as this one, simplifying assumptions and ertain gaps, both theoretical and operational, is inevitable. Here, however, and especially in view of the preliminary nature of these results, the uestion is not one of absolute accuracy of forecasts. The nature of the esponse of the whole system is the key result The demonstration of the lative sensitivity of land use forecasts to transportation network condiions, and of traffic on the networks to land use forecasts, and that they ach equilibrium, is the principal output to date.

There is, of course, room for improvement to this model system The lowing is but a sample of the questions to be explored in regard to the dequacy (or accuracy) of the system's forecasts.

a) What are the consequences of redistribution of basic employent? The PLUM and IPLUM models are derivatives of the LOWRY odel and as such the forecasts of basic employment are exogenous puts to the models. It is clear that transportation networks do have an fect on the location of some basic employment. Further, there are ong indications of decentralization of some basic employment in metrolitan areas To the extent that this may represent a significant shift in kplace location, there are impacts to be felt throughout the area How ould the model system deal with this problem (Putman, 1972)?

b) Both PLUM and IPLUM locate residences strictly as a function of trip distributions, available land, and subsequent. somewhat arbiary constraints. Work with other land use models indicates that local amenities of potential residential locations have a significant imct How can these models be modified to include these phenomena?

c) The networks now used in the system are highway only. Clearly, e of the important questions for future metropolitan development will be that of the utility of public (mass) transit systems How can the mode system be modified to include both highway and transit networks as wel as a modal-split procedure?

d) The level of sectoral detail throughout the model system is to highly aggregated. Where and how can disaggregations be accomplished?

e) What are the causes and consequences of the remaining instabilit of the model system?

All of these questions, plus others, require further investigation It i worth noting that some further work on the stability question was don which yielded encouraging results. During the ongoing testing of th model package, problems with the stability of the results continued to b encountered. After further analysis of the problem it appeared that one o the fundamental difficulties was the size of the projection interval In th existing model package a projection is made to the year 1980 from a bas year of 1965 in one step Successive iterations of the model package ar needed simply to attempt the finding of an equilibrium solution It wa thought that an alternative means of finding the equilibrium would be t approach it more gradually in a way analogous to making the projectio from 1965 to 1980 in several steps rather than in one great step It wa hoped that such a procedure would both eliminate the remaining insta bilities in the model system and be less expensive to operate by virtue o requiring fewer iterations of the model package The evidence available a this time is that it will indeed do what we hope.

A verbal description of this process, shown in Figure 7, is as follow

1. Assume that the base year is 1965 and the projection year is 1980

2. The base year (1965) trips are loaded on the 1980-network an zone-to-zone impedances are calculated.

3 With these impedances, one-third of the regional control totals and one-third of the 1965 → 1980 basic employment increments, a inputs, IPLUM is run and new trip matrices generated.

4 The 1965 trips are subtracted from the new trip matrices and the remaining new trips are loaded on the already partly loaded 1980-network New zone-to-zone impedances are calculated.

5. With these impedances, two-thirds of the regional control totals and an additional one-third of the basic employment increments, a inputs, IPLUM is run and new trip matrices are generated.

6 The trip matrices from Step 4 are subtracted from these new tri matrices and the remaining new trips are loaded on the partially loade 1980-network. New zone-to-zone impedances are calculated.

7. With these impedances, the full regional control totals, and the final one-third of the basic employment increments, as inputs, IPLUM is

run and new trip matrices are generated.

8 The trip matrices from Step 6 are subtracted from these new trip matrices and the remaining new trips are loaded on the partially loaded 1980-network. New zone-to-zone impedances are calculated

9. With these impedances, the full regional control totals, and zero basic employment increments, IPLUM is run and should then produce an equilibrium solution.

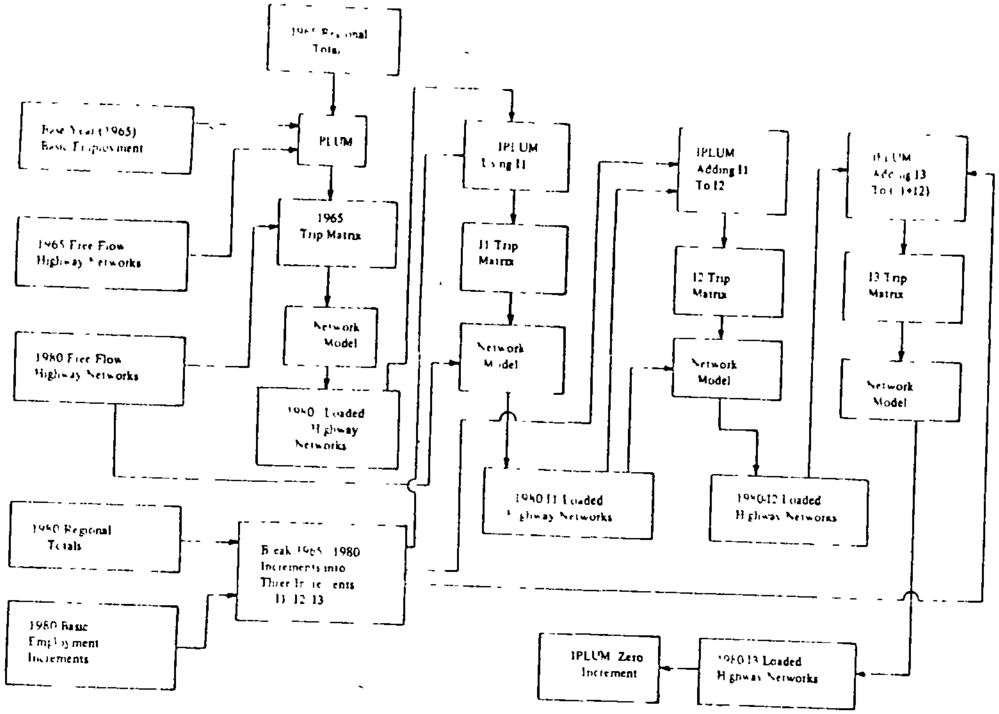


Fig 7 Schematic of New Process for Model Package.

This new process was run through step 8. The resulting population projections, where XA-1 = Step 3, XA-2 = Step 5, and XA-3 = Step 7, are shown in Figure 8. The dotted lines in Figure 8 represent the projections from Run IV 6AVG. It can be seen that the new process comes rather close after three recursions to where the old process was after six iterations. The fourth recursion of the new process was not run as described above, because of an unexpected operational deficiency in IPLUM The model is theoretically capable of projecting the response, in terms of spatial redistribution of local-serving employment and population, to changes in basic employment and/or transportation facilities. However, it cannot operate with a zero change in basic employment. Since there were

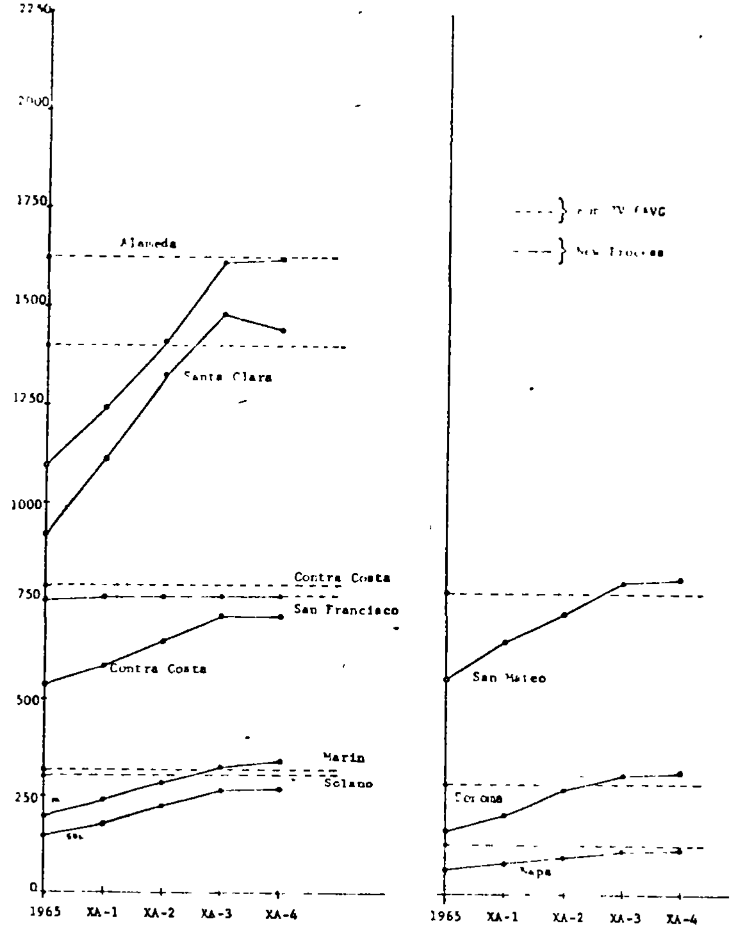


Fig 8 New Process, Population Projection for 1980

not sufficient resources (or time) on this project to correct this deficien in IPLUM, it was not possible to run the fourth recursion (Step 9) of new process as originally intended.

As an alternative test, the impedances from Step 8 were used a IPLUM was rerun as in Step 7. This means that the last growth increme was added to the previous two-thirds, but using the impedances deriv from the *prior* estimate of the last growth increment. The results of the two estimates of the spatial distribution of the last growth increment (o based on a two-thirds full network and the other on a three-thirds f network) should, if the new process had found an equilibrium in the fi three steps, be nearly identical. In fact, the largest difference, at county level, was less than 5%. These differences, shown as point XA-4 Figure 8, are:

San Francisco	+0.08%	Solano	+1.72%
Marin	+4.07%	Contra Costa	0.12%
San Mateo	+0.77%	Santa Clara	-2.64%
Sonoma	+1.76%	Alameda	+0.31%
Napa	+2.29%		

While further testing is clearly desirable, the implication of these results is that by this new process the whole system reaches an equilibrium point in just three steps as opposed to the five or six iterations required by the original procedure. Further work with the model package will be done with this new process.

Conclusions and Policy Implications

It will be recalled that the principal purpose of the research described in this paper was to investigate the interrelationships between transportation and land use. Further, it was desired to investigate the feasibility of balanced transportation development and land development. To do this, an integrated package of land use and transportation models was developed and subjected to limited testing. Though the concept has often been discussed, this effort represents, so far as we know, the first operational demonstration of the interrelationships, in particular, of the feedbacks, between transportation and land use. From the standpoint of method, the testing of this model system strongly suggests that the current, prevalent practice of using transportation and land use models without these feedbacks, is virtually certain to introduce significant errors in their results.

While much additional testing is necessary, the model tests suggest broad policy possibilities. Congestion of networks produces tendencies towards centralization of economic activities and land use in metropolitan areas. Uncongested networks, or improvements to congested networks, produce tendencies towards decentralization of economic activities and land use in metropolitan areas. Consequently, it is possible that policies which allow congestion of networks or make travel more difficult (and/or costly?), will assist in the reduction of metropolitan sprawl. Similarly it is possible that policies which relieve congestion and make travel less difficult will promote metropolitan sprawl and subsequently promote more travel and additional network congestion. Policies which limit land development in some areas, combined with policies which attempt to relieve congestion on some portions of the network, while allowing it on others, may be powerful instruments for maintaining a balance between transportation and land use. Such policies might prove capable of directing or channeling metropolitan growth and, for example, in combination with other forms of public investments could be useful in the development of new towns (or sub-nucleated metropolitan regions) and long-term relief of many of the pressing urban problems of today.

Acknowledgement

The research described in this paper was done as part of Contract No. DOT-FH--11 -7843 from the U.S. Department of Transportation, Federal Highway Administration,- Urban Planning Division to the Institute of Environmental Studies, Department of City and Regional Planning, University of Pennsylvania. One of the outputs of that contract was a report entitled *The Interrelationships of Transportation Development and Land Development.* Certain tables and figures in this paper were originally prepared for that report.

References

Goldner, W. (1968) *Projective Land Use Model (PLUM), BATSC Technical Report 219* Bay Area Transportation Study Commission, Berkeley, California.

Goldner, W., et al. (1972) *Projective Land Use Model* (in three volumes). Institute of Transportation and Traffic Engineering, University of California, Berkeley, California.

Harris, B., Nathanson, J., and Rosenburg, L. (1966) *Research on an Equilibrium Model of Metropolitan Housing and Locational Choice. Interim Report* Planning Sciences Group, Department of City and Regional Planning, University of Pennsylvania, Philadelphia, Pennsylvania.

Hill, D. M. (1965). "A growth Allocation Model for the Boston Region," *Journal of the American Institute of Planners* Vol. XXXI, No. 2, pp. 111–120.

Kuner, P. (1973a). "ITLUP Integrated Transportation and Land Use Package-Version 2 Computer Documentation," Appendix B-1, in Putman, S. H., et al., *The Interrelationships of Transportation Development and Land Development.*

Kuner, P. (1973b). "Investigation of Traffic Assignment Algorithms," Appendix C-1, in Putman, S. H., et al., *The Interrelationships of Transportation Development and Land Development.*

Lowry, I. S. (1964) *A Model of Metropolis RM-4035RC* RAND Corporation, Santa Monica, California.

Nathanson, J. (1970). *Basic Employment MODel. A Model for Intra-County Location of Basic Employment and Land, BATSC Technical Report 222* (Preliminary) Bay Area Transportation Study Commission, Berkeley, California

Putman, S. H. (1972) "Intraurban Employment Forecasting Models: A Review and a Suggested New Model Construct," *Journal of the American Institute of Planners* Vol. XXXVIII, No. 4, pp. 216–230.

Putman, S H , et al. (1973). *The Interrelationships of Transportation Development and Land Development* (in two volumes) Institute for Environmental Studies, Department of City and Regional Planning, University of Pennsylvania, Philadelphia, Pennsylvania

Putman, S H (1974). *Urban Land Use and Transportation Models A State-of-the-Art Summary* (in the press).

U.S Department of Transportation (1972). *Urban Transportation Planning—General Information and Introduction to System 360* Federal Highway Administration, Washington, D.C.

SAN JOSE NEWS
SAN JOSE, CALIF.
D. 75,531

AUG 2 4 1973

'Mouse Trap' Freeways

By JIM DICKEY
Staff Writer

Freeways are kind of like the proverbial mouse trap.

If you build a better road, the world will beat a path to its doorstep — or access ramps.

So says a University of Pennsylvania professor who has completed a study of freeways and land-use planning of the Bay Area, including Santa Clara County.

DR. Stephen Putman, associate professor of city planning at the university, concluded in his study that congestion on highways is not likely to be eased by widening existing roads or building new ones.

In fact, he said, it appears that traffic congestion will follow wherever new facilities take it, often accompanied by strips of undersirable commercial development.

Putman's study marks the first time computerized transportation land-use models have been integrated successfully, according to the university.

The professor used computers to "build" two imaginary roads to relieve congestion on existing roads.

The hypothetical roads ran between Oakland and Livermore Valley, and between Livermore and northwest San Jose.

The imaginary freeways were intended to centralize impending development in eastern Alameda County.

But the "roads" did something quite different from the plan.

Area residents used the highways for long distance commuting. This resulted in a general southerly shift of the region's employment and population.

The southeastern portion of Marin County was swamped with new residents and local-serving employment development.

In the absence of regionwide land use control policies, the highways did not produce the desired results, Putman said.

The professor also used computers to limit development in western Contra Cost County and central Alameda County by imposing strict land-use controls. The policy was successful in limiting growth in those two counties. But because the controls were not coordinated on a region-wide basis, development ran rampant in the adjacent counties of Marin, Sonoma, Napa and Solano.

"The message to the Bay Area is clear: land-use controls and transportation controls must be integrated on a regionwide basis if development is going to be accurately forecast and controlled," Putman said in his report.

The professor also explained how he believes the new highways/new developments cycle operates.

"As talk of a new road heard, speculators and developers consider purchasing adjacent land. Plans progress and developers become more assured and begin to purchase land. Once the road is under construction developers begin their own residential and commercial construction nearby. It is little surprise then that once the road is finished, demand is greater than anticipated."

Increasing estimates of demand and building wider roads will not prevent overcrowding, he believes, because "a new facility creates its own demand."

S. H. PUTMAN

UNIVERSITY of PENNSYLVANIA
NEWS BUREAU, Franklin Building
3451 Walnut Street, Philadelphia, 19174

For release:
Sunday, August 26, 1973

For information, call:
Chat Blakeman
Tel. (215) 594 8721

Congestion on the nation's highways is not likely to be eased by widening existing roads or building new ones. In fact, it appears that traffic congestion will follow wherever new facilities take it, often accompanied by strips of undesirable commercial development.

That's one of the basic conclusions reached by a University of Pennsylvania study of transportation and land use. Dr. Stephen Putman, associate professor of city planning at the University and principal author says that unless zoning controls are incorporated with plans for new transportation facilities on a region-wide basis, the nation's transportation system is not likely to become more convenient than it it is now, no matter what roads are built.

Putman's study marks the first time computerized transportation and land use models have been integrated successfully. Studies to date have usually assumed that one of the two variables remains constant. It is Putman's contention however that a dynamic relationship exists between the two and that to plan highway facilities without regard to land use controls, or vice versa, will lead to unexpected and uncontrolled growth.

(more)

The results of planning which didn't take account of both can be seen in highways which are overcrowded soon after they are completed and which suffer premature obsolescence. "The principal question addressed by the research is whether it is feasible, via integrated transportation planning and land-use planning, to avoid or ameliorate the occurance of this particular phenomenon in the future," the report states.

Putman's study focuses on the nine-county San Francisco Bay area, a region which is typical in experiencing considerable development over the last decade. (A summary of the Bay Area findings is attached.)

The typical scenario for highway construction in this area and others runs something like this, Putman says:

As talk of a new road is heard, speculators and developers consider purchasing adjacent land. Plans progress and developers become more assured and begin to purchase land. Once the road is under construction developers begin their own residential and commercial construction nearby. It is little surprise then that once the road is finished, demand is greater than anticipated.

Increasing estimates of demand and building wider roads will not prevent overcrowding because, says Putman, "a new facility creates its own demand." The most effective solution according to the report is to make land control an integrated part of transportation planning.

"Viewed in the large sense," he says, "we're talking about the process by which metropolitan areas expand. In more immediate terms we're talking about ugly urban sprawl."

(more)

Los Angeles probably offers the best example of what unplanned growth can mean. New highways encouraged outmigration, spreading the city's effective borders further and further and encouraging a decentralization which makes the creation of effective mass transportation facilities both difficult and exceedingly expensive. Other examples of highways which have led to sprawl include the Schuylkill Expressway in Philadelphia and the Shirley Highway running south from Washington, D.C.

If land use control were more closely integrated with transportation planning Putman believes "nucleated" cities would develop, separated by expanses of park and recreation land. This is the direction toward which the country should be planning Putman maintains, rather than allowing densely populated suburbs and congested roads to expand in all directions from cities.

Putman's report, sponsored by the Federal Highway Administration of the Department of Transportation, suggests two sets of planning alternatives. One, biased toward the status quo, suggests the government could continue to build roads literally until money or space runs out (demand will not decrease because a large number of people are still anxious to leave the city). Similarly, the government could continue to build roads and mass transit facilities simultaneously. This is by far the most popular solution, although Putman argues it is no solution at all. While it may have appeal for the short-run, such planning will cause new and greater problems in the long-run.

The second set of alternatives would seek to reduce and relocate demand. One possibility would be to allow roads to become so overcrowded and inconvenient that people would choose other means of transportation or eventually move closer to their places of employment. This is clearly not a viable solution.

It is necessary, and viable, Putman believes to make land control a part of transportation planning, limiting the amount of commercial and residential construction along new facilities. Widely employed by the British, this would ultimately lead to cities surrounded by "greenbelts" and discourage the type of urban sprawl so common today.

PRELIMINARY DRAFT

CALIBRATING URBAN RESIDENTIAL MODELS 2:

EMPIRICAL RESULTS

Stephen H. Putman

Federick W. Ducca

December, 1977

Urban Simulation Laboratory
Department of City and Regional Planning
University of Pennsylvania

INTRODUCTION

One of the critical tests of validity of an urban simulation model is its ability to perform well on a large number of different urban areas. One measure of "performing well" is the extent to which fitting the model to different urban areas yields similar parameters. To the extent that these simularities are found, and are consistent with the model's underlying theoretical assertions, they provide evidence of its general validity.

In Part 1 of this paper we described procedures and strategies for estimating the parameters (i.e. calibrating) a familiity of spatial interaction models, singly constrained, and used as location models. In this Part 2 of the paper we describe the results of fitting various forms of the Disaggregated Residential Allocation Model - DRAM (Putman, 1977) to a number of different urban area data sets.

CITY-TO-CITY COMPARISONS: FULL DRAM FORMULATION

The Disaggreagated Residential Allocation Model - DRAM derives from the evolutionary stream of Lowry derivative models, cast in terms of Wilson's entropy-maximizing formulation (Putman, 1978; Wilson, 1970). Before presenting the empirical results we give here a brief description of the model.

First we note that the residential location component of any Lowry type model may be written

$$N_i = b \sum_j E_j P_{ij} \tag{1}$$

where

N_i = employed residents living in zone i

E_j = number of persons employed in zone j

P_{ij} = the probability of a work trip between zones i and j

b = a scaling factor to ensure consistency between number of employed residents and number of employees.

The more sophisticated Lowry type models make the P_{ij} a multivariate function of the residential attractiveness of zone i and the time and/or cost of travelling between zones i and j. Thus we have

$$N_i = b \sum_j E_j \; W_i \; f(C_{ij}) \tag{2}$$

where

W_i = residential attractiveness of zone i

C_{ij} = a generalized cost of travel between zone i and zone j

By following the Wilson approach, one may derive a simple residential location model of the following form

$$N_i = \sum_j B_j W_i E_j f(C_{ij}) \tag{3}$$

where

$$B_j = \sum_i W_i f(C_{ij}) \tag{4}$$

Then, by substitution

$$N_i = \sum E_j \left[\frac{W_i f(C_{ij})}{\sum_i W_i f(C_{ij})} \right] \tag{5}$$

This then is the general form of DRAM, with two further embellishments.

First, in most of the Wilson applications

$$f(C_{ij}) = e^{-\beta C_{ij}} \tag{6}$$

where β is an empirically derived parameter. In DRAM,

$$f(C_{ij}) = C_{ij}^{\alpha} e^{\beta C_{ij}} \tag{7}$$

where both α and β are empirically derived parameters. Second, in most models W_i has been a single variable such as population or floor area of zone i. In DRAM

$$W_i = V_i^a P_i^b R_i^d N_{1,i}^q N_{2,i}^r N_{3,i}^s N_{4,i}^t \tag{8}$$

where

V_i = vacant, buildable, land in zone i

P_i = one plus the percentage of buildable land in zone i which has already been built upon

R_i = residential land in zone i

$N_{1,i}$ = one plus the percentage of employed residents of zone i who are in the lowest income quartile

$N_{2,i}$ = one plus the percentage of employed residents of zone i who are in the lower middle income quartile

$N_{3,i}$ and $N_{4,i}$ = upper middle and upper income quartiles

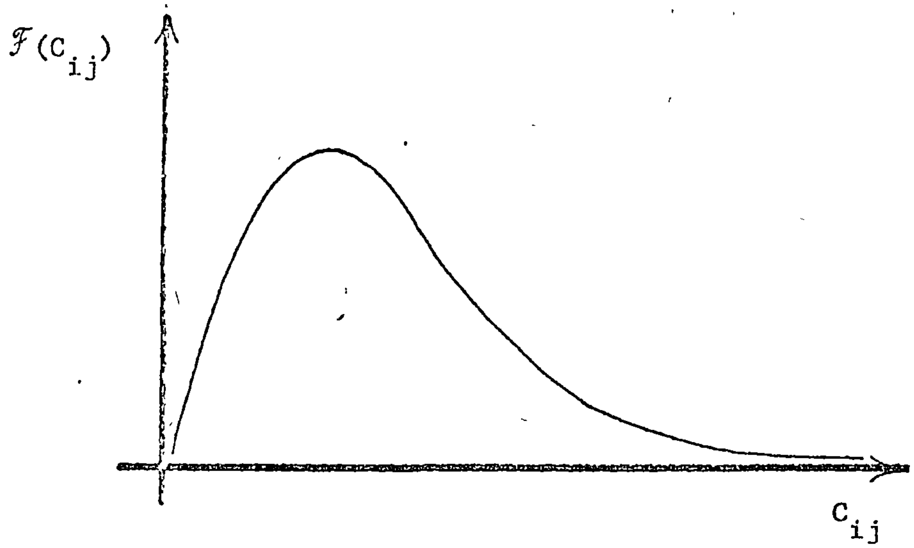
a,b,d,q,r,s,t = empirically derived parameters

Thus in estimating the parameters of DRAM we are estimating nine parameters. As the model directly forecasts the location of employed residents in each of the four income quartiles, we must therefore estimate four sets of nine parameters each. The procedures used in producing these estimates were discussed in Part 1 of this paper and will only be mentioned here as using a gradient search technique with a maximum likelihood criterion. We may next discuss our expectations as to the signs of these parameters.

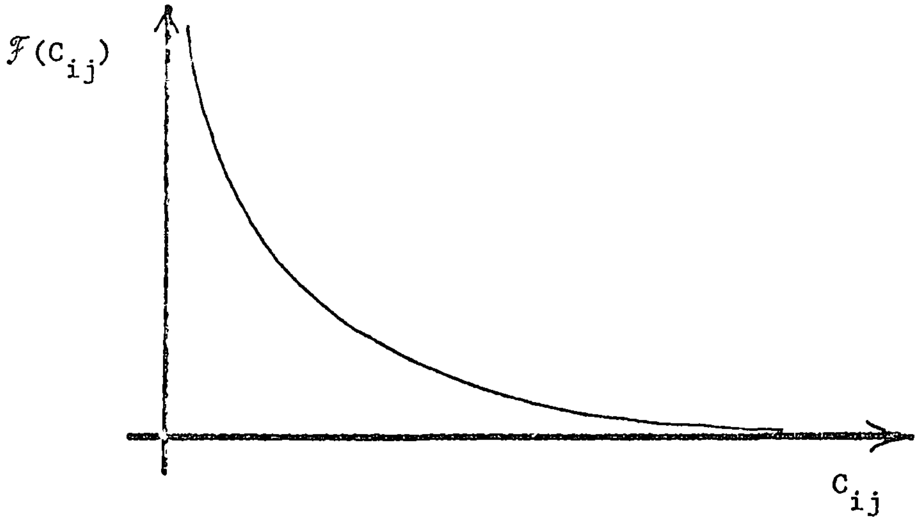
The trip function, equation (7), may be considered as being the product of two factors. The first factor, C_{ij}^{α}, is a measure of increasing numbers of residential location opportunities encountered with increasing distance travelled from the workplace in zone j. The second factor, $e^{\beta C_{ij}}$, is a measure of peoples willingness to make increasingly long (in time or cost) trips between work and home. Thus we expect that in virtually all cities $\alpha > 0$ and $\beta < 0$. In words, the greater the distance travelled from place of work, the more residential location opportunities encountered, and the longer the work-trip, the less willing persons will be to make it. Further, we expect higher income groups to be more willing to travel long trips than are the lower income groups. While we expect in most cases to observe $\alpha > 0$ and $\beta < 0$, it is possible to obtain other combinations. In our empirical work the overwhelming number of cases yielded the expected $\alpha > 0$, $\beta < 0$. A few cases turned up (to be described later in this paper) where $\alpha < 0$, $\beta < 0$ and where $\alpha < 0$, $\beta > 0$ but $\alpha >> \beta$. The effects of these different combinations of signs on the shape of the trip function are shown in Figure 1.

We turn next to the parameters of the three "land" variables. All three of these variables are expected to have positive parameters. The residential land variable and the vacant land variable are expected to operate together to describe the zones' development potential and to adjust for variations in size of zone. The percentage of buildable land already built on is expected to amplify or attenuate the first two by virtue of its being a surrogate measure of infrastructure development.

Figure 1: Alternate Shapes of the Tanner Function



a: where $\alpha > 0$, $\beta < 0$



b: where $\alpha < 0$, $\beta < 0$

or $\alpha < 0$, $\beta > 0$; and $|\alpha| >> |\beta|$

Further, we expect these parameters to range between 0.0 and 2.0, averaging somewhere near 1.0.

Finally, we have the four household composition variables. Here we expect the parameters for like household to like household to be significantly greater than those for each household to any other. Further, we expect the coefficients for adjacent household quartiles to be greater than those for dissimilar quartiles. Earlier results of work with DRAM contain further discussion of these expectations (Putman, 1977 and 1977a). As the title to this section indicates, we are concerned here with data sets which were sufficiently complete to allow calibration of the full DRAM formulation. In such cases parameters were separately estimated for each of the four employed resident types i.e. by income quartile. Thus the results described in this section of the paper, being for each of the four income quartiles for each of eleven metropolitan areas, represent forty-four fittings of DRAM to different data sets.

In Table 1 the values for α and β obtained for each income group of employed residents in each metropolitan area are given. The first result to notice is that in all cases, we get $\alpha > 0$ and $\beta < 0$, the expected result. City-to city differences in the α values follow no apparent pattern, nor do the within-city variations from income group to income groups. Work is currently underway on attempting to associate these parameter variations with differences in city sizes, ages, or general patterns of location of the different income groups. The within city, income group to income group variations in β have no readily discernable pattern either. The city-to-city variations do however yield an interesting result. The absolute values of β tend to decrease from city-to-city, moving from east-to-west across the continent. The inference to be drawn from this is that people in western cities are more willing to make a

Area	No. Zones	Year	Parameter	LI	LMI	UMI	UI
Atlanta, GA	183	1970	α	1.89	1.92	1.66	1.62
			β	-0.73	-0.81	-0.74	-0.72
Denver, CO	183	1970	α	2.56	2.37	2.90	2.88
			β	-0.46	-0.39	-0.18	-0.12
Minneapolis, MN	108	1970	α	1.45	2.29	2.99	2.21
			β	-1.09	-1.10	-1.27	-1.10
Philadelphia, PA	173	1970	α	1.47	1.72	1.87	1.81
			β	-1.62	-0.79	-0.47	-1.38
Phoenix, AZ	229	1970	α	1.30	1.42	0.94	1.02
			β	-0.51	-0.51	-0.35	-0.42
San Diego, CA	85	1970	α	2.06	2.10	2.08	2.09
			β	-0.15	-0.13	-0.15	-0.12
San Francisco, CA	291	1965	α	1.05	0.30	0.33	0.87
			β	-0.45	-0.26	-0.25	-0.31
Seattle, WA	227	1970	α	2.24	2.23	2.23	2.20
			β	-0.41	-0.50	-0.53	-0.56
Tulsa, OK	118	1970	α	0.46	0.30	0.36	0.18
			β	-0.41	-0.23	-0.25	-0.70
Vancouver, B.C.	163	1970	α	0.69	1.26	1.20	1.32
			β	-2.42	-0.80	-0.80	-0.70
Washington, D.C.	110	1970	α	2.55	2.63	3.90	3.18
			β	-1.00	-0.96	-1.41	-1.22

work-trip of a given length than are people in eastern cities, e.g. westerners are likely to do more daily travelling than easterners. This result is certainly consistent with commonly held opinion on the matter and suggests that there may be much to be learned by our continuing investigation of city-to-city and income group to income group variation in these parameters and the relationships of this variation to other measures of the urban areas.

In Table 2 the values obtained for a, b, and d of equation (8) i.e. the exponents of the vacant buildable land, one plus the percentage of developable land developed, and residential land variables, are tabulated. We note in Table 2 slightly less uniformity of parameters across cities than we found for the distance function variables in Table 1. The residential land parameter is consistent across all eleven cities. It is positive as expected and relatively close to 1.00, as expected. The percentage of developable developed parameters are, with only a few exceptions, also within the anticipated range of 0.0 to 2.0, with most of the values falling between 0.5 and 1.5, as expected.

The vacant buildable land parameter has not turned out according to expectations. While it is consistent across the data sets in being close to 0.0, it does show both positive and negative signs in different cases. There are several possible causes for these results; some of which are examined in greater detail in a later section of this paper. One possibility is that with the other two "land" variables in the attractiveness measure, this variable does not contribute significantly to the explanation of variance and thus is not statistically significant. Unfortunately there are no significance tests appropriate to models such as this, making it impossible to do a rigorous test of this hypothesis.

AREA	LOCATOR	ATTRACTOR		
		VAC. LAND	PER. DEV. DEV.	RES. LAND
Atlanta	LI	-0.068	0.093	0.817
	LMI	-0.095	-0.146	0.898
	UMI	-0.049	0.300	0.940
	UI	-0.011	0.421	0.874
Denver	LI	0.056	0.278	0.838
	LMI	0.062	0.301	0.867
	UMI	-0.039	0.217	0.865
	UI	-0.020	0.350	0.778
Minneapolis	LI	-0.068	0.716	0.693
	LMI	-0.096	0.717	0.841
	UMI	-0.130	0.531	0.904
	UI	-0.075	1.000	0.762
Philadelphia	LI	0.022	1.390	0.592
	LMI	0.030	1.432	0.503
	UMI	-0.020	1.533	0.465
	UI	0.079	1.484	0.678
Phoenix	LI	-0.099	-0.032	0.625
	LMI	-0.089	0.016	0.646
	UMI	-0.069	0.154	0.578
	UI	-0.068	0.167	0.613
San Diego	LI	-0.108	0.383	0.840
	LMI	-0.099	-0.242	0.940
	UMI	-0.065	0.247	0.779
	UI	-0.022	0.228	0.794
San Francisco	LI	0.001	1.240	0.774
	LMI	0.030	1.028	0.829
	UMI	0.041	1.041	0.668
	UI	0.027	1.109	0.630
Seattle	LI	-0.209	0.410	0.785
	LMI	-0.166	0.434	0.921
	UMI	-0.147	0.436	0.941
	UI	-0.175	0.412	0.959
Tulsa	LI	-0.110	0.166	0.567
	LMI	-0.137	0.165	0.533
	UMI	-0.106	0.337	0.627
	UI	-0.039	0.337	0.647
Vancouver	LI	-0.030	0.401	0.568
	LMI	-0.024	0.780	0.908
	UMI	-0.039	0.837	0.827
	UI	-0.078	0.748	0.717
Washington	LI	-0.088	n.a.	0.495
	LMI	-0.066	n.a.	0.444
	UMI	-0.123	n.a.	0.468
	UI	-0.015	n.a.	0.311

Another possibility is that the great variation in zone size which is present in these data sets forces the parameters of these land parameters to act as adjusters for zone size, rather than indicators of the extent to which each of these variables contributes to zonal attractiveness. Resolution of this question is a goal of ongoing research with the model and these data sets.

Finally, in Table 3, the values obtained for q, r, s, and t of equation (8) i.e. the exponents of the four income quartiles of employed residents, are tabulated. Referring back to equation (8) we recall that the residential attractiveness of a spcific zone has two general components, household composition and land measures. The household composition component is represented in the form of one plus the percentage of households, in the zone, in each of the four income quartiles. These variables act as surrogates for many zone specific attributes. Many measures of the socio-economic status of a zone's population will, no doubt, correlate with that zone's population income distribution. So too will housing quality, educational opportunities (i.e. school system quality), and the general attractiveness of the zone from both aesthetic and environmental viewpoints. Here then is a specific place in the model structure where the approach is intentionally "macro" as distinguished from, say, the micro-economic approach which would perhaps attempt to quantify and bring other, more specific, attribute variables into the model structure. From the point of view of presently practical forecasting procedures, this "macro" approach seems to be much the more sensible line of approach. It is more sensible not only in terms of thus avoiding the difficulties of collecting the data for these other variables, but even more so in avoiding the problems of having to update these variables as part of a recursive forecasting process.

		ATTRACTORS			
AREA	LOCATOR	LI	LMI	UMI	UI
Atlanta	LI	3.17	1.23	-1.63	-2.57
	LMI	0.96	3.09	0.79	-1.89
	UMI	0.11	0.81	3.68	-0.78
	UI	-0.50	0.31	1.69	2.54
Denver	LI	3.66	1.70	-1.05	-1.58
	LMI	1.12	1.93	0.63	-1.02
	UMI	-0.04	0.69	1.63	-0.06
	UI	-0.25	0.17	0.76	1.13
Minneapolis	LI	3.87	0.59	-2.16	-2.46
	LMI	1.47	2.12	-1.34	-1.61
	UMI	0.03	0.38	1.19	-0.64
	UI	-0.25	-0.64	-0.65	2.74
Philadelphia	LI	3.48	1.53	0.28	-1.02
	LMI	1.76	2.96	0.56	-1.02
	UMI	0.26	0.97	3.05	0.31
	UI	0.35	0.83	1.39	1.65
Phoenix	LI	3.32	1.54	-1.49	-2.72
	LMI	1.27	3.57	0.75	-1.74
	UMI	1.27	1.05	4.16	0.44
	UI	-0.78	-0.56	0.75	1.24
San Diego	LI	3.22	1.27	-0.86	-1.46
	LMI	1.77	3.04	0.57	-1.52
	UMI	-0.18	0.56	2.57	-0.10
	UI	-0.31	0.49	1.01	2.02
San Francisco	LI	2.22	1.11	-0.60	-1.72
	LMI	0.83	3.00	0.23	-1.22
	UMI	-1.10	1.00	3.05	0.01
	UI	-1.77	-0.54	1.19	2.26
Seattle	LI	1.65	1.13	-0.80	-1.77
	LMI	1.12	2.53	0.98	-0.73
	UMI	-0.18	0.54	2.50	1.15
	UI	-1.38	-0.51	1.48	3.02
Tulsa	LI	3.46	1.60	-0.23	-3.35
	LMI	1.35	3.84	1.10	-1.40
	UMI	-0.20	0.91	3.68	0.99
	UI	-0.91	0.24	1.30	2.43
Vancouver	LI	1.93	2.96	-1.42	-2.38
	LMI	0.50	3.08	-0.55	-2.51
	UMI	-1.20	0.86	2.41	-1.13
	UI	-1.62	-0.48	0.11	2.10
Washington	LI	5.42	2.36	-0.79	-2.08
	LMI	1.77	4.43	0.24	-1.45
	UMI	1.73	3.79	4.85	-0.27
	UI	-1.07	0.38	1.89	2.90

In Table 3 the parameters associated with the household composition variables are shown, arranged to show the importance of each household type as an attractor to each household type as a locator. So, for example, in Minneapolis for the location of Lower Middle income households, Lower income households have an attractiveness parameter of 1.47 while Upper income households have a negative, -1.61, attractiveness parameter. In each case, a locator in any given income class is most likely to be located with others in its same income class. Looking next at the relationship of each income class with other income classes, we find further regularities. Low income locators are most likely to locate, in all cities examined, with Lower Middle income households and most unlikely to locate with Upper income households. Lower Middle income households are most likely to locate with Lower income households and they too, are most unlikely to locate with Upper income households. Upper middle income households are most likely, in the majority of cases, to locate with Lower Middle income households. Upper income households are most likely to locate with Upper Middle income households and most unlikely to locate with Lower income households.

While these regularities are quite in keeping with one's expectations, it is most gratifying to have these expectations met by the results from so many different urban area data sets. It may also be interesting to speculate as to the city-to-city differences which may be reflected in the variations in the magnitudes of these parameters. This question is currently being investigated.

CITY-TO-CITY COMPARISONS: REDUCED DRAM FORMULATION

In addition to the data sets enumerated above, there were several data sets consisting of somewhat less disaggregated variables. We note that the Washington, D.C. data set included in the above discussion, while lacking the percentage of developable land developed, did have the household income distribtuion. The data sets to be presented here had no data upon which a household disaggregation could be based. Further, neither the nineteenth century Philadelphia data nor the Izmit, Turkey data had any information on the classes of land use within the zones. As it was still desired to explore whether these cities could be calibrated, at least in part, a reduced formulation of DRAM was used. The results of parameter estimating for these data sets are shown in Table 4.

The results for 1850 Philadelphia and 1960 Hazleton, PA are unique amongst all those obtained to data. Philadelphia in 1850 was a "walking city". There was, for all intents and purposes no form of mechanized transportation in Philadelphia prior to 1850. The spatial patterns of residences and workplaces clearly show virtually no spatial separation of the two. For this data set we obtained α = -2.28, and β = 0.06, suggesting a declining number of residence opportunities with increasing travel from the workplace, and a very small propensity to travel. This latter might be considered, within the known range of trip lengths (i.e. well under one mile for the bulk of the trips), as an indifference to or slight preference for 2 or 3 city block worktrips versus living and working in the same location (Putman, 1977a).

	#Zones	α	β	Vac. Land	Per. Dev. Dev.	Res. Land
Colorado Springs	81	1.95	-0.18	0.06	1.57	0.68
Hazleton, PA	90	-1.90	0.15	-0.09	1.00	0.95
Philadelphia 1850	105	-2.28	0.06	0.46		
Philadelphia 1860	105	3.09	-5.05	0.89		
Philadelphia 1870	105	1.06	-2.77	0.72		
Philadelphia 1880	105	1.92	-3.42	0.97		
Izmit, Turkey	23	0.66	-0.32	0.82		

Only total zonal
area available in
these data sets.

Hazleton, PA is a relatively small city with a population of 60,000 and with a developed area that extends only a few miles in any direction. The initial 1960 Hazleton data set had 108 zones and yielded estimates of $\alpha = -0.30$ and $\beta = -1.13$, but there were a number of zones with zero population and employment. When these zones were deleted, the new Hazleton data set had 90 zones and yielded $\alpha = -1.90$ and $\beta = 0.15$ thus suggesting that most of the worktrips are so short that Hazleton either is, or is like, a "walking" city. Further investigation of these results is now in progress.

The remainder of the results in this group of data sets correspond exactly with the results from the full data sets above. It is rather reassuring to find that for urban areas as widely separated in time and space as Philadelphia in 1860 and Izmit in 1960, the trip function parameters of DRAM behave as expected. Further we note that the change in shape of the trip function from 1850 to 1860 in Philadelphia corresponds quite well with historical accounts of the development of the city (Putman, 1977a).

EFFECTS OF SPATIAL AGGREGATION

It was sometimes hypothesized, in early urban simulation model studies, that one might safely estimate the parameters of a model for a particular urban area at one level of spatial detail and subsequently use those same parameters in the model to make forecasts for that urban area at some other level of spatial detail. We have never held this opinion. We expect parameters for a particular model to exhibit consistent patterns of signs and some consistency as to magnitude from one urban data set to another. We do not expect the values of the parameters to be equal either across urban areas, or within a given urban area at differing levels of spatial detail.

The San Francisco data set was available at three different levels of detail, 291 zones, 98 zones, and 30 zones. DRAM parameters were estimated from each of these data sets, and yielded the results shown in Table 5.

These results clearly illustrate two major points. First, all the patterns (i.e. signs and relative magnitudes within data sets) are consistent with the results presented above, thus further emphasizing the generality and robustness of the model. Second, there is no immediately obvious rationale for the changes in parameters from on level of areal detail to another. Again we are forced, for the present, to call this a matter for further research.

For readers of the original DRAM article (Putman, 1977) an error must be called to their attention. Early work with the San Francisco data at the highly aggregated 30 zone level of detail gave both $\alpha < 0$ and $\beta < 0$, yielding the exponential decay function of Figure 1.b above. It was thought probable that at the 30 zone level the average zone-to-zone travel time was great enough for the peaked portion of the curve shown in Figure 1.a to be obscured. In fact the average zone-to-zone travel time to adjacent zones is 3 minutes in the 291 zone system. This increases to 8 minutes for the 30 zone system. The previously published values of α and β for the 30 zone San Francisco data were as follows:

	LOW INCOME			LOW MIDDLE			UPPER MIDDLE			UPPER INCOME		
Variable	30 z	98 z	291 z	30 z	98 z	291 z	30 z	98 z	291 z	30 z	98 z	291 z
C_{ij}^{α}	1.65	2.64	1.05	1.67	2.64	0.30	0.94	2.65	0.33	0.75	2.60	0.87
exp (βC_{ij})	-0.68	-0.28	-0.45	-0.58	-0.38	-0.26	-0.46	-0.40	-0.25	-0.28	-0.55	-0.31
Vac. Land	0.05	0.05	0.00	0.06	0.14	0.03	0.06	0.14	0.04	-0.02	0.12	0.03
Per. Dev. Dev.	0.49	1.32	1.24	0.46	1.32	1.03	0.51	1.32	1.04	0.68	1.63	1.11
Res. Land	0.49	0.39	0.77	0.46	0.40	0.83	0.50	0.42	0.67	0.55	0.27	0.63
Low Inc.	3.37	2.54	2.22	0.67	1.20	1.11	-1.80	0.47	-0.60	-2.04	-0.23	-1.72
Low Mid. Inc.	0.73	1.20	0.83	3.19	2.57	3.00	0.18	0.50	0.23	-1.98	-0.21	0.01
Upper Mid. Inc.	-0.78	-0.21	-1.10	0.58	0.51	1.00	2.10	2.51	3.05	0.40	1.19	0.01
Upper Inc.	0.50	-1.39	-1.77	0.63	-0.53	-0.54	-0.30	1.53	1.19	1.98	3.07	2.26

	LI	LMI	UMI	UI
α	-2.06	-1.75	-1.76	-1.64
β	-0.57	-0.72	-0.76	-0.48

Our more recent work with this data set, in revised form, yields the results shown in Table 5 which confirm to the initial hypotheses regarding the signs of α and β, but initiate cur earlier hypothesis concerning possible changes in the shape of the trip function with changes in the zone system used for analysis.

SIGNIFICANCE OF VARIABLES

Models of economic and/or social processes have traditionally used multiple regression techniques to estimate equation parameters. It is not possible to linearize the formulation of DRAM so that its parameters might be estimated by these techniques. The search technique used to estimate the DRAM parameters are no more difficult, theoretically or practically, than standard regression procedures. There are, however, two important difficulties with the use of these non-standard estimation procedures. First, there is no way to test the statistical significance of parameters estimated in this way, as their distributional properties are unknown. Second, no useful comparison can be made of the values of the likelihood criterion achieved with one data set to the values achieved with another.

Taking the second problem first, we recall that the likelihood criterion is used in parameter estimation for DRAM because it is a more sharply peaked curve near the optimum than is the R^2 curve. In the work done to date the value of R^2 has been calculated in every case, but not used as the criterion of best fit. Consequently, for comparison purposes we are able to report the values of R^2 shown in Table 6. These results are given along with two caveats. First, it must be noted that the best fit parameters based on the likelihood criterion are not necessarily the best fit parameters which would be derived using the R^2 criterion. We have observed several instances where improvements in the likelihood criterion were accompanied by worsening R^2 values. The second caveat is more of a reminder that these results are based on unadjusted use of the data bases as supplied to us by various agencies. The use of locally informed persons to modify the raw data files and "fine-tune" the parameter estimates would no doubt improve the results. Despite these two caveats, these results, for such a variety of urban areas, are quite a satisfactory beginning.

Before returning to the question of significance of parameters, we wish to note that there has been some exploration done into a likelihood related measure which could be used to compare the results from one data set to those from another. The problem may be illustrated by assuming that for the Low Income sector in City A the best fit value of the likelihood criterion is -3,540, while for Low Income City B a value of -14,332 is obtained for the best fit. The R^2 could well be 0.75 for City A and 0.85 for City B. These results might be considered to be contrary, confusing, or both. We are presently exploring the properties of what might be termed a best/worst likelihood achievement ratio.

AREA	LOW INCOME	LOW MID. INC.	UPPER MID. INC.	UPPER INCOME
Atlanta	0.813	0.864	0.922	0.866
Denver	0.832	0.728	0.749	0.593
Minneapolis	0.909	0.873	0.889	0.867
Philadelphia	0.762	0.584	0.426	0.693
Phoenix	0.721	0.609	0.592	0.654
San Diego	0.806	0.830	0.824	0.830
San Francisco	0.614	0.612	0.693	0.684
Seattle	0.732	0.837	0.816	0.862
Tulsa	0.767	0.600	0.574	0.687
Vancouver	0.795	0.712	0.887	0.847
Washington, D.C.	0.772	0.713	0.594	0.442
Colorado Springs	0.605			
Hazleton	0.868			
Philadelphia 1850	0.557			
Philadelphia 1860	0.817			
Philadelphia 1870	0.732			
Philadelphia 1880	0.730			
Izmit, Turkey	0.711			

The notion of this ratio is to measure the improvement that the best fit parameters, according to a likelihood criterion, provide above some base level "worst fit" parameters. While there are numerous concepts of worst fit, one which is computationally attractive is the equiprobable location fit. This says, that all zones are equally probable locations for residents and thus the region's residents are distributed equally to all zones. The best/worst achievement ratio is merely a measure of the extent to which the best fit parameters are an improvement over a uniform distribution assumption.

Now, we return to the question of the statistical significance of the best fit parameters. In a multiple regression problem it is possible to calculate a standard error of each parameter (coefficient) and use a t-table of F-table to test the hypothesis that the given value of the coefficient could be due to chance even though the "population" value of the coefficient was zero. It is on this basis that the automatic inclusion and deletion of variables in stepwise multiple regression programs are performed. Unfortunately none of this has any meaning in the search procedures used for models such as DRAM.

As a small beginning towards understanding the significance of variables and their parameters in DRAM, an extensive set of calibrations was run. In these runs, utilizing Upper Income households from the Minneapolis data set, various combinations of variables were tested. These results are presented in Table 7.

First note that the results in Table 7 are grouped. The first two columns cover changes in the household (employed resident) variables. The second group shows changes in the land variables, while the last

Table 7 : DRAM Calibrations, Combinations of Variables, Minneapolis

Upper Income Households - 1970

Variable	Household Tests		Land Variable Tests				Distance Tests			All
T^{α}_{ij}	2.155	2.217	-0.933	-0.458	-0.234	0.611	*	-0.212	*	1.906
$\exp\beta D_{ij}$	-1.282	-1.169	-0.876	-1.517	-0.859	-0.573	*	*	-0.200	-0.887
Res.Land	0.971	0.819	*	*	0.697	*	0.763	0.753	0.724	0.697
Percent Dev.	0.946	0.987	*	*	*	1.816	1.636	1.630	1.581	1.802
Vac.Land	-0.029	-0.061	*	0.073	*	*	-0.076	-0.076	-0.041	-0.017
LI	*	*	-0.480	-0.877	0.894	-1.836	-0.546	-0.546	-0.546	-0.245
LMI	*	*	2.848	2.536	1.041	2.147	-0.903	-0.904	-0.901	-0.643
UMI	*	*	2.639	1.908	-0.227	2.042	-0.423	-0.422	-0.412	-0.649
UI	*	2.250	5.378	4.712	3.694	5.035	2.522	2.522	2.492	2.737
R^2	0.739	0.870	0.509	0.530	0.819	0.550	0.861	0.862	0.865	0.859
MaxL	-16382.	-9120.	-31669.	-30923.	-11644.	-29420.	-8543	-8012.	-8554.	-7976.
Best/Worst	.9721	.9845	.9461	.9474	.9802	.9499	.9855	.9864	.9854	.9864

group shows changes in the distance or trip function. The last column in the table gives the results with all variables included, as were reported above.

The most obvious conclusion is that signs and values of the model's parameters do change with the inclusion or exclusion of different variables. The question, of course, is whether these changes are important. The R^2 values are especially useless here as they actually run contrary to the likelihood criterion. Thus the R^2 for "All" variables is less than that for several of the partial variable sets, even though the best (maximum) value of the likelihood criterion is achieved when all variables are included. For comparison purposes, the best/worst achievement ratio is given in the last row of the table. Note that the uniform distribution value of -587,653 for the likelihood criterion was calculated. Many of these values are rather close to one another. This may mean that the measure does not discriminate well, or that there isn't much difference in the model fits for certain of the variable combinations.

All this again raises the question of what sorts of colinearities may exist between these variables? For the present we really don't know. There clearly are some problems which need to be resolved here. This, like several of the questions raised above, only serves to point out where work needs to be done, rather than to the problem's solution,

SUMMARY

This second of two papers has presented the results of applying the techniques described in the first paper to a number of different data sets. The conclusions may be simply stated.

a) There is eveidence, in the form of theoretically consistent parameter estimates over a wide variety of city types, to suggest that DRAM and similar models offer a robust and reliable macro descriptive model of urban residential location.

b) The level of zonal aggregation or disaggregation of a data set does affect the parameter values of a model estimated from the data sets. There is inadequate evidence to suggest exactly how this works.

c) There are several unresolved questions, when using gradient search techniques for parameter estimation of these models, as to statistical significance of parameter values, and measures of goodness of fit which allow proper comparison of one set of parameter results to another.

References

Putman, S. (1977) "Calibrating a Disaggregated Residential Allocation Model - DRAM" London Papers in Regional Science Vol. 7 pp. 108-124.

Putman, S. (1977a) "Calibrating a Residential Location Model for Nineteenth Century Philadelphia" Environment and Planning A Vol. 9 pp. 449-460.

Putman, S. (1978) Urban Residential Location Models: Evolution and Development Studies in Applied Regional Science, Martinus Nijhoff, Publications, Leiden Netherlands (forthcoming).

Wilson, A. (1970) Entropy in Urban and Regional Modelling Pion, London.

Regional Studies, Vol. 4, pp 473–490. Pergamon Press 1970 Printed in Great Britain

SfB Aa9
UDC 711.132.001.57
(74–5)

Developing and Testing an Intra-Regional Model

STEPHEN H. PUTMAN*

Department of City and Regional Planning, University of Pennsylvania, Philadelphia, Pa. 19104

(Received 31 March 1970; in revised form 26 June 1970)

PUTMAN S. H. (1970) Developing and testing an intra-regional model, *Reg. Studies* 4, 473–490. As part of the evaluation of the economic impact of alternative transport systems proposed for the Northeast Corridor region of the United States it was necessary to develop a model to forecast these impacts.

The model described here was developed by integrating portions of the concepts of economic-base, intersectoral input-output, and spatial accessibility. Forecasts are developed, by the model, of employment (by ten sectors), population, personal income (by six classes), land value, and land use, for 131 county-sized areas.

The theoretical development, parameter estimation, and some testing, of the model are described, and directions for future work are suggested.

Northeast Corridor of the U.S.
Multiple regression/correlation
Validation and testing of models
Spatial accessibility
Inter-sectoral flows
Transport sensitivity
Solution of model

I. INTRODUCTION

As AN integral part of the U.S. Department of Transportation's study of alternative transportation proposals for the Northeast Corridor of the United States an extensive series of studies of economic impacts was undertaken (CONSAD 1967, 1968, 1969a, 1969b, 1970). The purpose of these studies was to lend a further dimension to the evaluation of the transportation alternatives by including in the evaluation consideration of economic effects which might be induced by new or altered transport systems.

Initially an interrelated set of three models was proposed to forecast these induced effects. The first model was a large-scale (i.e., the whole region considered as divided into three sub-areas) econometric model (CROW, 1969). The second model which was to be in the system was an inter-regional input-output model, based upon the Leontief–Strout formulation (LEONTIEF and STROUT, 1963). This model was to capture inter-sectoral relationships and operate at a somewhat more detailed level of areal detail, the region here being considered as divided into ten sub-areas (CONSAD, 1968, 1969a). The final model in the system was an intra-regional location model which was to capture inter-area relationships at an even more detailed areal level, the region here being considered as divided into 131 sub-areas the size of counties (GLANCY, 1965; PUTMAN, 1967a).

While the conceptual interrelationships of the proposed system of models were rather complex, they were operationally straightforward. The forecasts from the econometric model became the predetermined control-totals for the inter-regional input-output model. Similarly the forecasts from the input-output model became the control-totals for the intra-regional location model.

After considerable investigation of the models and their interrelationships it was determined that the input-output model's contribution to the system was, at best, marginal. Consequently the inter-regional input-output model was deleted from the model system. Thus the forecasts from the econometric model became the predetermined control-totals for the intra-regional loca-

---

* Research Investigator—The principal development of this model was accomplished by the author while managing Contract No C-101-66 (Neg.) from the U.S. Department of Transportation to CONSAD Research Corporation

tion model. This paper is specifically devoted to describing the development and testing of the intra-regional location model called INTRA-I.

II. THE BASIC STRUCTURE OF THE MODEL

The specific purpose of this model is to forecast the response of the spatial distribution of economic and demographic activity in a region to changes in intra-regional transport facilities. This model proposes to achieve that goal by first producing independent forecasts of this response and then subjecting the sums of these to constraints imposed by the forecasts of the Crow econometric model. Levels of economic activity will be measured by employment levels and personal income, while demographic activity will be measured by sub-area resident population levels.

The structure of this model represents an attempt to draw together, within practical limits, aspects of economic-base-study models, sectoral input-output relationships, and the concept of accessibility. To a certain extent this attempt was successful, though further work, as shall become clear later in this article, remains to be done.

Building first upon the economic-base concepts (Isard, 1970; Lane, 1966; Nourse, 1968; Ullman and Dacey, 1960), the model structure discriminates between two principal classes of employment. Those types of employment which are primarily local population-serving, e.g. services or retail trade, are considered as non-basic. Those types of employment which primarily serve other industries, both local and non-local, e.g. mining or manufacturing, are considered as basic. An integral part of the structure of this model is the simplifying assumption that there is sufficient time lag in the real-world dynamics to allow non-simultaneous consideration of the basic employment *vis-à-vis* population and non-basic employment in the model. It is not now clear exactly how damaging this assumption may be, though other work (Putman, 1969; CONSAD, 1970) indicates that increased explanation of

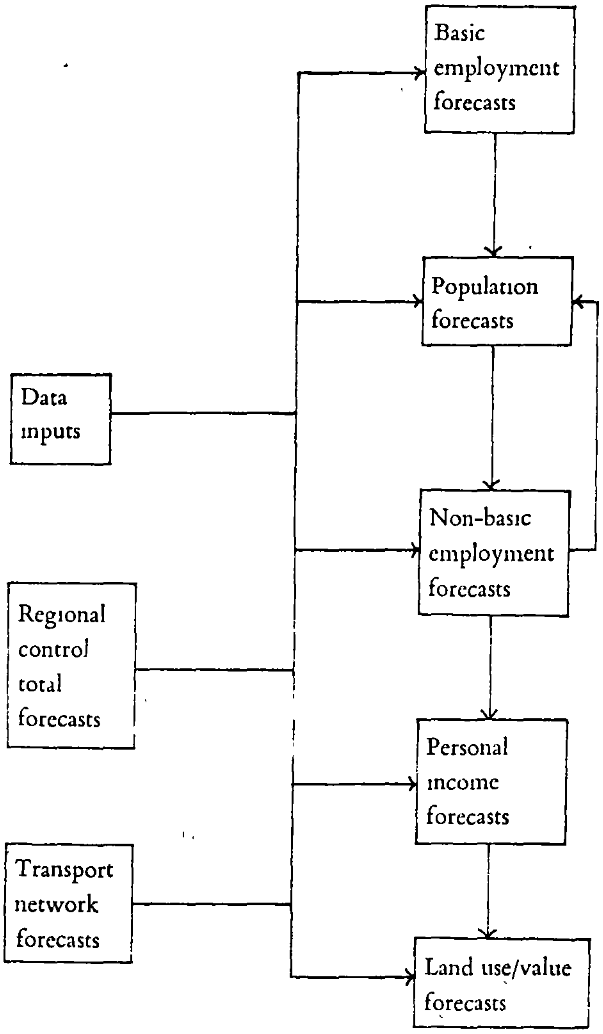


Fig. 1. Overall INTRA-I model structure.

variance may be obtained from a more sophisticated simultaneous equation formulation (at considerably increased computer cost). Further, it is asserted that any lag which may exist in the interrelationship between population and population-serving employment is short enough to allow their simultaneous consideration. Personal income should, in theory, be included in this simultaneous portion of the model; but, due to anticipated difficulty of forecasting it, plus the consequent likely unreliability of the forecasts, it was kept as a separate component of the model. Finally, it should be noted that significant consideration was given to likely data sources and the constraints thereby imposed during the preliminary development of this structure. The basic model structure is shown in Fig. 1.

III. THE BASIC-EMPLOYMENT EQUATIONS

The following four types of employment comprise the basic-employment class in the model: (a) Agriculture, Forestry, and Fisheries; (b) Mining, (c) Manufacturing; and (d) Transportation, Communication, and Public Utilities. A more detailed description of the industries included in these types may be found in the document from which the data were obtained (U.S. Department of Commerce, 1960, 1965).

In each sub-area the level of each type of basic employment is postulated to be a function of: (1) the spatial distribution of the suppliers of its inputs; (2) the spatial distribution of the consumers (markets) for its outputs; (3) the existing pattern of inter-industry relationships; (4) the existing pattern of transport facilities; (5) the level of that type of employment in that area in the prior time period; (6) the total of all types of employment in that area in the prior time period, and (7) gross population density in that area in the prior time period (Greenhut, 1956; Harris, 1954; Hoover, 1948, 1967; Hamburg and Lathrop, 1965; Thomas, 1969). The first four of these factors are measures of the ability of a given industry in a given area to obtain its required inputs and sell its outputs. The fifth variable represents the obtaining situation with regard to that type of industry in the given area and as such represents the culmination of many phenomena which produced that situation. The last two variables are measures of the area itself; e.g. a center of employment, or not; a residential or non-residential area.

The concepts of accessibility and inter-sectoral flows were included in two accessibility measures which were then developed. The first was a measure of accessibility to inputs (suppliers). This is of the form:

$$A^s_{i,g,t-1} = \sum_j \left[\frac{\sum_h (\alpha_{hg} p_h E_{j,h,t-1})}{Z^f_{i,j,t-1}}\right] \qquad (1)$$

where:

$A^s_{i,g,t-1}$ = a measure of the accessibility at time $t-1$ of area i to inputs required by industry g

$E_{j,h,t-1}$ = amount of employment in industry h in area j at time $t-1$

p_h = productivity ratio (output/employee) for industry h. (Ideally this would be area and time specific, but such data are unavailable. National figures are used in the model.)

α_{hg} = input–output coefficient—the number of units of inputs from sector h required in the production of one unit of output from sector g (O.B.E., 1965)

$Z^f_{i,j,t-1}$ = freight "impedance" between areas i and j at time $t-1$. (Ideally this would be commodity (sector) specific, but such data are unavailable.)

The numerator of this term can be seen to consist of the magnitude of supply of potential inputs ($p_h E_{j,h,t-1}$) available in area j times the relative "importance", by sector, α_{hg}, of those inputs to sector g's production. The denominator is a measure of the difficulty of obtaining those inputs. Note that price of inputs at their source is assumed constant, though if such data were available, the measure could be modified to include this such that an increase in price would result in a decrease in accessibility.

The second accessibility measure is a measure of market access. This is of a similar form:

$$A^m_{i,g,t-1} = \sum_j \left[\frac{\sum_h (\alpha_{gh} p_h E_{j,h,t-1})}{Z^f_{i,j,t-1}}\right] \qquad (2)$$

The difference between (1) and (2) can be seen to be the use of α_{gh} in the latter, representing the number of units of output from sector g used as input by sector h to produce one unit of output from sector h. Thus the numerator of (2) can be seen to be a measure of potential demand for the outputs of sector g.

The freight impedance term $Z^f_{i,j,t-1}$ is actually a composite of several transport facility variables. It is

$$Z^f_{i,j,t-1} = \sum_h [\beta_{k,t-1}(C^k_{i,j,t-1} + \gamma_k T^k_{i,j,t-1})] \qquad (3)$$

where:

$\beta_{k,t-1}$ = mode-split proportion representing the percentage of freight shipped via mode k at time $t-1$ (Ideally this should be both sector specific and i-j pair specific; however, such data were unavailable.)

$C^{k}_{i,j,t-1}$ = cost-per-unit to ship from area i to area j via mode k at time $t-1$. (Clearly this, too, should be sector and i-j pair specific.)

$T^{k}_{i,j,t-1}$ = time to ship from area i to area j via mode k at time $t-1$ (which should also be sector and i-j pair specific).

γ_k = implicit dollar value per unit time on mode k (which should be sector specific)

Thus the impedance variable used is a somewhat crude representation of the "cost" (both out-of-pocket and implicit) of shipping. This is one of the least accurate variables in the analysis and clearly calls for further analysis at such time that better data become available. Alternative formulations of the impedance variable were tested (CONSAD, 1969a, 1969b) and rejected.

These accessibility measures and the other variables were combined in a linear function as below:

$$E_{i,g,t} = a_1\,A^{s}_{i,g,t-1} + a_2\,A^{m}_{i,g,t-1} + a_3\,E_{i,g,t-1} \qquad (4)$$

$$+a_4\,E^{T'}_{i,t-1} + a_5\,(P_{i,t-1}/L_i) + a_6$$

where:

$E_{i,g,t}$ = employment of industrial sector g in area i at time t

$A^{s}_{i,g,t-1}$ = a measure of the accessibility at time $t-1$ of area i to inputs required by industry g

$A^{m}_{i,g,t-1}$ = a measure of the accessibility at time $t-1$ of area i to markets which consume the outputs of industry g

$E^{T'}_{i,t-1}$ = total employment minus industry g employment at time $t-1$ in area i

$P_{i,t-1}$ = population residing in area i at time $t-1$

L_i = land area of area i

The results of these estimates were as follows, using a single-equation least squares regression procedure:

Industry Sector 1: Agriculture, Forestry, Fisheries

$$E_{i,1,t} = \underset{(-4{\cdot}80)}{-0{\cdot}0269\,A^{s}_{i,1,t-1}} + \underset{(4{\cdot}96)}{0{\cdot}0347\,A^{m}_{i,1,t-1}}$$

$$+\underset{(30{\cdot}38)}{1{\cdot}0024\;E_{i,1,t-1}} + \underset{(3{\cdot}56)}{0{\cdot}0003\;E^{T'}_{i,t-1}}$$

$$\underset{(-5{\cdot}83)}{-0{\cdot}0134\,(P_{i,t-1}/L_i)} - 26{\cdot}739$$

$$r^2 = 0{\cdot}9219 \qquad F = 295{\cdot}2$$

Industry Sector 2: Mining

$$E_{i,2,t} = \underset{(3.36)}{0{\cdot}0118\,A^{s}_{i,2,t-1}} - \underset{(-3{\cdot}68)}{0{\cdot}0184\,A^{m}_{i,2,t-1}}$$

$$+\underset{(12{\cdot}45)}{0{\cdot}8717\,E_{i,2,t-1}} + \underset{(0{\cdot}31)}{0{\cdot}00005\,E^{T'}_{i,t-1}}$$

$$+\underset{(2{\cdot}77)}{0{\cdot}0111\,(P_{i,t-1}/L_i)} + 10{\cdot}504$$

$$r^2 = 0{\cdot}7358 \qquad F = 69{\cdot}6$$

Industry Sector 4: Manufacturing

$$E_{i,4,t} = \underset{(-0{\cdot}11)}{-0{\cdot}1090\,A^{s}_{i,4,t-1}} + \underset{(0{\cdot}16)}{0{\cdot}1501\,A^{m}_{i,4,t-1}}$$

$$+\underset{(45{\cdot}94)}{0{\cdot}9187\,E_{i,4,t-1}} + \underset{(2{\cdot}67)}{0{\cdot}0294\,E^{T'}_{i,t-1}}$$

$$\underset{(-3{\cdot}24)}{-0{\cdot}4546\,(P_{i,t-1}/L_i)} - 5147{\cdot}0$$

$$r^2 = 0{\cdot}9930 \qquad F = 3534{\cdot}7$$

Industry Sector 5: Transportation, Communication, Public Utilities

$$E_{i,5,t} = \underset{(-0{\cdot}17)}{-0{\cdot}0340\,A^{s}_{i,5,t-1}} + \underset{(0{\cdot}31)}{0{\cdot}0491\,A^{m}_{i,5,t-1}}$$

$$+\underset{(7{\cdot}88)}{0{\cdot}6855\,E_{i,5,t-1}} + \underset{(2{\cdot}20)}{0{\cdot}0132\,E^{T'}_{i,t-1}}$$

$$+\underset{(2{\cdot}23)}{0{\cdot}1744\,(P_{i,t-1}/L_i)} - 1771{\cdot}2$$

$$r^2 = 0{\cdot}9706 \qquad F = 824{\cdot}7$$

It should be noted that the numbers in parentheses beneath the regression coefficients are calculations of the t-statistic. For these estimates, where $n = 131$, a value of $t > 2{\cdot}34$ implies significance at the 1 per cent level or better, while $t > 1{\cdot}65$ implies significance at the 5 per cent level. The r^2 value given in each case is the coefficient of determination for the regression equation, and the F value is calculated for the ratio of explained-to-unexplained variance where for these data, $F > 4{\cdot}6$ implies significance at the 0·1 per cent level or better.

For all of these industrial sectors the correlation between either of the access variables and the dependent variables is positive. Further, either of these variables, regressed against the dependent variable, is statistically significant at the 1 per cent level. When both access variables are forced to enter the estimating equation, one of them always takes a negative coefficient and, for Sectors 4 and 5, both of them become statistically not significant. One explanation for the sign-changing behavior is the appearance of what is called the "suppressant variable" phenomenon

(PETERS and SUMMERS, 1968). This concept asserts that there is some aspect (component) of the variation, in this case, in $A^s_{i,g,t-1}$, for example, which is not present in $E_{i,g,t}$. When $A^m_{i,g,t-1}$ enters the regression equation with a negative coefficient, it is effectively subtracting out this undesired component of variation. Often, when this phenomenon occurs, both variables retain their statistical significance. Another possible cause of this sign-changing behavior is the presence of extreme multicollinearity between the two access variables (JOHNSTON, 1963). In this case the common variation of the two variables is so high as to render them both, when both are forced to enter the regression equation, not statistically significant, thus making the coefficients essentially meaningless in any case. This appears to be the situation found in Sectors 4 and 5.

When Sector 4, Manufacturing, is re-estimated with only one of the access variables, the supply accessibility variable appears most significant, the resulting estimated equation being:

$$\begin{aligned} E_{i,4,t} = {} & \underset{(3\ 35)}{0{\cdot}0363\, A^s_{i,4,t-1}} + \underset{(49{\cdot}52)}{0{\cdot}9197\, E_{i,4,t-1}} \\ & + \underset{(2\ 66)}{0{\cdot}0288\, E^{T'}_{i,t-1}} - \underset{(-3{\cdot}23)}{0{\cdot}4506\,(P_{i,t-1}/L_i)} - 5354{\cdot}5 \end{aligned}$$

$$r^2 = 0{\cdot}9930 \qquad F = 4452{\cdot}7$$

It is worth noting that only the coefficient of the supply accessibility shows any substantial change from the previous estimation of the equation, thus indicating that the multicollinearity problem is confined to the two accessibility variables

When Sector 5, Transportation *et al*, is re-estimated with only one of the access variables, the market accessibility variable appears most significant, the resulting estimated equation being

$$\begin{aligned} E_{i,5,t} + {} & \underset{(1\ 90)}{0{\cdot}0179\, A^m_{i,5,t-1}} + \underset{(7{\cdot}97)}{0{\cdot}6862\, E_{i,5,t-1}} \\ & + \underset{(2\ 15)}{0{\cdot}0131\, E^{T'}_{i,t-1}} + \underset{(2{\cdot}28)}{0{\cdot}1760\,(P_{i,t-1}/L_i)} - 1687{\cdot}4 \end{aligned}$$

$$r^2 = 0{\cdot}9706 \qquad F = 1038{\cdot}8$$

Again it can be seen that only the coefficient of the access variable shows any substantial change from the previous estimation of the equation. The complete equations (i.e. both access variables included) were used in the model runs described below. Further research effort will attempt to determine what effect is produced by using the simplified equations.

The coefficients of the other variables are basically what would be expected. The coefficients of prior employment in the same sector and area, $E_{i,g,t-1}$, are all positive and near unity, perhaps indicating slightly decreasing "returns-to-scale" of employment increases in areas where there are prior large quantities of that type of employment. The coefficients of prior total employment in the same area are also positive, thus indicating employment growth at centers of employment (perhaps agglomeration economies). The coefficients of the population density variable are somewhat less consistent. For Sectors 1 and 4 (Agriculture *et al* and Manufacturing) the negative coefficient implies declining employment in areas of higher gross population density. For Sector 5 (Transportation *et al.*) the positive coefficient implying growth of these types of employment in densely populated areas is quite reasonable. The same positive coefficient for Sector 2 (Mining) does not make sense in its implication that this type of employment will grow in densely populated areas. Further investigation of this result is clearly desirable. One possibility to be investigated is that this positive coefficient stems from the variable playing the role of suppressant variable for both $E_{i,2,t-1}$ and $E^{T'}_{i,t-1}$, both of which enter the equation with positive coefficients.

IV. THE NON-BASIC EMPLOYMENT EQUATION

The following four types of employment comprise the wholly non-basic employment class in the model: (a) Retail Trade; (b) Finance, Insurance, Real Estate; (c) Services; and (d) Government.

In each sub-area, at a given time, the amount of each type of wholly non-basic employment is postulated to be a function of (1) the total population in that sub-area at that time, which is primarily a measure of the market for the output of the employment, (2) the total employment in the sub-area at that time, which is also a "market" measure; and (3) the amount of that particular type of employment which was in the sub-area in the prior time period and which, as described above, represents the obtaining situation with regard to that type of industry in the given area.

Implicit in this postulated function is the assumption that for sub-areas at the geographical scale being considered here (i.e. counties), crossing of boundaries for the purpose of getting to these activities is rather infrequent, and what little crossing there is tends to be equalized from sub-area to sub-area. This assumption would certainly not be true for smaller geographical sub-areas and is probably not valid for a few of the areas in this study. The results of the estima-ing equations described below tend to confirm its validity for this particular case.

These variables were combined in a linear function as below.

$$E_{i,g,t} = a_1 E_{i,g,t-1} + a_2 P_{i,t} + a_3 E^T_{i,t} + a_4 \qquad (5)$$

where:

$E_{i,g,t}$ = employment of industrial sector g in area i at time t
$P_{i,t}$ = population of area i at time t
$E^T_{i,t}$ = total employment in area i at time t

Implicit in this function is an element of simultaneity of response of the level of non-basic employment to population and total employment. This is probably not quite true in respect to real world behavior. The data, however, on which the estimation of this model's parameters must be based, allow a choice between simultaneity or a 5-year lag. Though the 5-year lag was considered to be reasonable for basic employment, it was considered unreasonable for non-basic employment. Consequently the assumption of simultaneity was accepted in preference to assumption of a 5-year lag

The results of estimating the equations were as follows:

Industry Sector 7: Retail Trade

$$\begin{aligned} E_{i,7,t} = {} & \underset{(11\cdot89)}{0\cdot6969}\, E_{i,7,t-1} + \underset{(8\cdot44)}{0\cdot0089}\, P_{i,t} \\ & + \underset{(1\cdot65)}{0\cdot0168}\, E^T_{i,t} + 265\cdot59 \end{aligned}$$

$r^2 = 0\cdot9945 \qquad F = 7642\cdot5$

Industry Sector 8: Finance, Insurance, Real Estate

$$\begin{aligned} E_{i,8,t} = {} & \underset{(56\cdot60)}{1\cdot1320}\, E_{i,8,t-1} + \underset{(20\cdot34)}{0\cdot0108}\, P_{i,t} \\ & - \underset{(-10\cdot69)}{0\cdot0296}\, E^T_{i,t} - 185\cdot33 \end{aligned}$$

$r^2 = 0\cdot9979 \qquad F = 200\cdot19$

Industry Sector 9. Services

$$\begin{aligned} E_{i,9,t} = {} & \underset{(13\cdot96)}{0\cdot8375}\, E_{i,9,t-1} + \underset{(5\cdot82)}{0\cdot0086}\, P_{i,t} \\ & + \underset{(5\cdot04)}{0\cdot0458}\, E^T_{i,t} - 786\cdot46 \end{aligned}$$

$r^2 = 0\cdot9944 \qquad F = 7579\cdot0$

Industry Sector 10: Government

$$\begin{aligned} E_{i,10,t} = {} & \underset{(14\cdot84)}{0\cdot5937}\, E_{i,10,t-1} + \underset{(0\cdot69)}{0\cdot0021}\, P_{i,t} \\ & + \underset{(13\cdot10)}{0\cdot0786}\, E^T_{i,t} + 637\cdot72 \end{aligned}$$

$r^2 = 0\cdot9135 \qquad F = 446\cdot99$

With only one exception these results are completely in accord with the expectations postulated above. As was the case with the basic employment equation estimates, the coefficients of prior employment of the same sector in the same area are all positive and near unity, again, perhaps indicating slightly decreasing "returns-to-scale". The coefficients of current population in the same area are all positive and reflective of the surrogate "market" measure hypothesis. The lack of statistical significance of this coefficient for Sector 10 (Government) is apparently not due to multicollinearity (simple correlation with the other independent variables is always less than 0·79). It is possible that the market hypothesis is less relevant for this sector than for the others, with respect to population. Except for Sector 8 (Finance, Insurance, Real Estate) the coefficients of local total employment are also, according to the market measure hypothesis, positive. The reason for the negative coefficient obtained for this variable for Sector 8 is not at all obvious and warrants further research.

Two remaining types of employment are neither wholly basic nor wholly non-basic These two types are: (a) Contract Construction, and (b) Wholesale Trade.

The level of Sector 3 (Contract Construction) employment in a given area was postulated to be a function of. (1) total accessibility to markets and supplies, (2) employment and population densities, indicating the general character of the area; (3) the percentage change in total employment in the area, indicating (under the assumption that large employment increases are causes of increased building construction of all types) the potential local demand for new construction; and (4) the obtaining distribution of contract construction employment from the prior time period. These variables were combined in linear form, and the coefficients of the equation were estimated. Later tests of the equation in the model showed, however, that the presence of the percentage change of employment variable in the equation could lead, under certain special circumstances, to undamped oscillation of the Sector 3 forecast in the model. Consequently the equation used, deleting the offending term, was:

$$\begin{aligned} E_{i,3,t} = {} & a_1 A^T_{i,t-1} + a_2 E_{i,3,t-1} + a_3 (E_{i,3,t-1}/L_i) \qquad (6) \\ & + a_4 (P_{i,t-1}/L_i) + a_5 \end{aligned}$$

where:

$E_{i,3,t}$ = employment of Sector 3 in area i at time t
L_i = land area of area i
$P_{i,t-1}$ = population residing in area i at time $t-1$

$A^T_{i,t-1}$ = total accessibility to employment of area i at time $t-1$

and where:

$$A^T_{i,t-1} = \sum_j \left(\frac{\sum_g p_g E_{j,g,t-1}}{Z^f_{i,j,t-1}} \right)$$

where:

$Z^f_{i,j,t-1}$ = freight "impedance" between areas i and j at time $t-1$ as defined above

p_g = productivity ratio (output/employee) for industry g

The results of estimating the equation were as follows:

$$E_{i,3,t} = \underset{(1{\cdot}25)}{0{\cdot}0005\, A^T_{i,t-1}} + \underset{(35{\cdot}70)}{1{\cdot}1283\, E_{i,3,t-1}}$$

$$- \underset{(-6{\cdot}97)}{7{\cdot}6630\,(E_{i,3,t-1}/L_i)} + \underset{(2{\cdot}22)}{0{\cdot}0820\,(P_{i,t-1}/L_i)}$$

$$-405{\cdot}79$$

$$r^2 = 0{\cdot}9702 \qquad F = 102{\cdot}44$$

The negative coefficient of the gross-employment-density variable is unexpected. There is no obvious statistical reason for this result, e.g. multicollinearity or suppressant variable effect. One possible theoretical explanation has to do with the joint effect of the nature of the industry sector and the data source. Typically construction companies, while active in highly urban areas, have their offices (equipment storage yards, too) outside the urban areas. Since the data collected on persons employed by the firm considers all the employment to occur at the "offices", the negative coefficient may result from this mismatch between actual and reported workplaces

The level of Sector 6 (Wholesale Trade) employment in a given area is postulated to be determined by the same factors as those which affect the wholly non-basic employment (Sectors 7, 8, 9, 10) plus simultaneous accessibility to total employment (markets and supplies). The function estimated was:

$$E_{i,6,t} = a_1 A^T_{i,t} + a_2 E_{i,6,t-1} + a_3 P_{i,t} + a_4 E^T_{i,t} + a_5 \quad (7)$$

The results of the estimation were:

$$E_{i,6,t} = \underset{(1{\cdot}86)}{0{\cdot}0013\, A^T_{i,t}} + \underset{(11{\cdot}24)}{0{\cdot}5283\, E_{i,6,t-1}} - \underset{(-18{\cdot}44)}{0{\cdot}0129\, P_{i,t}}$$

$$+ \underset{(17{\cdot}12)}{0{\cdot}0685\, E^T_{i,t}} - 1477{\cdot}8$$

$$r^2 = 0{\cdot}9892 \qquad F = 2885{\cdot}9$$

In the case of these coefficients the only one which does not have the expected sign is that of population density in the given area i at time t. Again, while it is not clear why this should be the case, it is not totally unreasonable in that it may indicate a tendency for wholesale trade establishments (warehouses?) to locate away from densely populated residential areas. As was the case with other unexpected regression coefficients, there is room here for further research.

V. THE POPULATION EQUATIONS

The resident population of a given area i at time t is, by definition, the sum of: (a) the population of the area at time $t-1$; (b) the number of births in the given area since time $t-1$; (c) minus the number of deaths in the given area since time $t-1$; and (d) plus the net migration into the given area since time $t-1$. That is,

$$P_{i,t} = P_{i,t-1} + B_{i,\Delta t} - D_{i,\Delta t} + M_{i,\Delta t} \qquad (8)$$

where:

$P_{i,t}$ = resident population of area i at time t

$B_{i,\Delta t}$ = number of live births in area i from time $t-1$ to time t

$D_{i,\Delta t}$ = number of deaths of residents in area i from time $t-1$ to time t

$M_{i,\Delta t}$ = net migration into (or, if negative, out of) area i from time $t-1$ to time t

Each of these components of population change is itself forecast by a separate equation. Considerable conjecture but little empirical work has been published regarding determinants of birth and death rates in sub-areas of the size with which we are here concerned (N.B.E.R., 1960). After surveying what has been written, numbers of births and deaths by sub-area were postulated as being a function of: (1) per-capita personal income in the sub-area at the beginning of the relevant time period; (2) gross population density (representing urbanization) of the sub-area at the beginning of the relevant time period; and (3) number of births or deaths in the sub-area during the prior time period. These factors were combined in linear functions. These were:

$$B_{i,\Delta t} = a_1 (Y_{i,t-1}/P_{i,t-1}) + a_2 (P_{i,t-1}/L_i) + a_3 B_{i,\Delta t'} + a_4 \qquad (9)$$

and

$$D_{i,\Delta t} = a_1 (Y_{i,t-1}/P_{i,t-1}) + a_2 (P_{i,t-1}/L_i) + a_3 D_{i,\Delta t'} + a_4 \qquad (10)$$

where:

$Y_{i,t-1}$ = total personal income received by residents of area i at time $t-1$

$B_{i,\Delta t'}$ = number of live births in area i from time $t-2$ to time $t-1$

$D_{i,\Delta t'}$ = number of deaths of residents in area i from time $t-2$ to time $t-1$.

It was expected that for both births and deaths, the coefficient of the prior period's births or deaths would be positive and near unity, these being basically trend variables in both instances. In the births equation it was expected that the coefficient of per-capita personal income would be negative, reflecting lower birth rates in families of higher socio-economic status. Similarly, reflecting lower birth rates in urban areas, the coefficient of population density was also expected to be negative. The results of estimating the equation were:

$$B_{i,\Delta t} = \underset{(0{\cdot}63)}{1{\cdot}2813}(Y_{i,t-1}/P_{i,t-1}) - \underset{-(1{\cdot}53)}{0{\cdot}1878}(P_{i,t-1}/L_i)$$

$$+\underset{(41{\cdot}41)}{0{\cdot}8681}\,B_{i,\Delta t'} + 225{\cdot}74$$

$$r^2 = 0{\cdot}9696 \qquad F = 1256{\cdot}06$$

It can be seen that only the coefficient of the per-capita income variable was not as expected. The positive coefficient of per-capita income may be due to its being an inadequate (or incorrect) measure of socio-economic status. For example, a "family" with $20K/year income and two persons will have a very different per-capita income from a "family" with $20K/year income and five persons. The lack of statistical significance of this variable may or may not be due to this problem, also.

In the equation for deaths, it was expected that the degree of urbanization would be negatively related to deaths, because of (on the average) better living conditions and diet, and more adequate and accessible medical facilities in urban areas. It was not clear whether higher socio-economic status would be associated with fewer deaths because of better medical care, or more deaths because of increased "work pressure" The results of estimating the equation were:

$$D_{i,\Delta t} = \underset{(1{\cdot}73)}{1{\cdot}3319}(Y_{i,t-1}/P_{i,t-1}) - \underset{(-1{\cdot}89)}{0{\cdot}1095}(P_{i,t-1}/L_i)$$

$$+\underset{(49{\cdot}23)}{0{\cdot}9106}\,D_{i,\Delta t'} - 1227{\cdot}79$$

$$r^2 = 0{\cdot}9842 \qquad F = 2443{\cdot}9$$

Here, the coefficients repeated the sign pattern of those for the births equation. The positive coefficient of per-capita income is less surprising than it was in the births equation, but both of these phenomena warrant future research.

Net migration is the most difficult-to-forecast component of population change by sub-area. Also, and not coincidentally, it is the one for which the least data are available. There are two principal methods by which migration forecasts might be attempted in this model. The first of these would be to take a rather broad macro-behavioral approach and use some variation on gravity-interaction formulation (Anderson, 1955; Isard, 1960; Olsson, 1965, Stouffer, 1940). This approach was considered to be too aggregative in that it subsumes several phenomena into its formulation. Further, work with these formulations was precluded by the lack of data describing gross area-to-area migration flows. The second approach, the one which was taken, consists of identifying those measures (economic or socio-economic) which are causes (or surrogate measures thereof) of migration. A function may then be postulated, including measures of both "in" and "out" migration causes, which will explain (forecast) net area-to-area migration.

One such major determinant of migration is the availability of employment opportunities and a tendency for workers to migrate from areas of few opportunities to areas of many opportunities (Blanco, 1963; Raiman, 1962; Sjaastad, 1961). It is important to note that it is not necessary that there actually be greater numbers of opportunities at the destination area; rather, it is only necessary that the migrating individual expects them to be there. Another factor found to explain a significant portion of the variance of net migration was net change in population due to birth and death rates, i.e. "natural increase". Along with this variable, the percentage change in median income over a given time period in a particular area was also found to be related to the area's net migration (Lowry, 1966).

Given these theories and empirical results, it was possible to postulate a migration-forecasting equation for the model. Net migration into or out of a given area i during the period $t-1$ to t was postulated to be a function of: (1) population in the area at the beginning of the period; (2) total employment in the area at the end of the period, (3) natural increase of population in the area over the same period of $t-1$ to t; and (4) accessibility of the area to employment (opportunities) at time t. It must first be noted that both total employment in the area at time t (the end of the period) and accessibility to employment at time t are used, as was the case in certain of the employment equations, because of a preference of assuming simultaneity to that of assuming a 5-year lag. The function estimated for migration forecasting was.

$$M_{i,\Delta t} = a_1 P_{i,t-1} + a_2 E^T_{i,t} + a_3(B_{i,\Delta t} - D_{i,\Delta t}) + a_4 \phi_{i,t} + a_5 \qquad (11)$$

where:

$M_{i,\Delta t}$ = net migration into (or, if negative, out of) area i from time $t-1$ to time t

$P_{i,t-1}$ = population residing in area i at time $t-1$

$E^T_{i,t}$ = total employment in area i at time t

$B_{i,\Delta t} - D_{i,\Delta t}$ = "natural increase" in population of area i from time $t-1$ to time t

$\phi_{i,t}$ = accessible opportunities for employment of residents of area i at time t.

The measure of accessible opportunities is calculated as follows:

$$\phi_{i,t} = \sum_j (E^T_{j,t}/Z^P_{i,j,t}) \qquad (12)$$

where:

$E^T_{j,t}$ = total employment in area j at time t

$Z^P_{i,j,t}$ = passenger impedance between areas i and j at time t.

This passenger impedance term is a composite of several transport facility variables which is exactly analogous to the calculation of freight impedance in equation (3), with passengers and the passenger transport system being substituted for goods and the freight transport system.

The resident population in the area at the beginning of the period is expected to be positively related to out-migration from the area simply because the more people that live in an area, the more that can (and probably will) migrate, quite apart from area-to-area differences in rates. Total employment in the area was expected, as representative of local opportunities and of an urbanized area, to have a positive effect on in-migration and a negative effect on out-migration, thus subsuming both the local employment opportunities phenomenon and the historic rural-to-urban shift of population.

The natural-increase variable was expected to have a positive effect on out-migration, the assumption being that in areas with high levels of natural increase the overall population would be younger (i.e. low mortality and high fertility) and more mobile, and therefore more likely to migrate. Finally, the access to opportunities for employment is expected to exert a strong positive influence on in-migration. The results of estimating the equation were:

$$\begin{aligned} M_{i,\Delta t} = {} & \underset{(-5\ 78)}{-0\ 1127\, P_{i,t-1}} \; \underset{(-4\ 50)}{-\, 0{\cdot}0815\, E^T_{i,t}} \\ & \underset{(7\ 45)}{+\, 2{\cdot}0449\, (B_{i,\Delta t} - D_{i,\Delta t})} \; \underset{(3\ 97)}{+\, 2{\cdot}2675\, \phi_{i,t}} \\ & -13561{\cdot}0 \end{aligned}$$

$r^2 = 0{\cdot}5560 \qquad F = 37{\cdot}25$

The coefficient of lagged population is negative, as was expected. The coefficient of the access-to-opportunities measure is positive as was expected. The coefficient of present total employment was unexpectedly negative. This result may be due to the access-to-opportunities measure subsuming the local opportunities effects. If this were the case, then the negative coefficient of total employment might be indicative of the urban-to-suburban shift, i.e. a desire to live near, but not too near, the location of employment. The natural increase variable has an unexpected positive coefficient. It may be that this is an associative variable rather than a causal variable and that high levels of natural increase are associated with areas that the younger, more mobile segment of the population has moved to. In any case, on the occasion that additional data become available, it is clear that there is room for additional work on this topic.

Given the available data, it was then possible to make estimates of all these components of population change. This was done, and an estimate of population at time t was produced per:

$$\hat{P}_{i,t} = P_{i,t-1} + \hat{B}_{i,\Delta t} - \hat{D}_{,\Delta t} + \hat{M}_{i,\Delta t}$$

Since the actual values of $P_{i,t}$ were known for time t, it was possible to correlate $P_{i,t}$ with $\hat{P}_{i,t}$. Doing this yielded a simple correlation between them of 0·9978, which is equivalent to 99·58 per cent explanation of variance.

VI. THE PERSONAL INCOME EQUATIONS

The personal income forecasting equations are the least well developed portion of the model. At the time of development of this portion of the model no truly compatible income data were available for estimating such equations. A requirement imposed by the intended use of the model's outputs was that personal income be forecast on a where-received basis. The only income data available at the appropriate geographical level of detail were the 1960 Census of Population data. After some consideration it was decided, for this particular effort, to make income forecasts by primarily associative equations. Forecasts of personal income, by area, on a where-received basis, were made as a function of the employment-mix in the area. The general equation was:

$$F_{i,k,t} = \sum_m (a_m E_{i,m,t}) + a_n \qquad (13)$$

where:

$F_{i,k,t}$ = number of families in income class k in area i at time t

a_m, a_n = regression coefficients
$E_{i,m,t}$ = employment in area i of sector m at time t

The equations were estimated for six income classes which were defined as follows:

Class	Income
1	< \$1000
2	\$1000 → \$3000
3	\$3000 → \$5000
4	\$5000 → \$7000
5	\$7000 → \$10,000
6	> \$10,000

The results of the estimation were as follows:

Class 1: < \$1000

$$\begin{aligned} F_{i,1,t} = {} & \underset{(4{\cdot}06)}{0{\cdot}3251\,E_{i,6,t}} \underset{(-2{\cdot}75)}{-0{\cdot}0275\,E_{i,4,t}} \underset{(-2{\cdot}35)}{-0{\cdot}1650\,E_{i,8,t}} \\ & \underset{(-2{\cdot}97)}{-0{\cdot}1633\,E_{i,9,t}} \underset{(3{\cdot}32)}{+0{\cdot}2821\,E_{i,7,t}} \underset{(-0{\cdot}25)}{-0{\cdot}0377\,E_{i,3,t}} \\ & +505{\cdot}43 \end{aligned}$$

$r^2 = 0{\cdot}6614 \qquad F = 38{\cdot}74$

Class 2: \$1000 → \$3000

$$\begin{aligned} F_{i,2,t} = {} & \underset{(5{\cdot}63)}{0{\cdot}5853\,E_{i,6,t}} \underset{(0{\cdot}74)}{+0{\cdot}0096\,E_{i,4,t}} \underset{(-4{\cdot}39)}{-0{\cdot}3952\,E_{i,8,t}} \\ & \underset{(-0{\cdot}31)}{-0{\cdot}0213\,E_{i,9,t}} \underset{(4{\cdot}39)}{+0{\cdot}4604\,E_{i,7,t}} \underset{(-2{\cdot}35)}{-0{\cdot}4471\,E_{i,3,t}} \\ & +1482{\cdot}1 \end{aligned}$$

$r^2 = 0{\cdot}9216 \qquad F = 233{\cdot}21$

Class 3: \$3000 → \$5000

$$\begin{aligned} F_{i,3,t} = {} & \underset{(3{\cdot}75)}{0{\cdot}5251\,E_{i,6,t}} \underset{(5{\cdot}60)}{+0{\cdot}0952\,E_{i,4,t}} \underset{(-4{\cdot}29)}{-0{\cdot}5147\,E_{i,8,t}} \\ & \underset{(0{\cdot}27)}{+0{\cdot}0243\,E_{i,9,t}} \underset{(4{\cdot}67)}{+0{\cdot}6536\,E_{i,7,t}} \underset{(-1{\cdot}40)}{-0{\cdot}3513\,E_{i,3,t}} \\ & +2024{\cdot}7 \end{aligned}$$

$r^2 = 0{\cdot}9560 \qquad F = 431{\cdot}29$

Class 4: \$5000 → \$7000

$$\begin{aligned} F_{i,4,t} = {} & \underset{(-2{\cdot}27)}{-0{\cdot}3412\,E_{i,6,t}} \underset{(12{\cdot}89)}{+0{\cdot}2447\,E_{i,4,t}} \underset{(-2{\cdot}67)}{-0{\cdot}3481\,E_{i,8,t}} \\ & \underset{(-2{\cdot}36)}{-0{\cdot}2318\,E_{i,9,t}} \underset{(6{\cdot}46)}{+0{\cdot}9698\,E_{i,7,t}} \underset{(1{\cdot}14)}{+0{\cdot}3075\,E_{i,3,t}} \\ & +1390{\cdot}5 \end{aligned}$$

$r^2 = 0{\cdot}9716 \qquad F = 679{\cdot}10$

Class 5: \$7000 → \$10,000

$$\begin{aligned} F_{i,5,t} = {} & \underset{(-5{\cdot}61)}{-1{\cdot}2333\,E_{i,6,t}} \underset{(3{\cdot}99)}{+0{\cdot}2559\,E_{i,4,t}} \underset{(-0{\cdot}26)}{-0{\cdot}9496\,E_{i,8,t}} \\ & \underset{(-2{\cdot}22)}{-0{\cdot}3111\,E_{i,9,t}} \underset{(3{\cdot}99)}{+0{\cdot}8788\,E_{i,7,t}} \underset{(1{\cdot}78)}{+1{\cdot}7061\,E_{i,3,t}} \\ & +79{\cdot}08 \end{aligned}$$

$r^2 = 0{\cdot}9408 \qquad F = 315{\cdot}40$

Class 6: > \$10,000

$$\begin{aligned} F_{i,6,t} = {} & \underset{(-5{\cdot}68)}{-2{\cdot}0447\,E_{i,6,t}} \underset{(4{\cdot}62)}{+0{\cdot}2082\,E_{i,4,t}} \underset{(0{\cdot}53)}{+0{\cdot}1642\,E_{i,8,t}} \\ & \underset{(-0{\cdot}62)}{-0{\cdot}1507\,E_{i,9,t}} \underset{(1{\cdot}10)}{+0{\cdot}4059\,E_{i,7,t}} \underset{(5{\cdot}91)}{+3{\cdot}8465\,E_{i,3,t}} \\ & -1682{\cdot}4 \end{aligned}$$

$r^2 = 0{\cdot}8397 \qquad F = 103{\cdot}87$

VII. THE LAND VALUE—LAND USE FORECASTS

This portion of the model shares with the personal-income section the dubious honor of being based upon particularly poor data. The consequence here, as above, is that the forecasting equations are primarily associative.

The land value forecasts are actually forecasts of assessed value. Though it is well known that assessed value has certain non-economic (e.g. political) determinants, suitable sales value data are nonexistent. Consequently, assessed value and projections thereof are used in this model as a surrogate for sales value. The forecasting equation includes measures of economic activity as well as lagged variables from two time periods. It is worth noting that the inclusion of the second lagged assessed value variable does not significantly alter the equation's explanation of variance, but it does add stability to the forecast equation, which is necessary to dampen the sharp changes in assessed value which can result from political decisions. The equation used to forecast assessed value is:

$$\begin{aligned} V_{i,t} = {} & a_1(V_{i,t-1}+\delta V_{i,t-2})+a_2[(E^T_{i,t}-E^T_{i,t-1})/L_i] \\ & +a_3(P_{i,t}/L_i)+a_4 \qquad (14) \end{aligned}$$

where:

$V_{i,t}$ = total assessed value of land in area i at time t
δ = a constant; $0{\cdot}0 \leqq \delta \leqq 0{\cdot}5$
$E^T_{i,t}$ = total employment in area i at time t
L_i = total land area of area i

The results of estimating this equation were:

$$V_{i,t} = 1\cdot1394\,(V_{i,t-1} + 0\cdot2\,V_{i,t-2})$$
$$(13\cdot10)$$
$$+\,1641\cdot4\,[(E^T_{i,t} - E^T_{i,t-1})]L_i - 33\cdot1160\,(P_{i,t}/L_i)$$
$$(2\cdot73) \qquad\qquad (-1\cdot30)$$
$$+\,170472.$$

$r^2 = 0\cdot7198 \qquad F = 104\cdot47$

Forecasts of three categories of land use are made: (1) Industrial, (2) Commercial, (3) Residential. The equations attempt the use of accessibility, as reflective of land desirability, as an independent variable along with density of present use.

The equation for forecasting industrial land use was

$$U_{i,I,t} = a_1\,E_{i,4,t} + a_2\,(E^T_{i,t}/L_i) + a_3\,A^T_{i,t} + a_4 \quad (15)$$

The equation for forecasting commercial land use was

$$U_{i,C,t} = a_1\,(1\cdot0/L_i) + a_2\,(E_{i,6,t})^{a_3} + a_4 \quad (16)$$

The equation for forecasting residential land use was:

$$U_{i,R,t} = a_1\,(P_{i,t}/L_i) + a_2\,\pi_i + a_3\,A^P_{i,t} + a_4 \quad (17)$$

where:

$U_{i,I,t}$ = percent of industrial land in area i at time t
$U_{i,C,t}$ = percent of commercial land in area i at time t
$U_{i,R,t}$ = percent of residential land in area i at time t
π_i = 0 when $(P_{i,t}/L_i) < 10{,}000$
1 when $(P_{i,t}/L_i) \geqslant 10{,}000$
$E^T_{i,t}$ = total employment in area i at time t
L_i = land area of area i
$A^T_{i,t}$ = total accessibility to employment of area i at time t
$A^P_{i,t}$ = total accessibility to residential population of area i at time t
$= \sum_j (P_{j,t}/Z^P_{i,j,t})$
$P_{i,t}$ = population residing in area i at time

The results of estimating these equations were:

Industrial land use by county:

$$U_{i,I,t} = 0\;0020\,E_{i,4,t} + 0\cdot00082\,(E^T_{i,t}/L_i)$$
$$(9\;13) \qquad\qquad (4\cdot05)$$
$$+\,0\cdot000008\,A^T_i - 0\cdot449$$
$$(3\;28)$$

$r^2 = 0\cdot7534$

Commercial land use by county.

$$U_{i,C,t} = 56\;342(1\cdot0/L_i) + 1\cdot223(E_{i,6,t})^{0\;025} + 0\cdot1262$$
$$(2\cdot56) \qquad\qquad (5\cdot69)$$

$r^2 = 0\cdot6956$

Residential land use by county:

$$U_{i,R,t} = 0\cdot0554\,(P_{i,t}/L_i) - 58\cdot4531\,\pi_i$$
$$(6\;02) \qquad\qquad (-4\cdot98)$$
$$+\,0\cdot000003\,A^P_{i,t} - 5\cdot17$$
$$(6\cdot47)$$

$r^2 = 0\cdot7276$

VIII. COMMENTS ON EQUATION STRUCTURE

The model's equation structure may now be briefly summarized:

For Sectors 1, 2, 4, 5:

$$E_{i,g,t} = f(A^s_{i,g,t-1};\, A^m_{i,g,t-1};\, E_{i,g,t-1};\, E^{T'}_{i,g,t-1};\; P_{i,t-1}/L_i)$$

For Sectors 7, 8, 9, 10

$$E_{i,g,t} = f(E_{i,g,t-1};\, P_{i,t};\, E^T_{i,t})$$

For Sector 3

$$E_{i,g,t} = f(A^T_{i,t-1};\, E_{i,g,t-1};\, E_{i,g,t-1}/L_i;\, P_{i,t-1}/L_i)$$

For Sector 6

$$E_{i,g,t} = f(A^T_{i,t};\, E_{i,g,t-1};\, P_{i,t};\, E^T_{i,t})$$

Population

$$P_{i,t} = P_{i,t-1} + B_{i,\Delta t} - D_{i,\Delta t} + M_{i,\Delta t}$$
$$B_{i,\Delta t} = f(Y_{i,t-1}/P_{i,t-1};\, P_{i,t-1}/L_i;\, B_{\cdot,\Delta t'})$$
$$D_{i,\Delta t} = f(Y_{i,t-1}/P_{i,t-1};\, P_{i,t-1}/L_i,\, D_{i,\Delta t'})$$
$$M_{i,\Delta t} = f(P_{i,t-1};\, E^I_{i,t};\, (B_{i,\Delta t} - D_{i,\Delta t});\, \phi_{i,t})$$

Personal Income

$$F_{i,k,t} = f(E_{i,g,t})$$

Land Value

$$V_{i,t} = f(V_{i,t-1};\, V_{i,t-2};\, (E^T_{i,t} - E^T_{i,t-1})/L_i;\, P_{i,t}/L_i)$$

Land Use

$$U_{i,I,t} = f(E_{i,4,t};\, E^T_{i,t}/L_i;\, A^T_{i,t})$$
$$U_{i,C,t} = f(L_i;\, E_{i,6,t})$$
$$U_{i,R,t} = f(P_{i,t}/L_i;\, A^P_{i,t})$$

It will be remembered that there are constraints on:

$\sum_i E_{i,g,t} = E_{g,t}$ (Regional Employment Forecasts)
$\sum_i P_{i,t} = P_t$ (Regional Population Forecast)
$\sum_i Y_{i,t} = Y_t$ (Regional Income Forecast)

In order to obtain personal income in a given area i, the numbers of families in each income class are multiplied by the median income of the class (CONSAD, 1968)

It is now possible to trace certain types of impact through the model system.

A. Changes in freight transport facility characteristics at time t will not cause an impact until $t+1$. At that time they will have a direct impact on employment in Sectors 1, 2, 3, 4, 5, and 6 This will, in turn, produce indirect effects on employment in all other sectors, as well as on population, income, land value, and land use. All of these effects are, of course, cumulative, in that they will affect the forecasts for succeeding time periods.

B. Changes in passenger transport facility characteristics at time t will have no direct effect on employment in Sectors 1, 2, 3, 4, and 5 There will, however, be a direct impact on population (via migration). Thus, there will be indirect effects on employment in Sectors 6, 7, 8, 9, and 10, as well as on income, land value, and land use. Again, these effects are cumulative in that they will affect the forecasts for succeeding time periods.

C. Location of major facilities, either private or public (government), will have impacts in different parts of the model, depending upon the nature of the facility. For example, large factories or similar installations will have numbers of Sector 3 employees in the area during their construction, which will have direct impact on employment in Sectors 6, 7, 8, 9, and 10, as well as on population, income, land value, and land use Other types of exogenous effects may also be traced through the model once their initial representation, in terms of the model's variables, has been determined.

IX. SOLUTION OF THE MODEL

The solving of the model, i.e. making a forecast with it, consists of three main steps. First, employment in Sectors 1, 2, 4, and 5 is forecast. Second, population and all other sectors of employment (Sectors 3, 6, 7, 8, 9, 10) are forecast simultaneously. Third, personal income, land value, and land use are forecast. This structure is shown in Fig. 2.

The only portion of the above-mentioned structure which needs further clarification is Step Two, the Population and Sectors 3, 6, 7, 8, 9, 10 employment section. This is, in fact, a rather large set of simultaneous equations. To avoid the problems of solving such a system simultaneously, the solution is arrived at by use of an iterative technique. This technique is illustrated below.

Given, for example:

$$y = a_1 x + b_1 z + c_1 \tag{1}$$

$$x = a_2 y + b_2 z + c_2 \tag{2}$$

$$z = a_3 x + b_3 y + c_3 \tag{3}$$

The coefficients a_n, b_n, and c_n are known parameters. Initial values are assumed for x and z in equation (1). This equation may then be solved for a trial value of y as per:

$$\hat{y} = a_1 x_0 + b_1 z_0 + c_1$$

where

$\hat{y}$ = trial value of y

x_0, z_0 = initial values of x and z

Then in equation (2) the trial value of y and the initial value of z are used to calculate a trial value of x.

$$\hat{x} = a_2 \hat{y} + b_2 z_0 + c_2$$

Continuing the process,

$$\hat{z} = a_3 \hat{x} + b_3 \hat{y} + c_3$$

The solution continues by calculating a second trial value of y per.

$$\hat{\hat{y}} = a_1 \hat{x} + b_1 \hat{z} + c_1$$

Second trial values are also calculated for x and z:

$$\hat{\hat{x}} = a_2 \hat{\hat{y}} + b_2 \hat{z} + c_2$$

$$\hat{\hat{z}} = a_3 \hat{\hat{x}} + b_3 \hat{\hat{y}} + c_3$$

This process continues until the nth trial value is equal to (or within a tolerance range of) the $(n-1)$th trial value. For runs of the model, to date, the tolerance level has been set to 3 percent, and solution takes four to seven iterations (less than 2 minutes on an IBM 360/75). To test the uniqueness of the solution reached by this technique a test was made where the first trial values of x, y, and z were random numbers instead of the previous time period values normally used. The solution reached by the test problem, though taking more iterations, was the same as was obtained in non-test runs of the model

X. VALIDATION OF THE MODEL

Though all data necessary for a full validation of the model were not available, a partial validation was undertaken. A test forecast of all variables was made from 1960 to 1965. The 1965 County Business Patterns (U.S. Department of Commerce, 1966) were summed to provide employment control totals. The Bureau of the

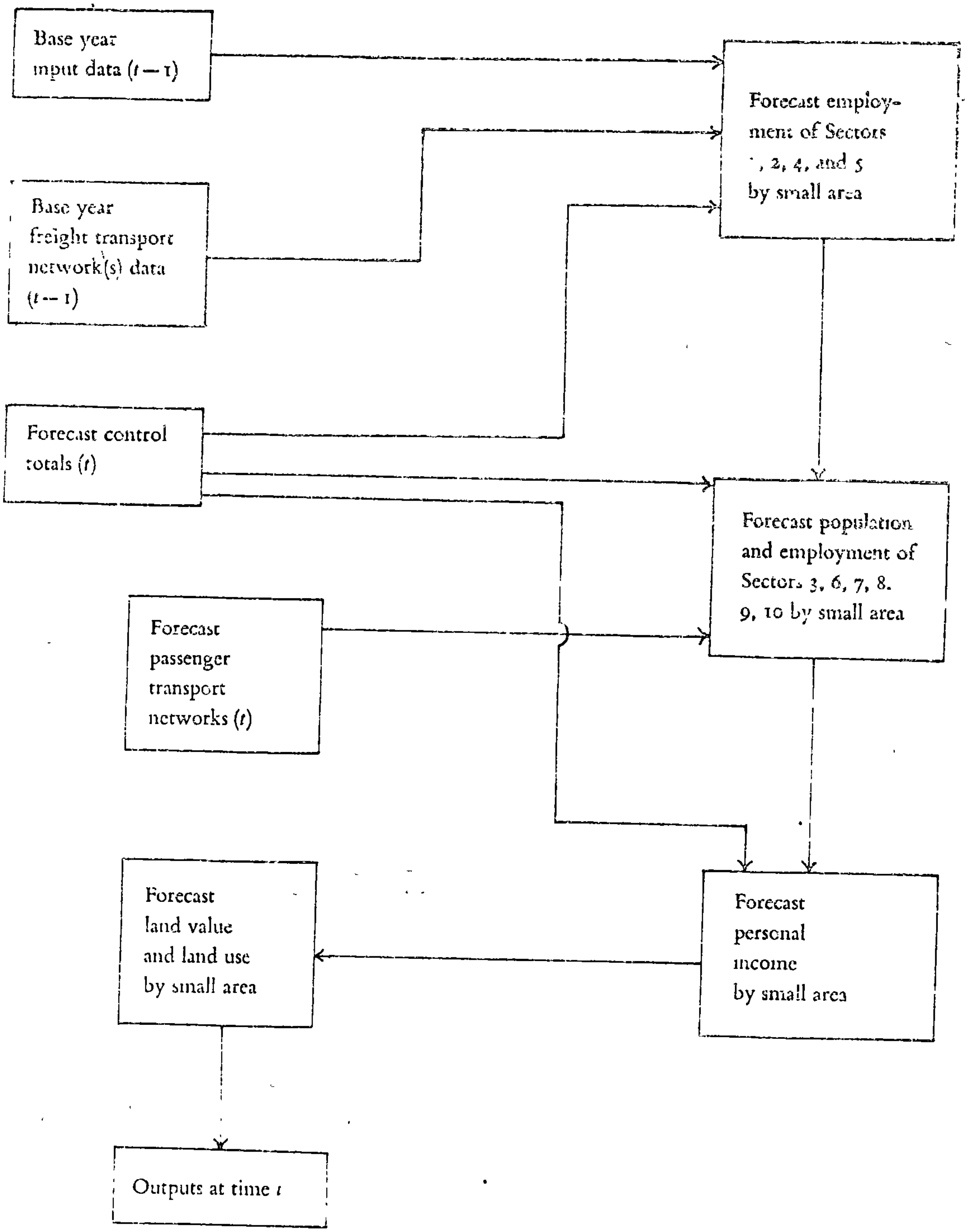


Fig 2 The solution procedure for the INTRA-I model

Census 1965 inter-censal estimates (Bureau of the Census, 1968) were summed to provide a population control total. The impedance data, 1965 passenger impedance and 1960 freight impedance, were supplied by the Northeast Corridor Transportation Project. The forecasts were made for the 131 districts and then aggregated to the 40 superdistricts. These forecasts were then plotted versus the actual 1965 data for comparison purposes. In these plots, the superdistricts were ordered in order of ascending magnitude of the actual data. The actual value of the variable, for each superdistrict, was plotted as an asterisk, and a line was drawn connecting these. The forecast variable was plotted as a ▼ for each superdistrict, and a line was drawn connecting these. In Fig 4 the population forecasts are compared with actual population In Fig. 3 the employment forecasts are compared with actual employment. As can be seen, in both cases the forecasts match the actuals rather closely.

In an attempt to further validate the model's predictive ability summations of the 1968 County Business Patterns employment data were used as sectoral control totals (except for Sector 1, Agriculture; Sector 5, Transportation, Communication, etc.; and Sector 10, Government; which are not fully covered by County Business Patterns and for which the special tabulations used to adjust 1960 and 1965 data were not available for 1968). The model's forecasts for the 131 districts, which were expected to be at least slightly in error due to the use of 1970 estimates as control totals of the three above-mentioned employment sectors and for population, were then correlated with the 1968 County Business Pattern data by county. The following results were obtained.

Sector	r^2
2. Mining	0·6878
3. Contract Construction	0·9510
4. Manufacturing	0·9717
6. Wholesale Trade	0·9647
7. Retail Trade	0·9577
8. Finance, Insurance, Real Estate	0·7505
9. Services	0·8821

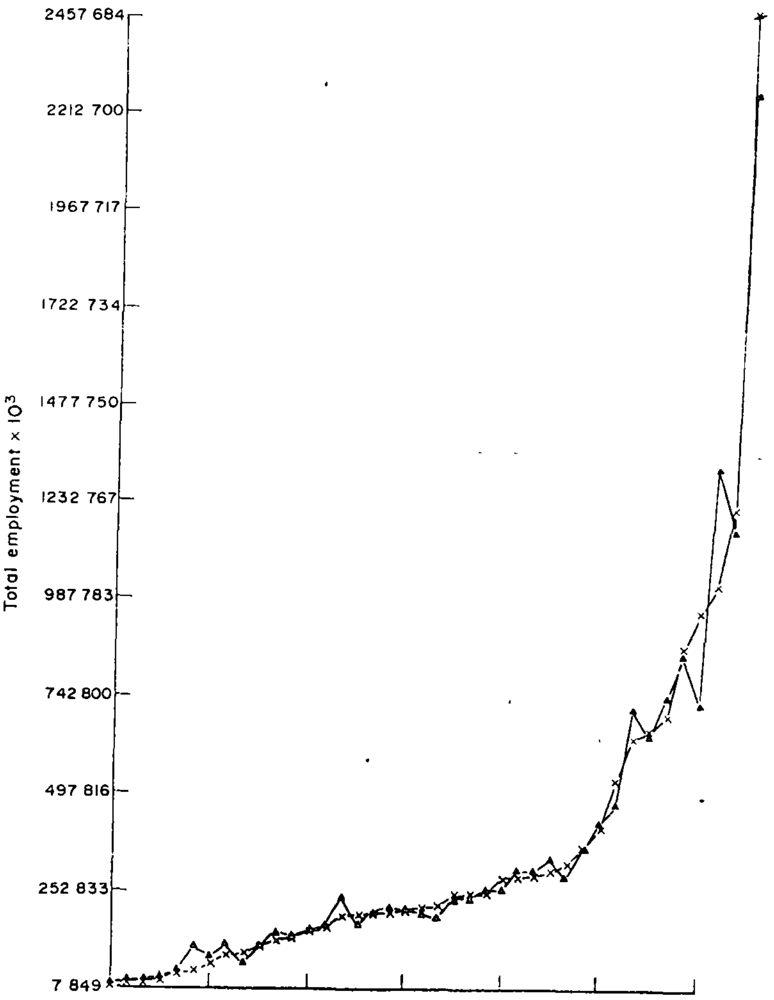


Fig. 3. *Actual superdistrict employment totals vs. INTRA-I forecasts* (1965)

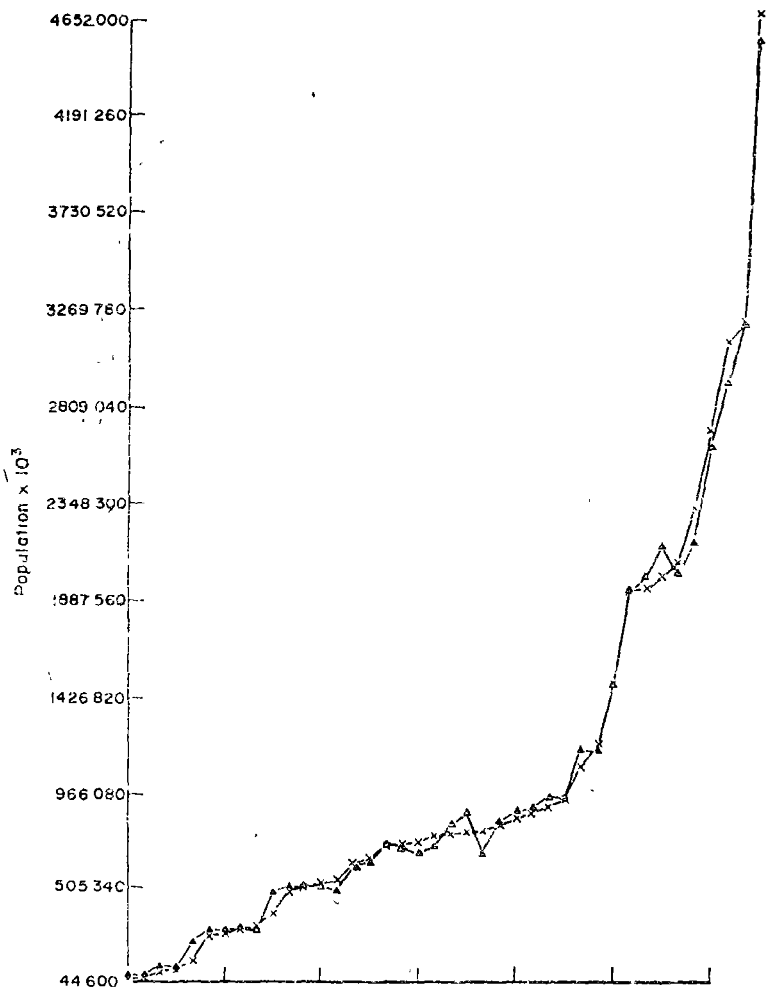


Fig. 4. Actual superdistrict population totals vs. INTRA-I forecasts (1965).

Under the circumstances, these results are reasonably good and lend additional weight to the argument that the model is capable of producing good forecasts.

It should be noted here that an inherent weakness of models developed from regression analyses is the problem of "outliers". These are observations included in the data (which in this case must later be forecast) that are very far from the regression line. A particularly noticeable outlying point through all the equations estimated for INTRA-I was District 59, New York, N.Y. (Manhattan). The development of this area, historically, is not adequately captured by the INTRA-I equations. The direct consequence of this is that, particularly for population, the model consistently underestimates in this area, as may be seen in the last points plotted on Figs. 3 and 4. Though underestimation is found in other areas, too, a peculiar combination of effects in Manhattan results in forecasts of precipitous population decline in the area after 1975. While there has been a decline in the area's population since 1950, the decline forecast by the model was considered far too steep to be acceptable. Consequently an artificial constraint was introduced which limits the decline in a given 5-year forecast period to 10 per cent of the prior period population level. Even though this constraint is called into operation in only a few instances, future work on intra-regional models of this sort should attempt to further explore this phenomenon.

A variety of input alternatives were prepared and tested with the model. Since these would be lengthy to describe in detail, only the conclusions from several of these tests are presented here.

Test I—An improvement (reduction) in freight impedance from one area to all others (the area tested was neither extremely urban nor extremely rural) produces the following forecast results:

(a) Increases in the given area and in the more industrialized of the adjacent areas (except, peculiarly, for Sector 8—Finance *et al.*).

(b) Slight decreases in the resident population of the given area, with slight increases in the less industrialized of the adjacent areas.

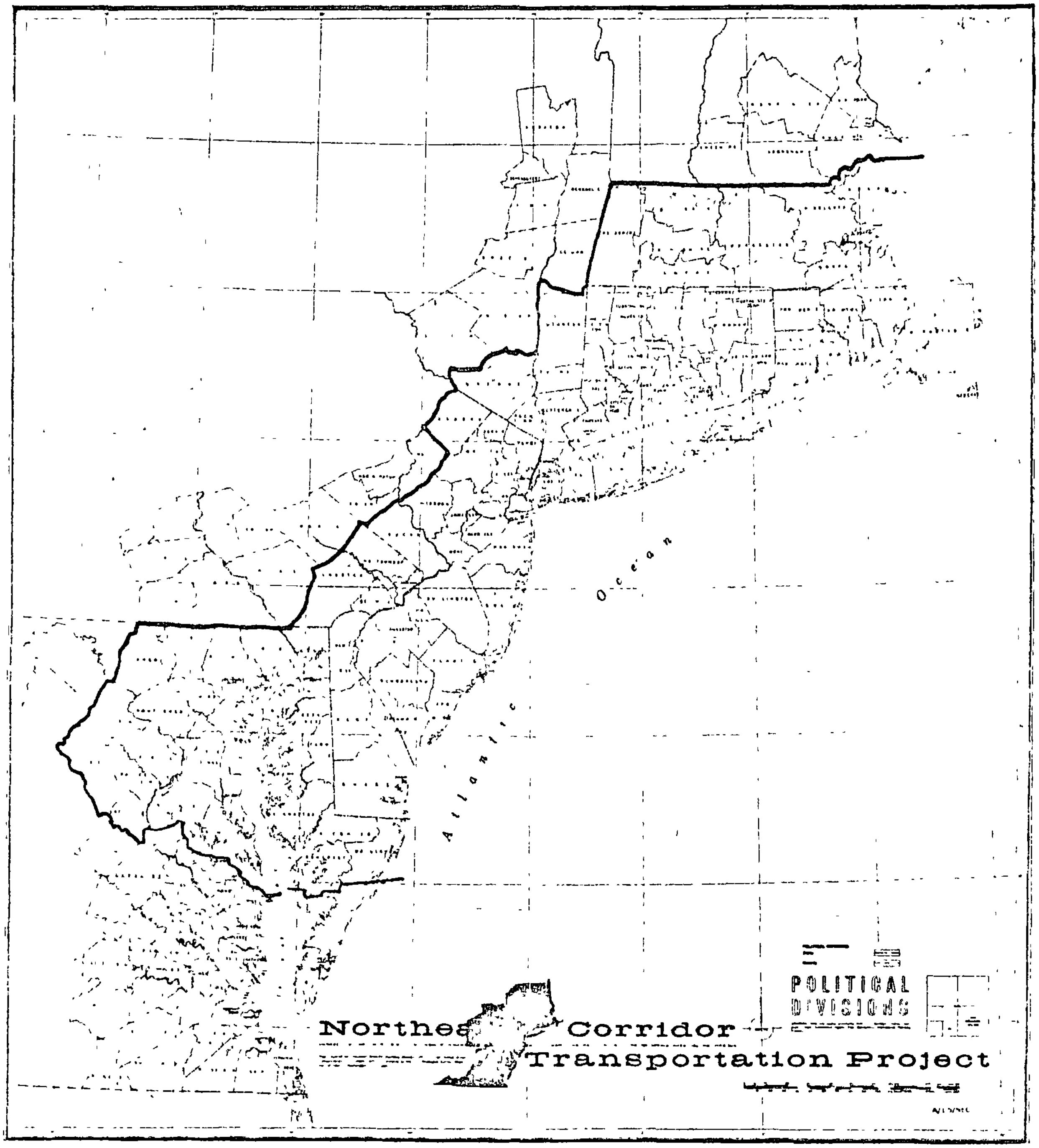


Fig 5 The Northeast Corridor "core area".

(c) Increased per-capita income level in the given area.

Test II—An improvement (reduction) in passenger impedance from one area to all others (same area as used in Test I) produces the following forecast results:

(a) Increases in resident population of the given area with virtually no change in the adjacent areas.
(b) Virtually no changes in employment.
(c) Decreased per-capita income level in the given area. (This result is somewhat suspect in that inspection of the model's equation structure indicates that as the resident population of an area increases, if the area's employment remains unchanged, the area's per-capita income will, by definition, decrease.)

Test III—An improvement (reduction in passenger impedance along the center of the "core" area (shown in Fig. 5) produces the following forecast results:

(a) A tendency toward decentralization of population and, subsequently, of employment, along the north–south axis followed by the improved transport system.

It should be noted that in all cases the impacts observed were for the first time period subsequent to the changed transport facilities' being affected. Subsequent time periods will obviously reflect both the changed transport facilities and the compounding of their induced effects.

Test IV—Changes in the intersectoral flow coefficients (the input–output coefficients) produced changes in the geographic distribution of employment.

Further tests are presently being conducted and will hopefully be reported at a later date.

XI. CONCLUSIONS

Numerous conclusions have already been presented in the body of this paper. Only three points need to be made here.

1. The model does appear to produce, subject to exogenous control totals, reasonably accurate and reliable forecasts, though there is clearly room for further development.
2. The model's results were used by the Northeast Corridor Transportation Project of the U.S. Department of Transportation as part of their major 1969 report to the Secretary of Transportation.
3. The computer program of the model has been made operational and extremely easy to use at the University of Pennsylvania's Computation Center for use by interested faculty and students in an attempt, among other things, to encourage further development and testing of it.

REFERENCES

Anderson T. R. (1955) Intermetropolitan migration. a comparison of the hypotheses of Zipf and Stouffer, *Amer. Sociol. Rev.* **20**, 287–291.

Blanco C. (1963) The determinants of interstate population movements, *J. Reg. Sci.* **5**, 77–84.

Bureau of the Census (1968) Estimates of the population of counties: 1 July, 1966, *Current Population Reports*, Series P-25, No. 401, August 1968.

CONSAD Research Corp. (1967) *Impact Studies: Northeast Corridor Transportation Project*, Vol. I, September 1967.

CONSAD Research Corp (1968) *Impact Studies: Northeast Corridor Transportation Project*, Vol. II, January 1968.

CONSAD Research Corp. (1969a) *Impact Studies: Northeast Corridor Transportation Project*, Vol. III January 1969.

CONSAD Research Corp. (1969b) *Impact Studies: Northeast Corridor Transportation Project* Vol. IVA, December 1969.

CONSAD Research Corp. (1970) *Impact Studies: Northeast Corridor Transportation Project*, Vol. IVB, February 1970.

Crow R. T. (1969) *An Econometric Model of the Northeast Corridor of the United States.* U.S Department of Transportation, Office of High Speed Ground Transportation Northeast Corridor Transportation Project, May 1969.

Fuchs V. R. (1962) *Changes in the Location of Manufacturing in the United States Since 1929* Yale University Press, New Haven, Conn.

Glancy D. M (1965) Description of the area system for the Northeast Corridor Transportation Project, U S. Department of Transportation, Transport Systems Planning Division: Technical Paper No. 2, December 1965.

Greenhut M L. (1956) *Plant Location in Theory and Practice*. The University of North Carolina Press, Chapel Hill, N.C

HAMBURG J R and LATHROP G T (1964) An opportunity-accessibility model for allocating regional growth, presented at the Seminar on Models of Land Use Development, Institute for Urban Studies, University of Pennsylvania, 22–24 October, 1964.

HARRIS C D (1954) The market as a factor in the localization of industry in the United States, *Ann Ass Am Geogr* **44,** 315–348.

HOOVER E M. (1948) *The Location of Economic Activity* McGraw-Hill, New York.

HOOVER E. M. (1967) Some programmed models of industry location, *Land Economics* **XLIII,** No. 3, August 1967.

ISARD W (1960) *Methods of Regional Analysis: An Introduction to Regional Science.* The Technology Press of the Massachusetts Institute of Technology, Cambridge, Mass.

LANE T (1966) The urban base multiplier: an evaluation of the state of the art, *Land Economics* **XLII,** No 3, August 1966.

LEONTIEF W and STROUT A (1963) Multiregional input-output analysis, in *Structural Interdependence and Economic Development* (edited by T BARNA) Macmillan, London.

LOWRY I S (1966) *Migration and Metropolitan Growth Two Analytical Models* Chandler Publishing Co, San Francisco, Calif.

National Bureau of Economic Research (1960) *Demographic and Economic Change in Developed Countries* Princeton University Press, Princeton, N J

NOURSE H. O (1968) *Regional Economics.* McGraw-Hill, New York.

Office of Business Economics (1965) The transactions table of the 1958 input-output study and revised direct and total requirements data, *Survey of Current Business* **45,** No. 9, September 1965.

OLSSON G. (1965) *Distance and Human Interaction* Regional Science Research Institute, Philadelphia, Pa.

PETERS W. S. and SUMMERS G W (1968) *Statistical Analysis for Business Decisions.* Prentice-Hall Inc, Englewood Cliffs, N J

PUTNAM S H (1967a) Models for determining indirect impacts of high speed ground transportation systems Northeast Corridor, *Papers—Eighth Annual Meeting Transportation Research Forum*

PUTMAN S H (1967b) Modelling and evaluating the indirect impacts of alternative Northeast Corridor transportation systems, *Highway Research Record,* No 180

PUTMAN S H (1969) An economic model for forecasting employment by industrial sector in small sub-areas of a large region Unpublished Ph.D. thesis, University of Pittsburgh, Pittsburgh, Pa.

RAIMAN R L (1962) Interstate migration and wage theory, *Rev. Econ Statist.* **44,** 428–438.

SJAASTAD L A. (1961) Occupational structure and migration patterns, in *Labor Mobility and Population in Agriculture.* Iowa State University Press, Ames, Iowa

STOUFFER S A (1940) Intervening opportunities: a theory relating mobility and distance, *Amer. Sociol Rev.* **5,** 845–867.

THOMAS M. D (1969) Regional economic growth: some conceptual aspects, *Land Economics,* **XLV,** No 1, February 1969.

ULLMAN E. L. and DACEY M. F. (1960) The minimum requirements approach to the urban economic base, *Papers and Proceedings, The Regional Science Association,* Vol. VI.

U.S Department of Commerce (1960) *County Business Patterns, First Quarter 1959.* Bureau of the Census, U S Government Printing Office, Washington, D.C

U S Department of Commerce (1966) *County Business Patterns, First Quarter 1965.* Bureau of the Census, U.S. Government Printing Office, Washington, D.C

Draft

THE INTEGRATED FORECASTING OF TRANSPORTATION AND LAND USE

Prepared for presentation at a Seminar on
EMERGING TRANSPORTATION PLANNING METHODS

Sponsored by the U.S. Department of Transportation

Daytona Beach, Florida - December 1976

Stephen H. Putman
Associate Professor of City and Regional Planning
University of Pennsylvania
Philadelphia, Pa. 19174

Abstract

The past five to ten years research on urban transportation and land use models has produced important advances in operational planning techniques. Two of these advances are closely related and are discussed in this paper. The first of these was the demonstration of both the feasibility and superiority of an integrated transportation and land use model package. The second of these was the development of a more general form of urban land use model along with the procedures necessary for its calibration. Put together, these two advances make integrated transportation and land use forecasting and policy analyses a reasonable operational analysis technique.

This paper first describes the general development of an integrated transportation and land use model. The bulk of the paper is then devoted to the development of a simple numerical example of one of the new land use models along with a procedure for estimating its parameters. The concluding section of the paper points out the connection of this new land use model to transportation models.

Introduction

Serious efforts to construct urban land use simulation models began in the early 1960's. In all of these efforts the relationships between transportation and land use were only partially represented. In each case, a matrix of zone-to-zone impedances was input to a land use model. These impedances were measures of the difficulty of interaction between activities located in different zones. Depending upon the availability of data, travel times, travel costs, or airline distances (or various combinations thereof) were used as impedances. The impedances were the only link between transportation and land use in these modelling efforts.

During this same period the modelling of transportation (especially highway) networks progressed to the point of rather routine application of trip generation, distribution and assignment program packages. The connection from the land use distributions to these transportation models was simply the input of the land use data to the network model package.

As part of its response to an increasing concern with the premature obsolescence (over-congestion) of urban highways, in March 1971 the U.S. Department of Transportation* issued a request for a proposal to do a research study entitled "Interrelationships of Transportation Development and Land Development". The contract for the study was awarded to Professor Stephen Putman at the University of Pennsylvania, with the work beginning in November of 1971. It was evident from the start of the work that the premature obsolescence question was but one facet of the more general question of urban spatial configurations.

---

*Federal Highway Administration, Office of Highway Planning, Urban Planning Division.

The proposed method of analysis for the general problem was an integrated set of land use and transportation models. As the project emphasis was on the integration of these models and their consequent capability for analyzing land use-transportation interactions, it was decided to use modified versions of existing models rather than constructing wholly new ones. The land use model portion of the package was a much modified version of PLUM, called IPLUM. The transportation part of the package was an "in house" network package incorporating the capacity restrained, incremental tree-by-tree assignment. In early tests of the model package, in order to assure eventual compatability with FHWA software, a trial linkage was constructed and several test runs were also made using the FHWA-UTPS network programs.

This project, completed in the summer of 1973, demonstrated the first successful integration of a land use model with a transportation model, incorporating feedbacks.[1] Trip generation was accomplished by extracting and converting the implicit trip matrices from the land use model activity allocations. Congested network times resulting from these trips were then fed back into the land use model. The entire process was run, in an iterative manner, to equilibrium.[2]

Though this first project did not involve extensive policy testing, there were some interesting policy conclusions. In particular, the simulation results demonstrated that attempts to deal with highway congestion by means of new highway facility construction will, in the absence of stringent land development controls, lead to even greater highway congestion in future years. In addition it was shown that a

---

[1] Putman, S. H. (1973) "The Interrelationships of Transportation Development and Land Development" Urban Planning Division, FHWA, Department of Transportation, Washington D.C. (in 2 Volumes, revised and reprinted in 1976).

[2] Putman, S.H. (1974) "Preliminary Results From an Integrated Transportation and Land Use Model Package" Transportation Vol. 3, pp. 193-224.

"do nothing" policy of allowing existing facilities to congest might lead, in the long run, to lower levels of congestion. As part of these conclusions it was clearly demonstrated that new highway facilities, in the absence of land use controls, lead to increased urban sprawl, while congested facilities tend to inhibit sprawl. Finally, these test runs indicated relationships throughout large metropolitan regions; e.g. the effects of land use controls imposed in one county on development in adjacent counties, that would probably pass unnoticed with more traditional analysis techniques.

In May 1973 a grant was received from the National Science Foundation for a study titled Laboratory Testing of Predictive Models. The purpose of this work was to develop a method for testing and evaluating predictive land use models, and to apply this method to existing land use models. After a thorough review of existing land use models, the efforts of the project were devoted to work with the EMPIRIC model and a version of IPLUM (which resulted from the D.O.T. work described above) as a representative of Lowry derivative models.

The research strategy in this new work was to first attempt to estimate both models' parameters for a common data set, and then to do a series of sensitivity test of both models. No difficulties were encountered in estimating parameters for EMPIRIC. When it came to estimating the parameters for IPLUM it was discovered that there was no formal technique in U.S. modelling practice for estimating the parameters of Lowry derivative models. In dealing with this dilemna, the research was eventually led to consideration of the entropy-maximizing approach of Alan Wilson.[1] Through this approach IPLUM was reformulated and subsequently calibrated (i.e. its parameters were properly estimated). The resulting model was sufficiently

---

[1] Wilson, A. G. (1970) Entropy in Urban and Regional Modelling , Pion, London.

different from IPLUM to be named Disaggregated Residential Allocation Model - DRAM.[1]

The two models, DRAM and EMPIRIC, were then put through a series of tests of their responses (i.e. changes in outputs) to different types of change in inputs. The results of these test showed DRAM to be much more responsive to changes in inputs than was EMPIRIC. This, along with DRAM's better articulated theoretical structure suggests a clear superiority over EMPIRIC for use in planning policy tests and evaluations. This work was completed with the preparation of a draft final report in the Summer of 1975.[2]

June 1975 saw the beginning of a second round of work on the Integrated Transportation and Land Use Package. While some further work had been done in the preceeding two years,[3] a National Science Foundation grant titled Development of an Improved Integrated Transportation and Land Use Model provided the funds necessary to carry the development of ITLUP to its present state. The chief goal of this effort was to upgrade ITLUP tp include DRAM, and to add several additional models to the package. Also to be explored were a basic employment model, a modal split procedure, and an air pollution model. The whole package, including parameter estimation procedures was to be prepared for dissemination to experienced users in the field.

Additional funds from the Office of the Secretary, D.O.T. were included in this grant to cover the preparation of a number of policy tests using the improved ITLUP package on an actual data set. These policy tests are currently under way.

---

[1] Putman, S. H. (1977) "Calibrating a Disaggregated Residential Allocation Model - DRAM" London Papers in Regional Science Vol. 7, pp. 108-124.

[2] Putman, S. H. (1976) "Laboratory Testing of Predictive Land-Use Models: Some Comparisons" Office of Transportation Systems Analysis and Information, DOT, Washington D.C.

[3] Putman, S. H. (1976) "Further Results from the Integrated Transportation and Land Use Model Package (ITLUP)", pp. 165-173.

As part of the work of the current grant, many calibrations of DRAM were performed, including some for a 19th century data set,[1] with highly encouraging results confirming the general form of the model. Similarly, a simple model of basic employment location has been developed and yields rather promising results. All of this work is nearing completion and will be written up in the near future.

The importance of this work, quite apart from the various interesting theoretical-empirical findings, lies in the preparation of advanced planning technology for distribution and use to practitioners in the field. Use of these methods can reasonably be expected to improve planning analyses, while the results of such uses can be used to direct future research efforts towards the development of even more accurate and useful methods.

These new techniques, along with numerical examples of their derivation and application, are presented in the following pages.

---

[1] Putman, S. H. (1977) "Calibrating a Residential Location Model for Nineteenth-Century Philadelphia" Environment and Planning A Vol. 9.

Integrating Transportation and Land Use Models:

It is common knowledge that in the past two decades there has been enormous public investment in highway construction. This same period has also witnessed large scale shifts in population from rural areas to urban areas; and within the urban area, from center-city to suburb. Even without attempting here to define the explicit causal relationships between these two developments, it is possible to describe a related phenomenon which is observed, not infrequently, when the construction of a new section of roadway is followed, all too soon, by very heavy usage and subsequent congestion. Specifically, the nature of this process seems to be that: a) Due to the inadequacy of existing facilities a decision is taken in a metropolitan area to improve transport facilities (usually by road building) for a particular part of the area. Then b) assuming that the decision is approved, in anticipation of the new roadway, land developers and/or speculators become involved with properties near the proposed route. As construction of the facility begins, some homeowners and businesses may consider and even act upon location decisions. Upon completion of the facility, additional location decisions are made. Finally c) a relatively short time after completion of the facility, the demand for its use greatly exceeds the demand which existed prior to the decision to construct it. Consequently the design capacities are soon reached and often exceeded, resulting in congestion and premature obsolescence of the facility. While this is a rather generalized description of a very complex process, it is reasonably accurate. The question is whether it is feasible, via integrated transportation planning and land use planning, to avoid or ameliorate the occurrence of this particular phenomenon in the future. Further it is necessary to analyze this process and to determine whether a) balanced development of transportation facilities and land use is feasible, and b) if it is feasible, what means are available to accomplish it.

There is considerable evidence indicating that the demand for highway travel is rather sensitive to changes in highway capacity. This sensitivity, as described above, frequently results in heavy utilization and congestion of new facilities soon after their construction. It has been argued that the solution to this problem is to construct more facilities. It is possible that at some point, if this policy were followed, an equilibrium situation would indeed be reached. This conclusion can be supported by asserting that the elasticity of demand for highway travel is finite. However, if the "population" generating that demand continually increases at the same time, it is not clear that the total demand can easily be satisfied in this manner. The limit of this strategy, at the extreme, would be reached when so much land is converted to roads that land development for other purposes is restricted, with a consequent limit on trip generation and road use. It is obvious that this "equilibrium" is not the desired solution and is hopefully not feasible in any case. It will therefore be necessary to analyze possible "intermediate" solutions.

A balance between transportation facility development and land development implies a market equilibrium of the demand for use of the transportation facility and its supply, i.e. its speed and capacity characteristics. There are two basic alternatives available to the planner who wishes to modify an existing rransportation and land use situation sufficiently to achieve such a balance, though it is likely that the best strategies will be mixtures of the alternatives. The first alternative is to allow demand for transportation to fluctuate freely, with no interference, and attempt to cope with it. This would be accomplished, as in the past, by new highway construction or by implementation of mass transportation systems. The second alternative is to attempt to restrict the demand for transportation so that facilities do not become overloaded. This could be accomplished in three ways: 1) the existing transportation facility could be made more costly to use, e.g. by the imposition of tolls or by allowing congestion to develop, thus imposing a time penalty on users; 2) land development controls could be imposed, thus reducing (or slowing the growth of) trip generating activities; and 3) a mixture of these two actions could be implemented. Finally, a mixture of the two basic alternatives could be attempted, where an attempt would be made to cope with a certain amount of transportation demand (by improving transportation facilities) at the same time that an attempt was being made to control its increase. To summarize, the essence of this problem is controlling (i.e. altering, directing, and modifying) the spatial organization of the metropolis. In particular, the concern is with its spatial expansion and with the feasibility of balanced development of land use and transportation facilities.

The inherent complexity of a comprehensive analysis of both transportation and land use, along with the requirement that the analysis method be capable of providing a self-consistent procedure for testing the sensitivity and response to integrated transportation and land use control policies strongly suggest the development of integrated transportation and land use models. Serious efforts to construct urban land use simulation models began in the early 1960's. In all of these efforts the relationships between transportation and land use were only partially represented. In each case, a matrix of zone-to-zone impedances was input to a land use model. These impedances were measures of the difficulty of interaction between activities located in different zones. Depending upon the availability of data, travel times, travel costs, or airline distances (or various combinations thereof) were used as impedances. The impedances were the only link between transportation and land use in these modelling effors.

During this same period the modelling of transportation (especially highway) networks progressed to the point of rather routine application of trip generation, distribution and assignment program packages. The connection from the land use distributions to these transportation models was simply the input of the land use data to the network model package.

In the trasnportation and land use literature of the 1960's and early 1970's one will find occasional references to the desirability of more fully integrating land use and transportation models (not to mention planning). This integration of models would attempt to eliminate the principal failing of contemporary land use or transportation studies, by explicitly including feedback loops. Typically, a transportation study assumes a future land use pattern as given, and designs a transportation system to cope with it. This procedure ignores the redistributive effects which are produced by the construction of the system. Transportation systems obviously do not just suddenly appear but are constructed in stages with consequent redistribution of activities all during the period of construction. The typical land use study accepts a transportation system as given, and then estimates the consequent distribution on the network. An integrated model package would attempt to capture the interrelatedness of the transportation system and the distribution of activities.

The integration of a transportation network model with a land use model is, in principle, a rather straightforward matter. Consider the problem of forecasting the future distribution (e.g. spatial pattern) of population and employment in an urban area. We are given a description of the transportation network which will exist at that future time, along with regionwide projections of population and employment. Further we know, for a specified base year, the population and employment distributions and the trip patterns.

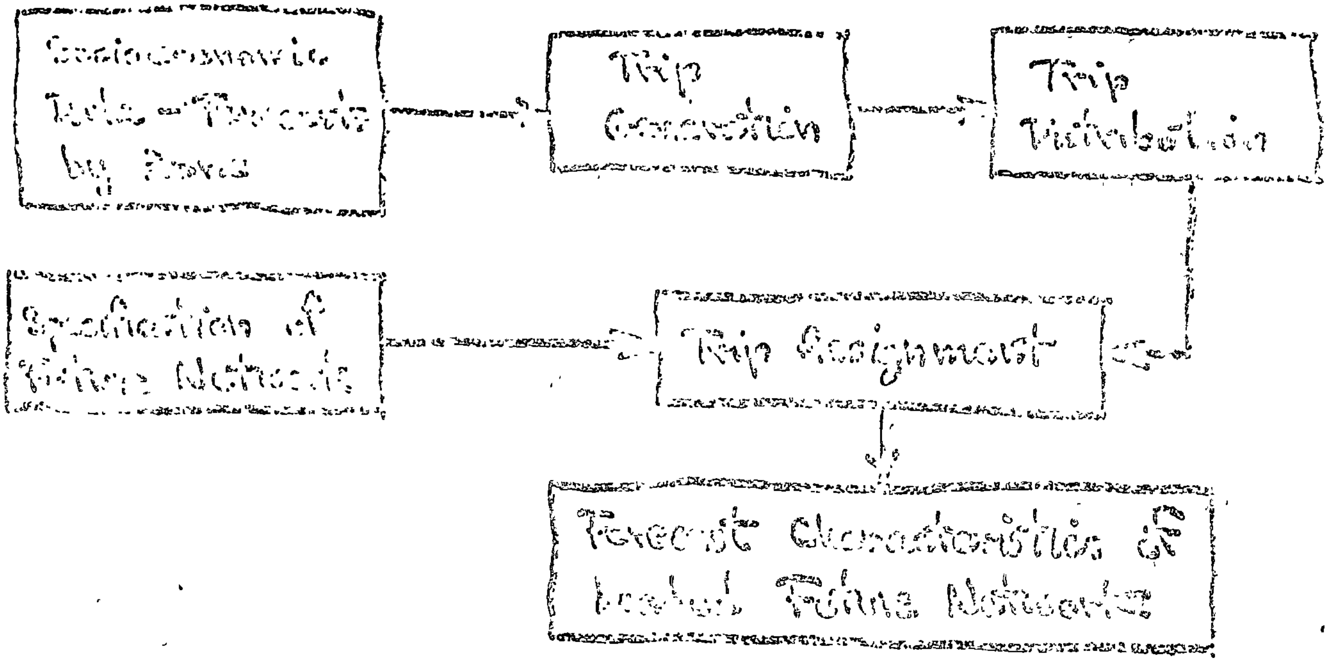
We begin the forecasting process by loading the current trip pattern on the projection year network using a capacity restrained assignment procedure. This yields an estimate of travel times on the future network in the unlikely event that both the region's activity levels and distributions remained unchanged from the base year. These travel times are used as input (in the form of a zone-to-zone impedance matrix) to the land use model, which produces the first estimate of the forecast year activity distributions. The activity distribution is, in turn, used to make the first estimate of the forecast year trip pattern (origin - destination matrix).

The second iteration begins with the assigning of the first estimate of the forecast year trip matrix to the forecast year network (the base year trips having been previously removed). This yields a second, and more likely, estimate of travel times on the forecast year network. These travel times are then input to the land use model which produces a second, revised, estimate of the forecast year activity distributions. The activity distributions are them used to generate the second, revised estimate of the forecast year trip pattern. The entire process is then repeated, until an equilibrium is reached.

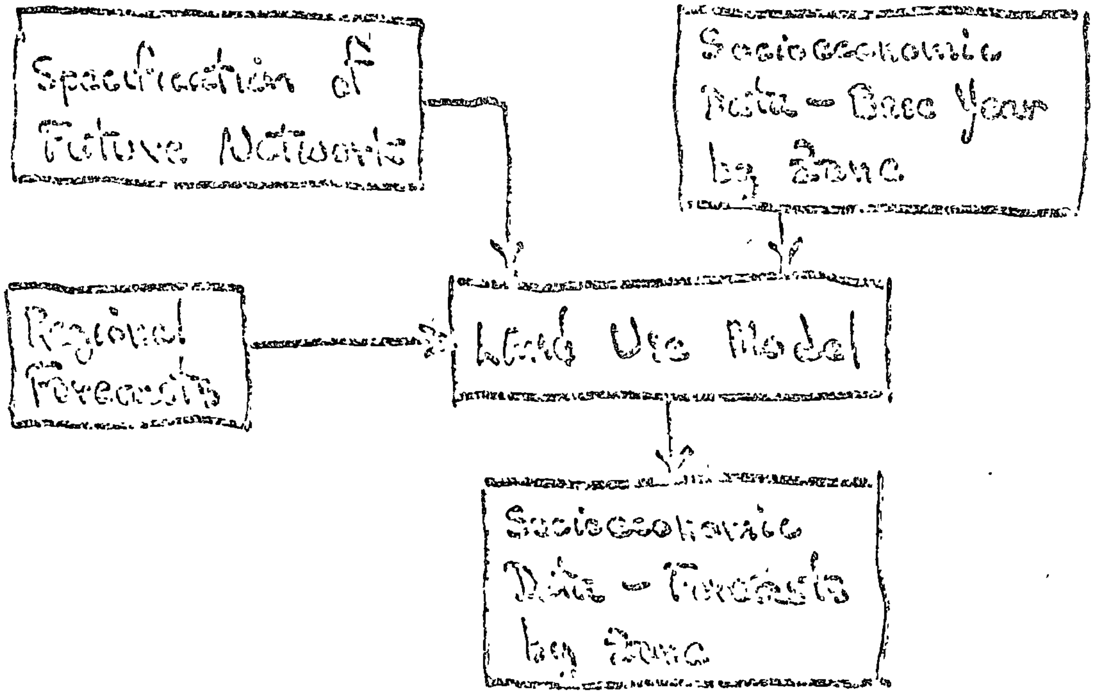
It may be helpful to consider this process with the aid of the following block diagrams. First, in Figure 1 we see simplified representations of a typical transportation planning process and a typical land use

Figure 3.

TRANSPORTATION PLANNING



LAND USE PLANNING

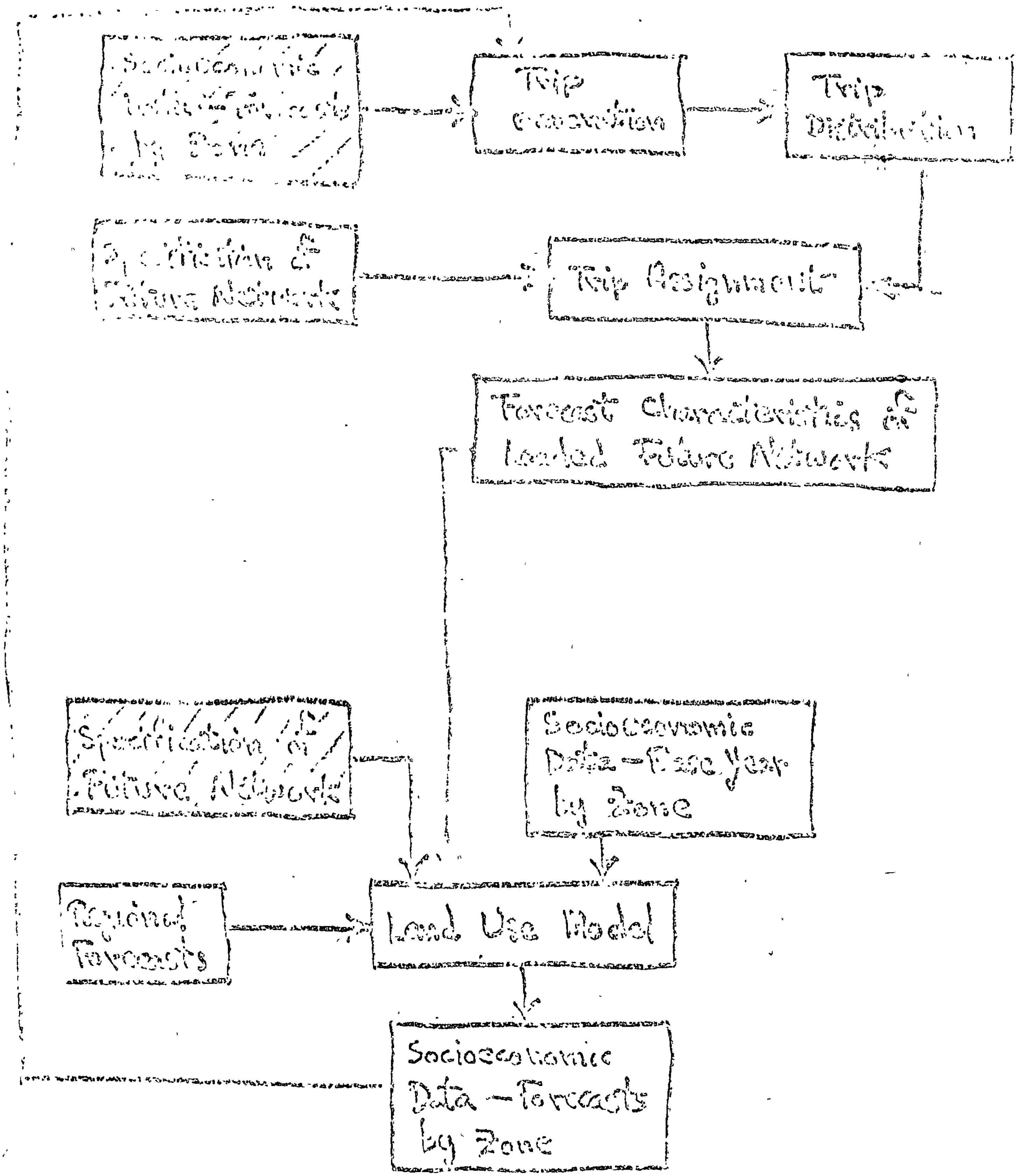


planning process.

In Figure 2 we see how these two processes may be linked, first by using the outputs of the transportation planning process as input to the land use planning process, and second, by using the outputs of the land use planning process as input to the transportation planning process. Having shown this connection, it becomes clear as to why it may be necessary to iterate (i.e. repeat) the steps in the process several times before coming to an equilibrium solution.

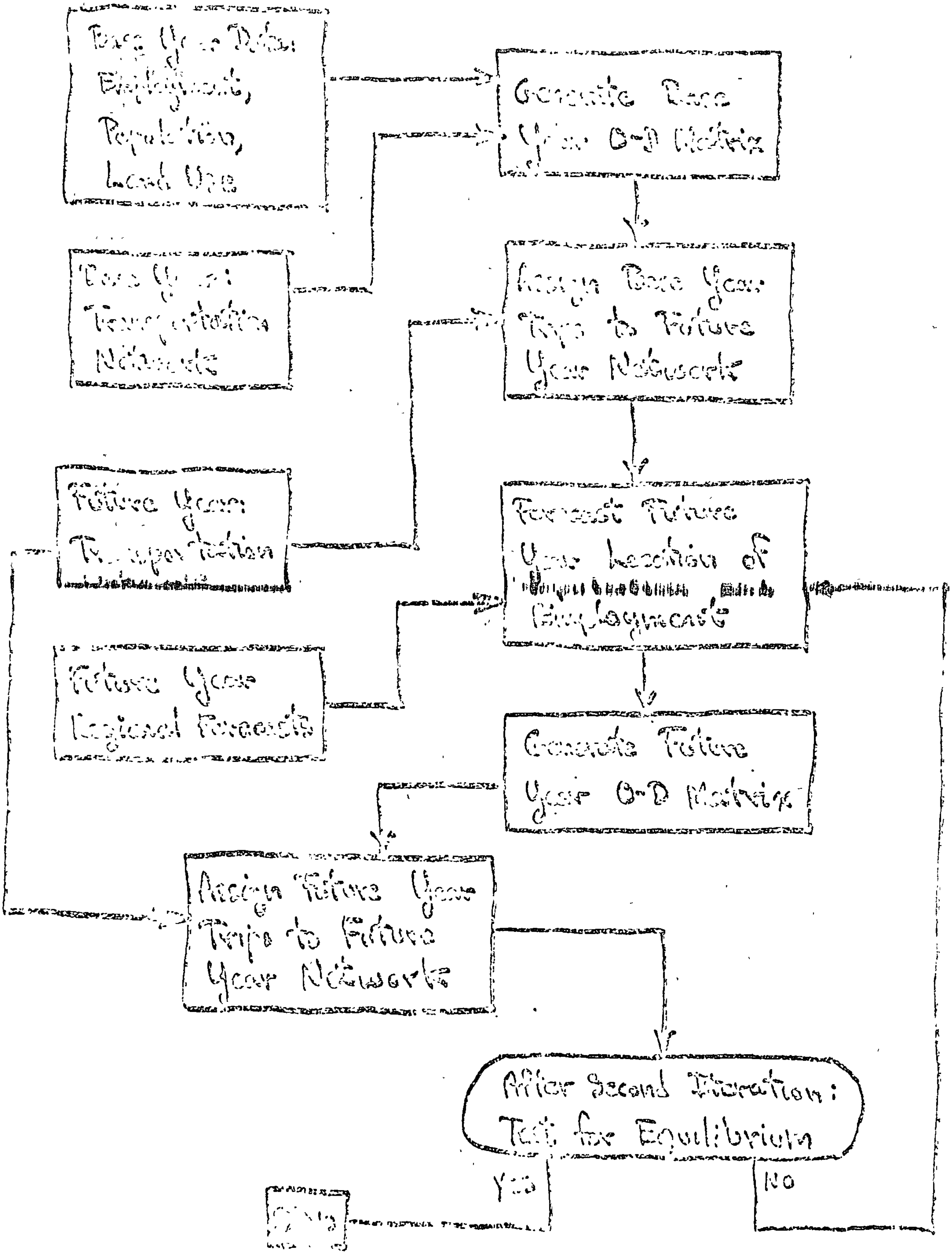
In Figure 3 we see a translation of the integrated transportation planning and land use planning processes into a set of model procedures. First a set of base year data on both land use and transportation is used to estimate a base year O.D. trip matrix. These trips are then loaded on (assigned to) the future year transportation network. The time and cost characteristics of this partially congested network are used, along with regional economic forecasts, as input to forecasts of the future year land use pattern. From these future year land use forecasts, we obtain a future year O.D. trip matrix. These estimates of future year trips are then assigned to the guture year network. It is certain that at least two cycles through this process will be required. After these first two cycles the results (forecasts) are checked for equilibrium, after which the process is terminated.

Figure 2



INTERACTION OF LAND USE & TRANSPORTATION

Fig. 2


Base Year Data: Employment, Population, Land Use
Base Year Transportation Network
Generate Base Year O-D Matrix
Assign Base Year Trips to Future Year Network
Future Year Transportation
Forecast Future Year Location of Population and Employment
Future Year Regional Forecasts
Generate Future Year O-D Matrix
Assign Future Year Trips to Future Year Network
After Second Iteration: Test for Equilibrium
Yes
No
Stop

ITLUP

The first operational version of ITLUP followed the above procedure rather closely. In that version of the package, the network model used an incremental tree-by-tree, capacity restrained assignment algorithm. The basic employment distribution was an exogenous input. The population and non-basic employment distributions were estimated with a Lowry-derivative model (a modified version of the incremental PLUM model).

Policy and sensitivity tests were run with this version of ITLUP. The most important methodological conclusion from this work was a clear demonstration of the need for an integrated package to make accurate forecasts. Use of the same land use model in an unintegrated non-iterative forecasting procedure yielded less accurate forecasts than those obtained from ITLUP. The most important substantive conclusion from this work was that improvements in a region's transportation system will, in the absence of coordinated land use controls, usually lead to further urban sprawl and congestion along with continual demand for new transportation improvements which will serve to repeat the cycle. The reverse was also found to be true; allowing the development of transportation congestion reduced tendencies to urban spral and led to the development of a regional pattern of clusters of denser development.[1]

Since the publication of the initial results from ITLUP, considerable additional work has been done.[2] Subsequently other researchers have

---

[1] Putman, S. H. (1973) "The Interrelationships of Transportation Development and Land Development";

Putman, S. H. (1974) "Preliminary Results from an Integrated Transportation and Land Use Models Package".

[2] Putman, S. H. (1976) "Further Results from the Integrated Transportation and Land Use Model Package (ITLUP)".

published findings which corroborate the ITLUP results.[1]

In 1973 another project was begun, under N.S.F. sponsorghip, with the intent of comparing various operational land use forecasting models.[2] During this project the residential submodel of ITLUP was reformulated to allow for its proper calibration.[3] This reformulation was based on the Wilson maximum entropy approach.[4] Further work with this form of the model, called DRAM (Disaggregated Residential Allocation Model) has produced important results not only from the standpoint of parameter estimating and forecasting, but for the more direct integration of transportation and land use models.

There appears to be a great deal of misunderstanding of the maximum entropy derivation of these new transportation and land use models. Further, it seems that some of their more important implications are quite overlooked in U.S. planning practice. Consequently, the next section of this paper will present a simple and hopefully lucid discussion of the underlying basis for these models while the following section will present an illustration of one such model.

Spatial Interaction Models and Maximum Entropy

It is well known that the initial development of transportation and

---

[1] Berechman, J. (1976) "Interfacing the Urban Land-Use Activity System and the Transportation System", Journal of Regional Science, Vol. 16, No. 2, pp. 183-194.

Peskin, R. L. and J. L. Schofer (1976) "Assessment of Transportation Energy Conservation Strategies through Simulation" presented at Fourth Inter-society Conference on Transportation, Los Angeles, California, July.

[2] Putman, S. H. (1976) "Laboratory Testing of Predictive Land-Use Models: Some Comparisons".

[3] Putman, S. H. (1976) "Calibrating a Disaggregated Residential Model-DRAM".

[4] Wilson, A. G. (1970) Entropy in Urban and Regional Modelling.

land use models was largely based on the observed fit of these phenomena to the Newtonian gravity model. While the descriptive validity of these gravity models was reasonably good, a number of persistant doubts remained. One of the most serious of these, from a theoretical point of view, was the question of why human spatial interactions should resemble the interactions of planetary bodies.

The entropy maximizing derivation of spatial interaction models completely obviates the use of the gravity model analogy. The derivation thus depends not on a rather implausible analogy to a physical phenomenon, but on an analogy with statistical mechanics. Further, it has been shown in the literature that several models which derive from concepts of micro-economic behavior, e.g. travel cost minimizing and to a certain extent market-clearing behaviors, are compatible with the maximum entropy approach.[1]

The principles of the maximum entropy approach may be shown with the following simple example.[2] Imagine six employed persons living in a residence zone i and commuting to various work zones j. Let there be one residence zone, i=1 and three work zones corresponding to j=1, 2, 3. Suppose that the six workers are named A,B,C,D,E,F. We may now specify the origins and destinations of the worktrips of each worker. Each possible, fully described, system of a) origin, b) three destinations, and c) six worktrips with their origins and destinations may be called a microstate of the system. Six of these possible microstates are shown in Figure 1. There are obviously very many microstates of even this simple system.

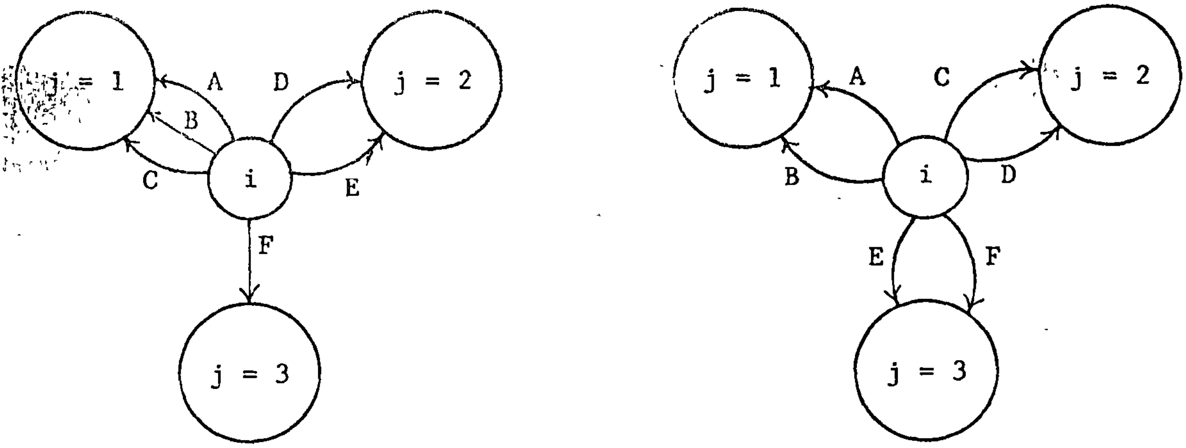
---

[1] Senior, M. S. (1973) "Approaches to Residential Location Modelling 1: Urban Ecological and Spatial Interaction Models (A Review)" Environment and Planning Vol. 5, pp 165-197.

Nijkamp, P. (1975) "Reflections on Gravity and Entropy Models" Regional Science and Urban Economics, Vol. 5, pp 203-225.

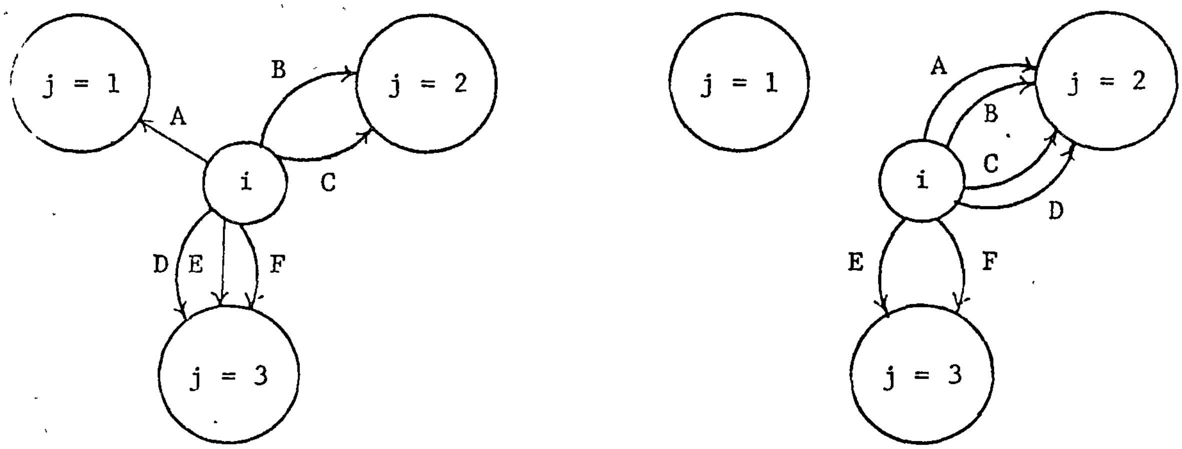
Choukroun, J. M. (1975) "A General Framework for the Development of Gravity-Type Trip Distribution Models" Regional Science and Urban Economics Vol. 5, pp 177-202.

[2] Drawn in part from Senior (1973), op.cit.



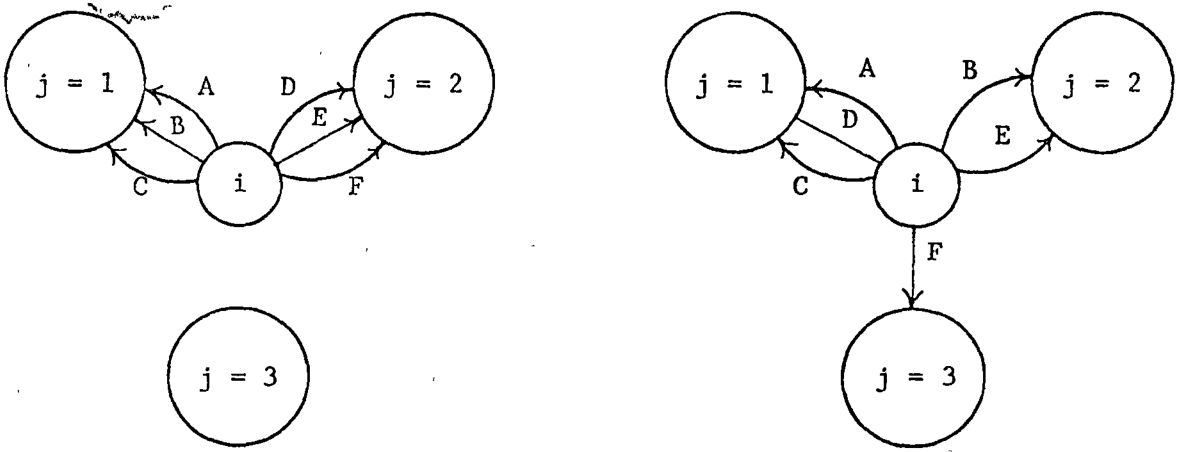
Microstate 1

Microstate 2



Microstate 3

Microstate 4



Microstate 5

Microstate 6

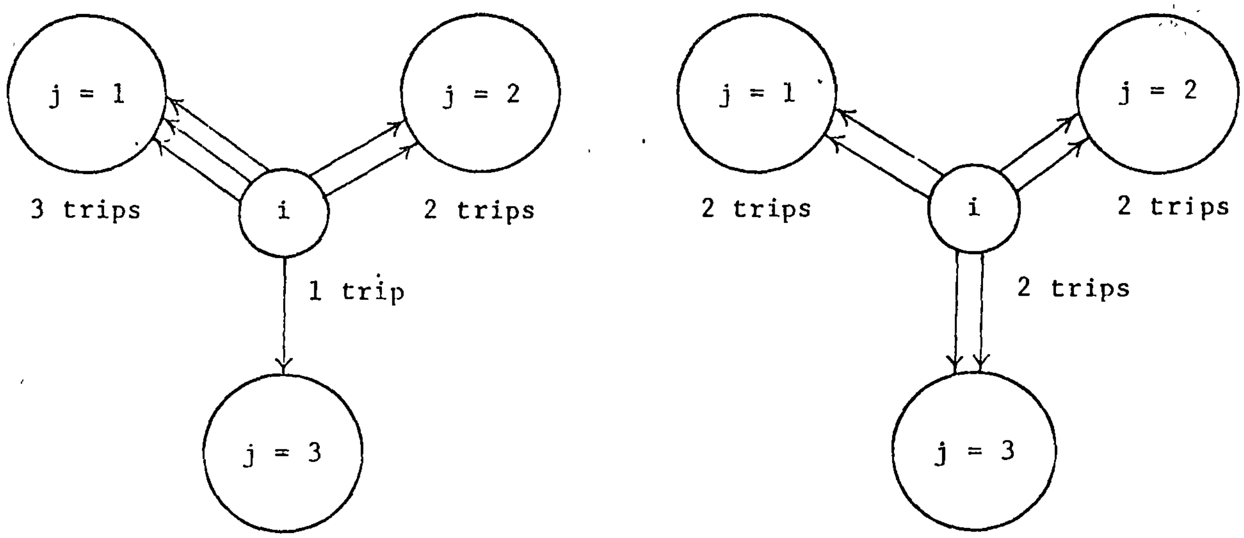
Figure 1: System Microstates

Let us now consider Microstate 1 where the trips between i and $j=1$ are 3; between i and $j=2$ there are 2 trips, and the total trips between i and $j=3$ are 1. Microstate 6 may also be seen to have his same distribution of trips, from i to $j=1$ there are 3, $i \rightarrow 2=2$, $i \rightarrow 3=1$. Clearly there are many microstates which could be drawn and which would have this same arrangement of total trips. This particular arrangement of zone-to-zone trips, if described independently of <u>which</u> worker is making which trip, may be called a mesostate of the system. Four mesostates of the system are shown in Figure 2.

Comparing Figures 1 and 2, it can be seen that Microstates 1 and 6 are possible microstates of Mesostate A. Microstate 2 is a possible microstate of Mosostate B. Microstate 5 is <u>not</u> a possible microstate of Mesostate D. Thus each mesostate describes a set of possible microstates.

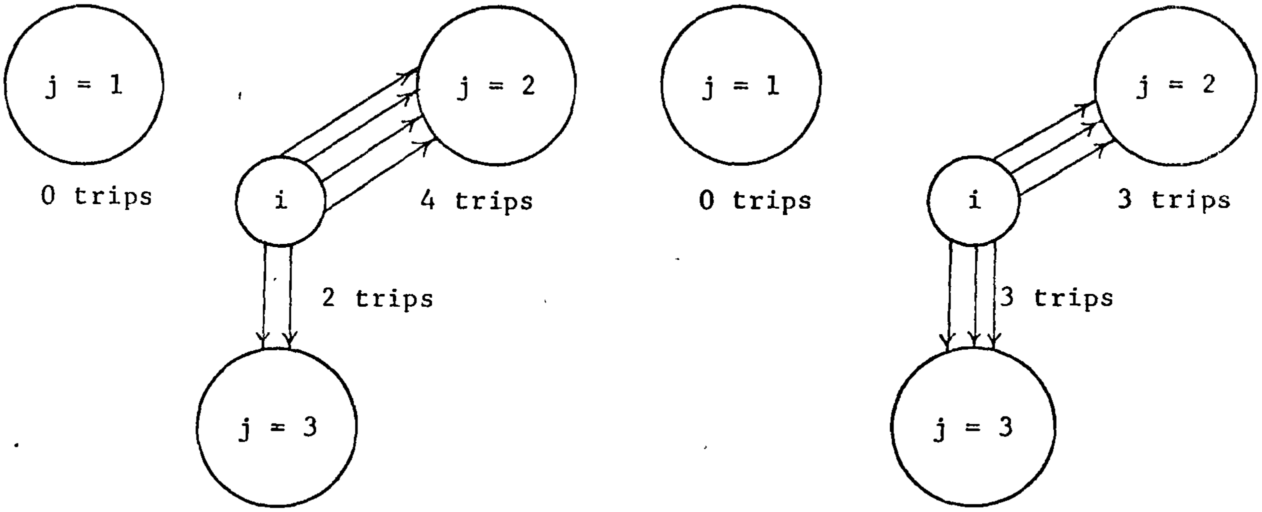
If we now consider that there might be several residence zones in addition to the one which has been used in this example then a more aggregate level of dexcription of the system would be, the total trips leaving each origin and the total trips arriving at each destination. Let us assume that two workers live in zone $i=2$ and four workers live in zone $i=3$ in addition to the six already defined as living in $i=1$. Further assume that these additional workers are named G,H,I,J,K,M.

A microstate of this newly expanded system would be a list of the origins and destinations of the worktrips of each of the twelve workers. A mesostate of this system would be a list of the total number of worktrips from each origin zone to each destination zone. Finally, a <u>macrostate</u> of this expanded system is a list of the total trips leaving each origin and the total trips arriving at each destination. Figure 3 shows four macrostates of the expanded system.



Mesostate A

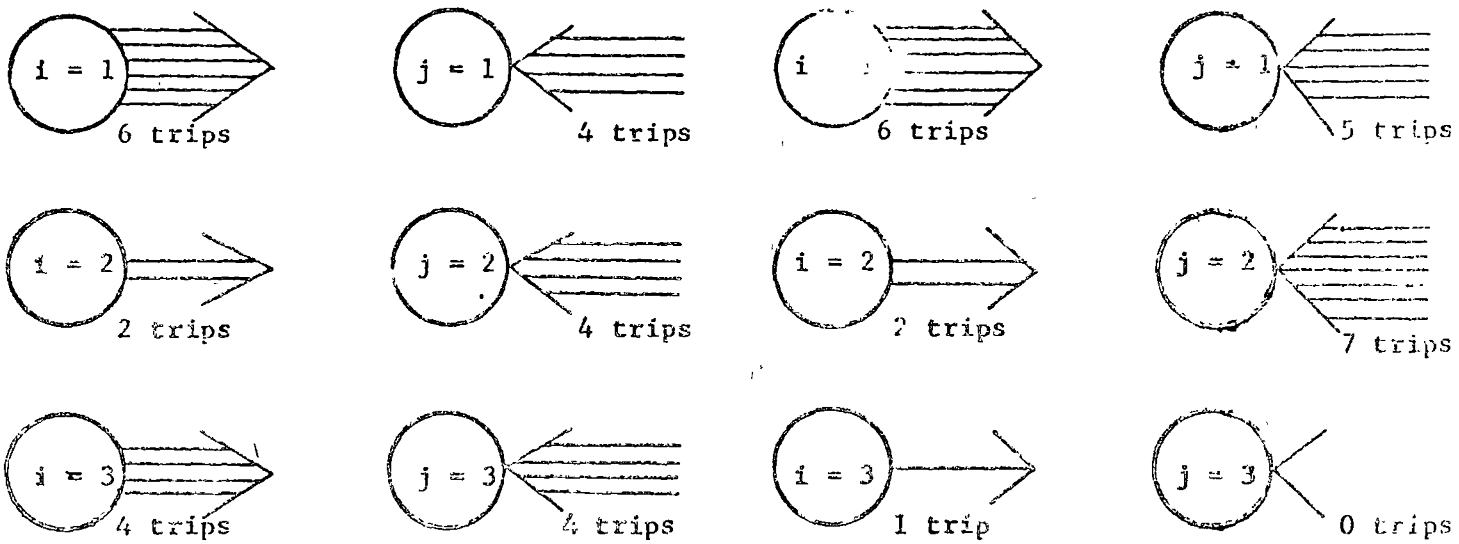
Mesostate B



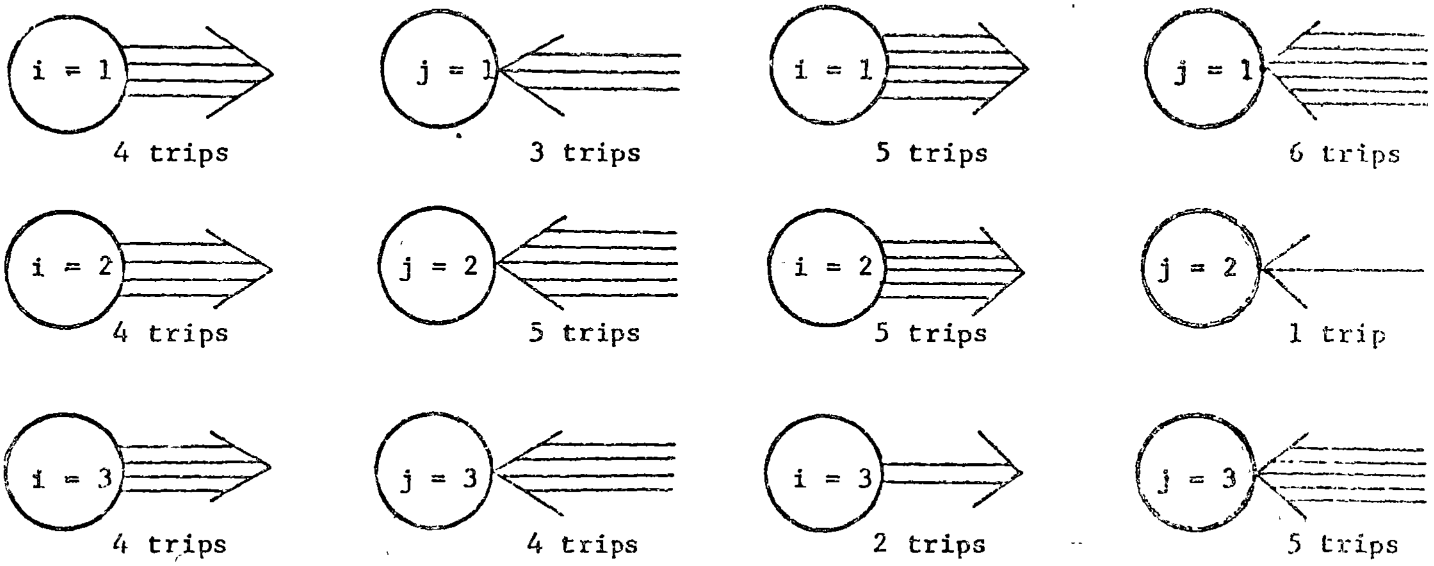
Mesostate C

Mesostate D

Figure 2: System Mesostates



Macrostate 1 Macrostate 2



Macrostate 3 Macrostate 4

Figure 3

Referring to Figure 3, Macrostates 1 and 2 with six trips leaving i=1 contain all the previous examples of microstates and mesostates. Macrostates 2 and 3 with the trips leaving i=1 not equal to six, correspond to other system states which do not include the microstates and mesostates given as examples. We should also note in passing that one could have defined a macrostate for the example of a single origin used at the start of this discussion. This would have been in a sense, a degenerate case, as the trips leaving the single origin would always have been equal to six.

In operational urban simulation models virtually no attempt is made to simulate at the level of microstates i.e. individual behavior. Most of the models operate at the mesostate and/or macrostate level. The entropy approach deals with these, and requires two key assumptions. First, all microstates are assumed to be equally probable. Second, the most likely, mesostate or macrostate is assumed to be the one with the greatest number of possible microstates.

We may now easily derive a spatial interaction model.[1] First, defining the variables.

T_{ij} = the number of workers living in i and working in j (this is the variable to be estimated)

O_i = the number of workers living in i (a given)

D_j = the number of jobs in j (given)

C_{ij} = the cost of travel from i to j (given)

C = the total travel expenditure of the system (given)

In addition there are several constraints on the system. First, the sum over all j destination zones of all trips arriving from zone i is equal to the total trips leaving i. That is

---

[1] Wilson, A. G. (1970) Entropy in Urban and Regional Modelling Pion Limited, London pp 1-14.

$$\sum_j T_{ij} = O_i \tag{1}$$

Second, the sum over all i origin zones of all trips going to j is equal to the total trips arriving in j. That is

$$\sum_i T_{ij} = D_j \tag{2}$$

Finally, the total travel cost is equal to the sum of, the trips between each i-j pair times the cost of travelling between that pair.

That is

$$\sum_j \sum_j T_{ij} c_{ij} = C \tag{3}$$

Now, the most probable mesostate is the one with the maximum number of possible microstates, subject to constraints (1), (2), and (3). Consequently what is desired is the description (i.e. equation) of the matrix $\{T_{ij}\}$ which has the greatest number of microstates, i.e. the greatest number of ways, $W(\{T_{ij}\})$, of getting $\{T_{ij}\}$ given the constraints. The number of ways of getting $\{T_{ij}\}$ is a problem of combinations of individual workers to given origin-destination pairs. The microstates in Figure 1 represent six of many possible combinations of one origin, three destinations, and six workers. If T is the total number of workers, i.e.

$$T = \sum_i O_i = \sum_j D_j = \sum_i \sum_j T_{ij}$$

then by combinational theory

$$W(\{T_{ij}\}) = \frac{T!}{\prod_{ij} T_{ij}!} \tag{4}$$

For the Figure 1 example:

i=1 j=3 T=6

Then[1]

$$W(\{T_{ij}\}) = \frac{6!}{(T_{11}!)(T_{12}!)(T_{13}!)}.$$

And, if we wish to know the number of microstates for Mesostate A in Figure 2.

$$W(\{T_{ij}\}) = \frac{6!}{3!\ 2!\ 1!} = \frac{720}{6(2)} = 60$$

The number of microstates for Mesostate B in Figure 2.

$$W(\{T_{ij}\}) = \frac{6!}{2!\ 2!\ 2!} = \frac{720}{8} = 90$$

By trial and error, one may substitute values in the denominator of this equation and discover that the minimum value of the denominator (subject to the constraint that the sum of all the trips equals six) is at 2!2!2! Thus the maximum value of W is 90, which suggests that in the absence of any further information about the system of our example, the most probable mesostate is when the six trips are evenly distributed to the three destinations.

We recall however that the it is the equation for the T_{ij} which we are trying to determine. Consequently we must maximize $W(\{T_{ij}\})$ as given in equation (4) subject to the constraints (1), (2), and (3). This is mathematically possible, and involves the use of Lagrangian multipliers and Stirling's approximation to the numerical value of large factorials.[2] The result gives

$$T_{ij} = A_i B_j O_i D_j \exp(-\beta c_{ij}) \qquad (5)$$

where

$$A_i = \frac{1}{\sum_j B_j D_j \exp(-\beta c_{ij})} \qquad (6)$$

---

[1] Note 6! means 6 factorial, which equals 6x5x4x3x2x1=720. Also note that 0! is defined to be 1. Further note that $\prod$ is the product operator, as $\sum$ is the sum operator.

[2] Stirling's approximation for large values of x yields; $\ln x! = x \ln x - x$

and

$$B_j = \frac{1}{\sum_i A_i O_i \exp(-\beta c_{ij})} \qquad (7)$$

where β is an empirically derived parameter, or if C in equation (3) were known, β could be solved for numerically.

Before proceeding, we should note that the name maximum entropy for this derivation stems from the fact that equation (4) is defined, in statistical mechanics, to be the entropy of the system. It is, of course, equation (4) and therefore the system entropy which we maximize to derive equations (5), (6) and (7). At this same point, it must be mentioned that there are alternative interpretations of entropy which may be used. From a practical point of view the results lead in the same direction which is that a spatial interaction model whose equations resemble the traditional gravity model may be derived from assumptions which do not include any reference to Newtonian gravitational phenomena, but which do refer to probability statements and "most probable"distributions.

Having shown the derivation of a spatial interaction model via the maximum entropy approach, we may also show its relationship to a transportation cost minimizing approach of the sort which economists assert should constitute the underlying behavior of spatial interaction. A simple cost minimizing approach may be constructed as follows. Suppose that there are a given number of trips D_j, terminating in all zones. Suppose that each trip T_{ij} has a cost c_{ij} of travelling between i and j. The problem may then be stated as:

$$\text{Minimize:} \quad \sum_i \sum_j T_{ij} \; c_{ij}$$

$$\text{Subject to:} \quad \sum_i T_{ij} = D_j$$

$$\sum T_{ij} = 0_i$$

This is known as the "transportation problem" of linear programming and often has additional constraints as part of the problem formulation.

Now, consider the entropy maximizing approach as presented above. It may be shown algebraically that maximizing W is equivalent to maximizing ln W and by further manipulation that

$$\text{Maximizing:} \quad W\left(\{T_{ij}\}\right) = \frac{T!}{\Pi_{ij} \; T_{cj}!}$$

is equivalent to

$$\text{Maximizing:} \quad -\sum_i \sum_j T_{ij} \ln T_{ij}$$

Further we note that in general, maximizing a function is the same as minimizing its negative. That is

$$\text{Max: } f(x) = \text{Min: } -f(x)$$

Consequently, the entropy maximizing approach may be rewritten as

$$\text{Minimize:} \quad \sum_i \sum_j T_{ij} \ln T_{ij}$$

$$\text{Subject to:} \quad \sum_i T_{ij} = D_j$$

$$\sum_j T_{ij} = 0_i$$

$$\sum_i \sum_j T_{ij} \; c_{ij} = C$$

Suppose now that for a particular problem situation there is a minimum numerical value M, of entropy for this problem. Then we may consider the relationship between the cost minimizing formulation and the entropy minimizing formulation in the following problem statement where we simply add one more constraint to the cost minimizing problem.

$$\text{Minimize:} \quad \sum_i \sum_j T_{ij} c_{ij}$$

$$\text{Subject to:} \quad \sum_i T_{ij} = D_j$$

$$\sum_j T_{ij} = O_i$$

$$\sum \sum T_{ij} \ln T_{ij} \geq M$$

As this last constraint is non-linear, the solution of the problem in this form is quite difficult. However, its statement in this form shows that the entropy constraint simply adds a "noise" level into the cost minimizing problem. Obviously, if $M = 0$, ther is no "noise" and thus the expected spatial interactions will exactly equal the cost minimizing solution.

At the level of detail appropriate to land-use and transportation modelling, say 100 to 200 zones for a metropolitan area of several counties, the entropy derivation seems relatively satisfactory. Thought of in terms of viewing the metropolitan area from about a mile above it, the statistical mechanics, most probable distribution, is a very reasonable way to describe the observable phenomena.

In the next section, a residential allocation model based on these principals will be developed. This will be followed by a discussion and numerical example of its proper calibration.

A Simple Residential Location Model

In developing equations (5), (6), and (7) it was assumed that both O_i (number of workers living in area i) and D_j (number of workers employed in area j) were known. In a residential location model we are trying to estimate O_i. We therefore replace O_i in those equations with W_i, a measure of the residential attractiveness of area i. When we do this we eliminate the need for the origins balancing factor A_i. This gives the following equations:

$$T_{ij} = B_j W_i D_j \exp(-\beta c_{ij}) \tag{8}$$

where

$$B_j = \frac{1}{\sum_i W_i D_j \exp(-\beta c_{ij})} \tag{9}$$

But, T_{ij} is the number of workers living in area i and working in area j. We wish N_i, the number of workers living in area i, to be the dependent variable, so we sum over all workplaces. This gives

$$N_i = \sum_j T_{ij} = \sum_j B_j W_i D_j \exp(-\beta c_{ij}) \tag{10}$$

If we now substitute equation (9) into equation (10) we get

$$N_i = \sum_j \left[\frac{W_i D_j \exp(-\beta c_{ij})}{\sum_i W_i D_j \exp(-\beta c_{ij})} \right] \tag{11}$$

and, simplifying this equation, we get

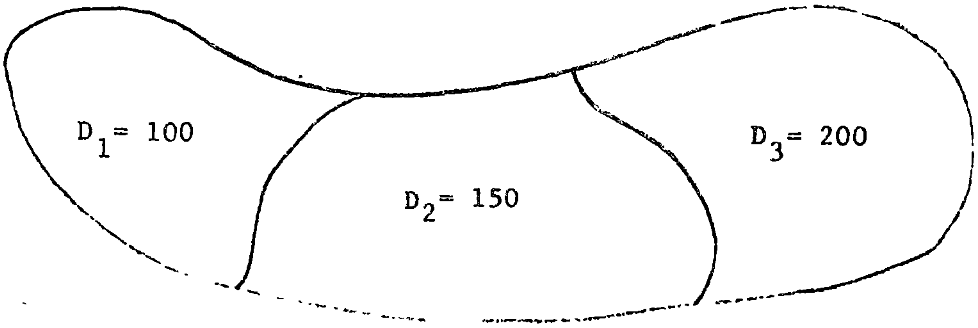
$$N_i = \sum_j D_j \left[\frac{W_i \exp(-\beta c_{ij})}{\sum_i W_i \exp(-\beta c_{ij})} \right] \tag{12}$$

It may be helpful to note that the term (fraction) within the brackets in equation (12) is simply the relative attractiveness-accessibility of an individual area i compared to all other areas in the region being modelled.

Thus equation (12) is merely stating that the number of workers living in i is the sum over all j of the workers employed in each j times the probability that workers employed in that j will live in i.

An obvious question here is how to measure the attractiveness W_i of each area i? Many such measures have been proposed. Current work with a version of this model indicates that a description of the prior (base year) household composition of area i along with a measure of the available land and its degree of development provide an adequate description of the area's attractiveness.[1] At this point a numerical example of the model may be helpful to understanding its operation.

Assume a region divided into three zones. In each of these zones there is a known quantity of employment.



Suppose further that the cost of travel between zones is given by the following matrix.

Zone	1	2	3
1	1.5	2.5	4.0
2	2.5	2.0	3.5
3	3.0	3.5	1.5

---

[1] Putman, S. H. (1976) "Calibrating ...", op.cit.

Next, we assume that some measure of residential attractiveness has been defined for each zone, based perhaps on its population composition and, say, density. So, $W_1 = 3$, $W_2 = 4$, $W_3 = 5$.

Finally, without discussing how we found its value, let us assume that the value of β is 2.0.

We may first calculate the matrix of $\exp(-2.0c_{ij})$. For example when c_{ij} equals 1.5 then $\exp(-2.0\text{x}1.5) = \exp(-3.0) = 0.0498$. So, filling in the new matrix of $\exp(-2.0c_{ij})$

Zone	1	2	3
1	0.0498	0.0067	0.0003
2	0.0067	0.0183	0.0009
3	0.0025	0.0009	0.0498

Starting with the first zone, we may now calculate its employed residents.

$$N_1 = D_1\left[\frac{W_1\exp(-2.0c_{11})}{W_1\exp(-2.0c_{11}) + W_2\exp(-2.0c_{21}) + W_3\exp(-2.0c_{31})}\right] +$$

$$D_2\left[\frac{W_1\exp(-2.0c_{12})}{W_1\exp(-2.0c_{12}) + W_2\exp(-2.0c_{22}) + W_3\exp(-2.0c_{32})}\right] +$$

$$D_3\left[\frac{W_1\exp(-2.0c_{13})}{W_1\exp(-2.0c_{13}) + W_2\exp(-2.0c_{23}) + W_3\exp(-2.0c_{33})}\right]$$

Numerically this is

$$N_1 = 100\left[\frac{3(0.0498)}{3(0.0498) + 4(0.0067) + 5(0.0025)}\right] +$$

$$150\left[\frac{3(0.0067)}{3(0.00.67) + 4(0.01.83) + 5(0.0009)}\right] +$$

$$200\left[\frac{3(0.0003)}{3(0.0003) + 4(0.0009) + 5(0.0498)}\right]$$

$$N_1 = 100 \left[\frac{0.1494}{0.1887}\right] + 150 \left[\frac{0.0201}{0.0978}\right] + 200 \left[\frac{0.0009}{0.2535}\right]$$

$$N_1 = 79.17 + 21.58 + 0.71 = 110.7$$

Similarly we may calculate

$N_2 = 129.3$ $N_3 = 210.0$

Thus we have 111 employed residents in zone 1, 129 in zone 2, and 210 in zone 3. Note that without using a regional control total for employed residents, their sum is 450, the number of persons employed in our closed region. To get total population from employed residents regional or zone-specific multipliers are used.

Policy tests may be made with this model by changing its inputs. Suppose, for example that a transportation improvement between zones 1 and 2 produces a 0.5 reduction in travel cost between them, yielding the following cost matrix

Zone	1	2	3
1	1.5	2.0	4.0
2	2.0	2.0	3.5
3	3.0	3.5	1.5

Recalculating the employed residents yields 123 in zone 1, 117 in zone 2, and 210 remaining unchanged in zone 3.

Another policy might result in 50 jobs moving from zone 3 to zone 2. This gives $D_1 = 100$, $D_2 = 200$, and $D_3 = 150$.

If, using the original travel costs, this new employment distribution was used as input the resulting distribution of employed residents would be 121 in zone 1, 166 in zone 2, and 162 in zone 3.

Calibration of the Simple Residential Model

In the numerical example above, the value of β in equation (12) was assumed to equal 2.0. In model calibration the value of β, plus other parameters which may appear in more complex forms of this model, is precisely what needs to be estimated.

Consider a situation where the employed residents, residential attractiveness, employment and zone-to-zone travel cost are known for a given set of zones at a given point of time. We are interested in finding the value of β. Due to the nonlinearity of equation (12) it is not possible to use a standard regression procedure to find β. What is required is some form of non-linear search procedure which finds a best value of β in an efficient fashion.

Briefly stated, let $\hat{N}_i$ be an estimate of N_i as per,

$$\hat{N}_i = \sum_j D_j \left[\frac{W_i \exp(\hat{\beta} c_{ij})}{\sum_i W_i \exp(\hat{\beta} c_{ij})} \right] \tag{13}$$

Then we must define a criterion to describe how closely N_i is matched by $\hat{N}_i$, and then find the value of $\hat{\beta}$ which produces the best possible value of the criterion.

There are several different criterion functions which might be chosen, such as R^2, maximum likelihood, or minimum relative error squared. There are arguments for and against different criteria in different problems. We will continue the discussion based on R^2 as a criterion. This may be defined as:

$$R^2 = 1 - \left[\frac{\sum_i (\hat{N}_i - N_i)^2}{\sum_i (N_i - \bar{N}_i)^2} \right] \tag{14}$$

where, again

N_i = observed employed residents in zone i

$\hat{N}_i$ = estimated employed residents in zone i

$\bar{N}_i$ = mean of N_i

Given this function one may find a $\hat{\beta}$ which maximizes R^2, by a trial and error procedure. In actual practice this would be a tiresome procedure. A more efficient alternative would be to use gradient search.

The gradient of any function may be calculated from the partial derivatives of that function with respect to each of its variables. For example, for a given function, say $\mathcal{F}(x,y,z)$, the gradient $\nabla\mathcal{F}$ is given by

$$\nabla\mathcal{F} = i\frac{\partial\mathcal{F}}{\partial x} + j\frac{\partial\mathcal{F}}{\partial y} + k\frac{\partial\mathcal{F}}{\partial z}$$

Note that each term in the above vector has been multiplied by a constant (the directional unit vector) in each coordinate direction. Now, suppose we let

$$w = \mathcal{F}(x,y,z) = x^2 + y^2 - z$$

then

$$\frac{\partial w}{\partial x} = 2x; \quad \frac{\partial w}{\partial y} = 2y; \quad \frac{\partial w}{\partial z} = -1$$

If the directional unit vectors in each coordinate direction are defined as i, j, and k respectively, then

$$\nabla w = i2x + j2y - k$$

Thus, the gradient is a vector orthogonal (at right angles, or perpendicular) to a particular mathematical surface. If the gradient is projected back on the mathematical surface, it points in the direction of steepest ascent. More simply, gradient search is a sophisticated "hill climbing" procedure.

In doing a gradient search the gradient, ∇, of the criterion, with respect to the parameter(s), must be found. For our case,

$$\nabla R^2 = \frac{\partial R^2}{\partial \hat{\beta}} = \frac{\partial R^2}{\partial \hat{N}_i}\left(\frac{\partial \hat{N}_i}{\partial \hat{\beta}}\right) = -2\left[\frac{\sum_i \left[(\hat{N}_i - N_i)\left(\frac{\partial \hat{N}_i}{\partial \beta}\right)\right]}{\sum_i (N_i - \bar{N}_i)^2}\right] \tag{15}$$

To continue, we refer to equation (13) and to simplify the notation, let

$$W_i \exp(\hat{\beta} c_{ij}) \rightarrow L_{ij} \tag{16}$$

$$\sum_i W_i \exp(\hat{\beta} c_{ij}) \rightarrow M_j \tag{17}$$

Now, substituting equations (16) and (17) into (13),

$$\hat{N}_i = \sum_j D_j \left[\frac{L_{ij}}{M_j}\right] = \sum_j D_j L_{ij} M_j^{-1}$$

Then, taking the partial derivative

$$\frac{\partial \hat{N}_i}{\partial \beta} = \sum_j D_j \left[M_j^{-1}\left(\frac{\partial L_{ij}}{\partial \hat{\beta}}\right) + L_{ij}\left(\frac{\partial (M_j^{-1})}{\partial \hat{\beta}}\right)\right] \tag{18}$$

Again, taking the partial derivative

$$\frac{\partial L_{ij}}{\partial \hat{\beta}} = W_i c_{ij} \exp(-\hat{\beta} c_{ij}) \tag{19}$$

and substituting again from equation (16) .

$$\frac{\partial L_{ij}}{\partial \hat{\beta}} = c_{ij} L_{ij} \tag{20}$$

Again, taking the partial derivative

$$\frac{\partial (M_j^{-1})}{\partial \hat{\beta}} = - M_j^{-2} \sum_k W_k c_{kj} \exp(-\hat{\beta} c_{kj}) \tag{21}$$

and substituting again from equation (16)

$$\frac{\partial (M_j^{-1})}{\partial \beta} = - M_j^{-2} \sum_k c_{kj} L_{kj} \tag{22}$$

Now, substituting equations (20) and (22) into (18),

$$\frac{\partial \hat{N}_i}{\partial \beta} = \sum_j D_j \left[M_j^{-1} c_{ij} L_{ij} + L_{ij} \left(-M_j^{-2} \sum_k c_{kj} L_{kj} \right) \right] \tag{23}$$

$$= \sum_j D_j L_{ij} M_j^{-1} \left[c_{ij} - M_j^{-1} \sum_k c_{kj} L_{kj} \right] \tag{24}$$

This result can then be substituted back into equation (15) to yield

$$\nabla R^2 = -2 \left[\frac{\sum_i \left\{ (\hat{N}_i - N_i) \left[\sum_j \left(\frac{D_j L_{ij}}{M_j} \right) \left(c_{ij} - \frac{\sum_k c_{kj} L_{kj}}{M_j} \right) \right] \right\}}{\sum_i (N_i - \bar{N}_i)^2} \right] \tag{25}$$

Now, a numberical example may serve to clarify the purpose of these machinations.

Let us begin with the following "observed" data.

$D_1 = 100$

$D_2 = 150$

$D_3 = 200$

$W_1 = 3$

$W_2 = 4$

$W_3 = 5$

$N_1 = 111$

$N_2 = 129$

$N_3 = 210$

and

$C_{11} = 1.5$ $\quad$ $C_{21} = 2.5$ $\quad$ $C_{31} = 4.0$
$C_{21} = 2.5$ $\quad$ $C_{22} = 2.0$ $\quad$ $C_{23} = 3.5$
$C_{31} = 3.0$ $\quad$ $C_{23} = 3.5$ $\quad$ $C_{33} = 1.5$

This data set represents the data which would normally be collected for use in calibrating an actual land use model. The question in such a case would be to determine a value of β that would give the best fit between the models estimates of $\hat{N}_i$ and the observed N_i values. In this example, we know that a value of -2.00 for β would give a perfect fit. (Because, the observed N_i in this example were obtained in the previous pages by actually working through the model with a value of -2.00 for β.) We may begin with a trial value of -1.00 for β. With this we get:

$$\hat{N}_i = 101 \qquad \hat{N}_2 = 132 \qquad \hat{N}_3 = 217$$

We may then calculate the R^2 between these estimated $\hat{N}_i$ and the N_i observed. When we do, we get $R^2 = 0.97455$ and, if we also calculate ∇R^2, we get -0.01882 as the value of the gradient of R^2. This tells us to move the value of β in a negative direction in order to improve the R^2. Suppose, then we try a new value of -1.52 for β. With this we get:

$$\hat{N}_i = 107 \qquad \hat{N}_2 = 129 \qquad \hat{N}_3 = 214$$

If we then calculate the R^2 between this second set of estimated N_i and the N_i observed, we get $R^2 = 0.99216$ and we get a value of -0.3306 for the gradient of R^2. This tells us that the R^2 has improved, but that it can be further improved by moving β further in the negative direction.

Suppose that we now take a new value of -2.52 for β. Then we get

$$\hat{N}_1 = 113 \qquad \hat{N}_2 = 131 \qquad \hat{N}_3 = 206$$

We then calculate $R^2 = 0.99486$ and 0.01556 for ∇R_2. Again, we have an improvement in R^2, but now the gradient tells us to make a positive increment in β in order to get a better R^2.

At a value of -2.12 for β, we get

$$\hat{N}_1 = 112 \qquad \hat{N}_2 = 130 \qquad \hat{N}_3 = 208$$

We calculate $R^2 = 0.99963$ and $\nabla R^2 = 0.00597$. So, we take another positive step. At a value of -1.96 for β, we get

$$\hat{N}_1 = 111 \qquad \hat{N}_2 = 129 \qquad \hat{N}_3 = 211$$

And, we calculate $R^2 = 0.99995$ and $\nabla R^2 = -0.00234$.

Finally, a small negative step to -2.00 for β yields

$$\hat{N}_1 = 111 \qquad \hat{N}_2 = 129 \qquad \hat{N}_3 = 210$$

And, $R^2 = 1.00000$ and $\nabla R^2 = -0.00000$. Thus, we are at the optimum or maximum value of R^2, as a function of β.

There are, of course, a number of ways to make this procedure more efficient, but the principle remains the same. There are also choices to be made as to search strategy. For example we could search to maximize the value of R^2, or we could search to minimize the value of ∇R^2. We should note also that for actual data sets we do not expect to get an $R^2 = 1.00$, rather we continue the search procedure until we get a $\nabla R^2 \approx 0.0$ and assume that this is the best set of parameters which can be gotten for the given model equation and data set. In practice, the use of these search programs is no more difficult than the proper use of standard multiple regression packages.

Integrating Transportation and Land Use Forecasting

The simple residential model derived above illustrates the general structure and an appropriate calibration procedure for more sophisticated operational models of the same form. One such model is DRAM, mentioned at the beginning of this paper. Developed from the general principles described above and calibrated as shown, DRAM has yielded consistent sets of parameters and reasonably good data fits for five different U.S. metropolitan areas in the 1960's or 1970's plus a most interesting (and still consistent) set of results for nineteenth century Philadelphia.[1] It is this model which is now included in the Integrated Transportation and Land Use Package - ITLUP.

The direct connection of the land use and transportation models should now be easily seen. Any Lowry type residential model allocates employees to place of residence. In so doing it generates an implicit matrix of work trips. If such a model is not described via entropy-maximizing it is nevertheless possible to extract this implicit work trip matrix, apply conversion factors to put it in terms of vehicle trips, and use these trips to load the network. (We note that shopping trips may be derived in a similar fashion.) This is what was done in the earlier versions of ITLUP where a modified PLUM model was used for residential locations. If an entropy derived model is used for residential location, then a similar trip matrix is an explicit part of the model formulation and operation (see equation (8), for example). In such models this trip matrix, again modified by conversion factors from person-trips to vehicle trips, may be taken directly from the model run and loaded on (assigned to) the trans-

---

[1] Putman, S. H. (1977) "Calibrating a Residential Location Model for Nineteenth Century Philadelphia" Environment and Planning, forthcoming.

portation network.

The simplicity of these statements is, in a sense, anticlimactic. This is the way to link land use and transportation models. Once the structure of these land use models is understood, the linage to transportation is both obvious and inescapable. While the addition of land use models along with the use of capacity restrained networks and feedback from congested networks, may seem to be an awesome increase in model complexity, our working experience with the model package suggests otherwise. We anticipate the eventual inclusion of similar land use models in standard transportation planning program packages such as the UTPS. Once the normal feelings of uncertainty about new techniques are overcome, their use will be seen to be quite straightforward.

The ability of such integrated model packages to properly represent important metropolitan level phenomena normally overlooked by separate land use and transportation models suggests they will play an important role in metropolitan regional planning in the coming decade.

LABORATORY TESTING OF PREDICTIVE LAND-USE MODELS:

Some Comparisons

Report of Results From
NSF-GI-38978

STEPHEN H. PUTMAN
Associate Professor of City and Regional Planning
University of Pennsylvania
Philadelphia, Pa. 19174

U.S. DEPARTMENT OF TRANSPORTATION
Office of Transportation Systems Analysis and Information
October 1976

Preface and Acknowledgement

During the past twenty years millions of dollars have been spent on design, development, and application of computer simulation models to urban problems. Despite the fact that at least one epitaph for urban simulation has been delivered, work continues to be done on this topic, which represents perhaps the only hope for comprehensive analysis capability for urban systems.

Due to the circumstances under which these models were developed, for the most part by private consulting firms, it has never been possible to make any true comparison of the many models which were attempted. Beginning in the late 1960's Professor Britton Harris and Professor Stephen H. Putman have collected, at the University of Pennsylvania, more than two dozen of these simulation models and their data sets. Due in part to this unique resource the National Science Foundation initiated the grant (NSF-GI-38978) which resulted in this report. The results reported herein are based on tests, utilizing actual urban data bases, of representatives of the two most frequently used forms of urban simulation model.

During the first half of the work under this grant the effort was directed jointly by Professors Harris and Putman. The final portion of the project and the preparation of this report were largely under Professor Putman's direction.

An external advisory board to the project provided valuable advice and especially helpful reviews of the draft of this report. The members of this board were:

Daniel Brand
William Garrison
David Gendell
Robert Goldman
William Goldner
Walter Hansen
John Hodge
Gregory Ingram
Philip Patterson
Wilbur Steger

Thanks are owed also to Professors David E. Boyce and Tony E. Smith of the Department of Regional Science, who provided useful comments and advice.

Many graduate students were involved with this project at one time or another. The following were responsible for significant components of the work:

Frederick Ducca
Charles Schramek
Ahmet Acar
Erhan Gencer
Rob Mathie
Charles Sawyer

Programming tasks were accomplished by:

Peter Kuner
Josefa Romanach
David Sanford

Secretarial work was ably provided by:

Pamela Riley
Margarita Gesualdo
Viriyane Norodom

Last, but far from least, we acknowledge the patient and helpful guidance of our NSF project monitors. The project was initiated with the help of Dr. Robert Goldman. After Dr. Goldman left NSF, Dr. David Seidman provided able and understanding guidance to the completion of the work.

1. INTRODUCTION AND SUMMARY

2. THE MODELS AND CALIBRATIONS

3. TESTS OF MODELS

Appendix I: Calibrating a Disaggregated Residential Allocation Model - DRAM

Appendix II: Tables of Model Outputs

CHAPTER ONE: INTRODUCTION AND SUMMARY

Introduction

The purposes of the research described here were to first review the status, and second develop a framework for comparison and testing of urban simulation models.

The development history of these models in the past twenty years has been cyclical. The initial model attempts began in the early 1960's and built to a crest of activity over a five year period. There followed, in the late 1960's and early 1970's a decline in modelling efforts in the wake of the many disappointments of the early work. Since the early 1970's there seems to have been a slow resurgence of model use and development.

Real progress in urban modelling can only be accomplished by a continuing process of model hypothesis and development, and subsequent model application and evaluation. Concentration on any one aspect of this process to the exclusion of the others probably causes no real harm, but may be an inefficient use of resources. Theory construction and statistical inference can never wholly substitute for empirical research. When this is attempted one is soon confronted by increasingly complex theoretical structures which simply cannot be supported by existing empirical foundations. Thus periodical evaluations and winnowing of previous results is a necessary part of the model development process.

Review and Selection of Models

There have been a number of good review articles published in past years.[1]

---

[1] Batty, M. (1972) "Recent Development in Land Use Modelling: A Review of British Research", Urban Studies, Vol. 9, No. 2, pp. 151-177.

Brown, H. J. et. al. (1972) Empirical Models of Urban Land Use New York: Columbia University Press.

Some further review work was done for this project and has been published elsewhere.[1] At the conclusion of the review it was decided that only the principal location algorithm(s) of the models would be tested. Thus more and less elaborate post-processing procedures, often found attached to these models, would be removed, allowing explicit evaluation of the model's basic construct.

The models reviewed, virtually all those for which any published descriptions were available, were classified into four broad groups.

A. The Lowry derivative models - this is now a large group of models[2] based on a straightforward set of relationships between place-of-work, place-of-residence, and in some cases, shopping-place. Most of these models deal with both residential location and non-basic or population-serving types of employment. All these models require an exogenously provided set of basic employment location estimates.

B. The EMPIRIC models - this is a somewhat smaller group of many applications of the same model.[3] The model is a set of linear difference equations with no explicit theoretical structure. The model applications involve statistical analyses of an urban data base, with the specific variables used in each application being determined as a consequence of their results. The models include both residential location and the location of all types of employment.

---

[1] Putman, S.H. (1975) "Urban Land Use and Transportation Models: A State-of-the-Art Summary" <u>Transportation Research</u>, Vol. 9, No. 2, pp. 187-202.

[2] Goldner, W. (1971) "The Lowry Model Heritage" <u>Journal of the American Institute of Planners</u>, Vol. XXXVII, No. 2

[3] Hill, D. M (1965) "A Growth Allocation Model for the Boston Region" <u>Journal of the American Institute of Planners</u>, Vol. 1, pp. 278-287.

C. The research models - a small assortment of models with potential for application at some future time, but currently in the development or pilot application stage. Examples of these are the revised Herbert-Stevens model, the National Bureau of Economic Research (NBER) model, and the Birch model.[1]

D. A miscellany of other models - a group of models proposed but not implemented, implemented but not successfully, implemented but too complex or tailor-made to a particular circumstance to allow application elsewhere, and others simply not worth pursuing further.

Having grouped the models this way, it is quite clear that models from the last two groups were not appropriate for further investigation. The most useful comparison of models then appeared to be a comparison between an application of EMPIRIC and an application of a Lowry derivative model.

The EMPIRIC model has been applied in a dozen or more major cities of the U.S. Any one of these applications would have been suitable for our comparison purposes.

There have been almost as many applications of Lowry derivative models in major U.S. cities. Of these, the most frequently applied model has been the Projective Land Use Model (PLUM) in any of its several versions. Consequently the initial intent of this project was to compare a version of PLUM to one of the EMPIRIC applications.

---

[1] Wheaton, W. Jr. (1974) "Linear Programming and Locational Equilibrium" Journal of Urban Economics, Vol. 1, pp. 278-287.

Ingram, G. et. al. (1972) The NBER Urban Simulation Model New York: Columbia University Press.

Birch, D. et. al. (1973) "The New Haven Laboratory: A Testbed for Planning" Report to the U.S. Dept. of Housing and Urban Development.

Comparison of Models: Parameter Estimation

The first comparison of the two models was to be with respect to the difficulties and relative success in estimating their parameters. This aspect of the project led to one of its major research findings. After careful review of all Lowry derivative applications in the U.S. it was discovered that in all but one case[1] the model parameters had not been properly estimated. A careful investigation was subsequently made into British modelling practice, where this problem had been identified and largely resolved. This led to a reformulation of the version of PLUM originally scheduled to be used in the project, and the subsequent development of a new form of the model, called Disaggregated Residential Allocation Model (DRAM). DRAM was used throughout the remainder of the project; it is further described in Chapter 2 and Appendix 1 of this report.

The estimation of parameters for EMPIRIC was more straightforward and was accomplished using the same procedures as had been used in its various applications. EMPIRIC's parameters were reestimated for Boston, Minneapolis-St. Paul, and Washington, D.C. DRAM's parameters were estimated for San Francisco and Minneapolis-St. Paul. Work was also done to secure other data sets for further parameter estimations in future project efforts. In most cases EMPIRIC yielded a slightly better fit to the base year data than did DRAM. This was accomplished by use of a much more extensive set of independent variables, but in the absence of any behavioral structure to the model. DRAM, with slightly lower base year data fits, but with much reduced data input requirements, is likely to produce better long term forecasts than EMPIRIC. More detailed results of these parameter estimations are presented in Chapter 2 of this report.

---

[1] The Voorhees Urban Systems Model (USM) application in the Dallas-Fort Worth, Texas region.

Comparison of Models: Sensitivity Tests

The parameters of both EMPIRIC and DRAM (plus an associated employment estimating model, EMPMOD) were thus estimated for the Minneapolis-St. Paul data set. Following this, the models' responses to both arbitrary changes in inputs as well as to simulated policy inputs were tested. Many of these tests were performed to test the models' responses to varying circumstances. Very important differences showed up in the models' performances in these tests. EMPIRIC showed no population response to changes in base year population or employment, and employment response only to base year employment changes. EMPIRIC showed some response to zone specific accessibility changes, but no response to regionwide changes. DRAM, in all these cases showed what appeared to be proper responses to all these changes in inputs. These results are presented in more detail in Chapter 3 of this report.

Conclusions

The principal conclusions of the research effort are enumerated here. All are described in more detail in the following chapters.

1. Both models require substantial data preparation prior to their use.
2. The parameters of either model can be adjusted to yield rather close statistical fits to observed data.
3. Based on these fits, both models appear to be capable of making forecasts of urban form in the absence of attempted policy manipulations. EMPIRIC may have a slight advantage over DRAM in this respect.
4. DRAM is clearly superior with respect to its response to changes in inputs. This suggests a clar advantage over EMPIRIC whenever policy tests are contemplated.
5. Testing different models on common data sets appears to be a powerful means for comparing models.

CHAPTER TWO: THE MODELS AND CALIBRATIONS

The Models Briefly Described

A brief description of the models used in the project is presented here, prior to discussing their calibration.

The EMPIRIC model is basically a set of linear equations. The variables are all expressed in terms of regional shares. For example, taking population by zone;

$$p_i = P_i / \sum_i P_i$$

where

p_i = the share of the region's population found in zone i

P_i = actual population of zone i

The dependent variables in EMPIRIC are expressed as changes in a zone's share of the variable from time t to time t+1. For example, again taking population.

$$\Delta p_i = p_{i,t+1} - p_{i,t}$$

Finally, the EMPIRIC equation structure is simultaneous, each dependent variable being a function of other dependent variables plus several predetermined variables (either lagged variables or exogenously determined variables). More specific discussion of the actual variables used will be found in the next section of this chapter.

These definitions comprise the full extent to which EMPIRIC has any structure. The variables to be included in each equation are not prespecified. Their selection is generated as the output of statistical analyses in each model application. It is precisely this lack of a substantive theoretical form which justifies the econometricians' contention that the model is not properly specified, and the urban modellers' contention that the model is non-behavioral.

The DRAM model is a sophisticated variation on the basic Lowry model theme. The hypotheses of the Lowry model assert that, given a spatial distribution of employment, and a description of zone-to-zone travel times (or costs) it is possible to estimate the location of the employees' residences. This location is taken to be a result of trip length probabilities, and in the more complex variants of the model, of residence zone characteristics. This may be written in equation form as,

$$P_i = \sum_j E_j \, p_{ij} \, A_i$$

where

P_i = population of zone i

E_j = employment in zone j

p_{ij} = the probability of a work trip as long as the travel time (or cost) between zones i and j

A_i = a measure of the residential attractiveness of zone i

The important differences between the Lowry variants result from different functional forms to generate p_{ij} from travel times (or costs) and different ways of measuring A_i. Further differences stem from structuring the model in static or dynamic form.

In DRAM,

$$p_{ij} = f\{D_{ij}^{\alpha} \ \exp \ \beta D_{ij}\}$$

where

D_{ij} = travel time between zone i and zone j

α, β = empirically derived parameters

and

$$A_i = f\{X_{1i}^{\delta 1}, \ X_{2i}^{\delta 2}, \ \ldots, X_{ni}^{\delta n}\}$$

where

$X_{ni}^{\delta n}$ = various measures of zonal characteristics including population composition and level of development

δn = empirically derived parameters

The process of finding numerical values for the parameters α, β, and δn is described later in this chapter and involves both equation fitting and, due to the explicit structure of the model, hypothesis testing.

All the Lowry variants require some external source of basic employment estimates. In past practice these sources have been quite varied, ranging from hand prepared estimates to rather complex models. In order to skip these complications in the present work, a straightforward multiple regression model, EMPMOD, was assembled for the purpose. The development of EMPMOD was incidental to the prime focus of this work. While the estimates produced are reasonably good (and are discussed later in this chapter) EMPMOD should be considered a means to the project's end.

The remainder of this chapter is devoted to a discussion of the models' parameter estimations and their implications.

Calibrating the Models

In the development of models of urban and regional systems the analyst is irrevocably trapped in the problems of drawing inferences from non-experimental statistics. It is not possible, say, to have two San Francisco Bay Areas on which to run controlled experiments. A direct consequence of this vexing situation is that we can never prove the ultimate correctness of any given model formulation as opposed to any other. We may eliminate many possible formulations on the grounds that they conflict with accepted theory and/or empirical results. Once we have narrowed down to a few likely candidates for further testing there really is no way to prove that one is better than another. We are, however, willing to assert that a model which achieves good fits to data is more worthy of further investigation than one which does not fit as well.

Considerable progress has been made in developing methods for finding "best-fit" parameters of urban simulation models. The process of finding a set of numerical parameters for a specific equation (or set of equations) which produce the best fit of those equations to a given data set has come to be called calibration of the equation(s) or model(s). In a particular example of a given data set and a given set of equations, any procedure for adjusting parameters to fit the equation to the data may properly be called a calibration procedure. The important questions here, given the data and equation(s), are first whether the procedure is computationally efficient and consistent with the theoretical structure (equations) of the model, and second, how to measure the goodness-of-fit of the model to the data.

It is not always possible to determine the best calibration procedure. It is often possible to eliminate some procedures as being clearly less adequate than others. One example of an inadequate procedure would be the practice of fitting parameters to one part of a model without taking into account their interactions with other parameters in the model. Another such example would be the practice of arbitrarily assigning parameter values without testing the consequences of such values. Neither of these procedures could produce a "proper" calibration of a model.

It is likewise not possible to specify a best fit criterion which is applicable under all circumstances. The coefficient of determination R^2, is often used as a measure of goodness of fit. Yet, this measure is, strictly speaking, inappropriate for the nonlinear models often encountered in urban simulation. Other more appropriate criteria, such as maximum likelihood estimates, are so little known to model practicioners as to be viewed with some trepidation. Criteria such as root mean square error, or standard error of estimate, do not provide a convenient basis for comparing one model to another in the absence of identical data sets.

The procedures and criteria used in this study are described along with the discussion of calibration results which follows.

Calibration of EMPIRIC

The EMPIRIC model was first described more than a decade ago, and has since seen application in more than a dozen U.S. cities. Peat, Marwick, Mitchell & Company (hereafter referred to as PMM) have been the principal proponents and practicioners of EMPIRIC. In past years they have generously supplied reports and data from these applications to the Principal Investigator of this study. Consequently there were detailed descriptions of previously estimated EMPIRIC models available for this study. These reports were available for the Atlanta, Boston, Denver, Puget Sound, Twin-Cities (Minneapolis-St. Paul), and Washington, D.C. metropolitan areas. In addition there were packages of computer programs and data sets available for Boston, Twin-Cities, and Washington. An idea of the sizes of these metropolitan areas as modelled may be obtained by reference to Table 1.

Reviewing each of these applications led to the conclusion that while many of the variables used were similar from one application to the next, (the equation structure was, of course, identical), the specific variables used were different in each application. The dependent variables were always expressed in terms of change in regional share. Population was always divided into four groups, by income, approximating quartiles. These groups are referred to as Lower Income, Lower Middle Income, Upper (or Higher) Middle Income, and Upper (or High) Income.

The five EMPIRIC equation sets were then examined for evidence of consistencie or inconsistencies from one model application to the next. In each application there were, typically, four or five population sectors and five or six employment sectors being forecasted. The precise sectoral definitions differ from one application to the next, but are generally similar.

As above, the population is usually defined as household income quartiles or groups approximating quartiles, while employment usually consists of a few basic and a few non-basic sectors. For each sector, the dependent variables are change in the zone's share of the region's total amount of the particular activity. The independent variables are of four types. First, there are lagged, or base year, values of the dependent variables and second, there are the other dependent variables. The third type of independent variable is the accessibility and/or land use variables of which there are usually several. Finally there are the public utility variables such as sewer and water availabilities.

Name of Region	Population	Employment	Year	Counties in Study Region
Atlanta, Ga.	1.0 million	605 thousand	1961	7
	1.4 million	n.a.	1970	
Boston, Mass.	3.4 million	n.a.	1960	n.a.
Denver, Colo.	0.9 million	388 thousand	1960	5
	1.2 million	533 thousand	1970	
Puget Sound, Wash.	1.7 million	610 thousand	1970	4
Twin-Cities, Minn.	1.5 million	610 thousand	1960	7
	1.9 million	850 thousand	1970	
Washington, D.C.	2.1 million	1146 thousand	1968	7

Table 1: Comparative Sizes of EMPIRIC Application Regions

The general procedure involved in applying the EMPIRIC model involves first, the preparation of a large file of raw (i.e. corrected, but unmodified) and constructed (i.e. combinations or modifications of raw) variables. A selection is then made of variables, generally those which have worked well in prior applications, for use in the preliminary regression analyses. The completion of the model calibration is then a matter of testing alternative variables until a best fit set of equations and parameters is obtained. EMPIRIC is, in a sense, very much an opportunistic model in that the final selection of variables to be used is largely based on the results obtained in the regression analyses. Those variables which produce the best fit being the ones used in the model. The regression fits obtained by this means are generally good, with coefficients of determination ranging upwards from 0.55, many of them being in the range of 0.70 to 0.90.

The measure of goodness of fit used in the EMPIRIC applications was the multiple coefficient of determination R^2. These results are tabulated for the various studies in Table 2. Note that there are two sets of results for most regions. These represent the R^2 from calibration or fitting the model to the data set, and the R^2 from reliability tests. The reliability tests consisted of using the fitted model to forecast the second data point (e.g. 1970) from the first (e.g. 1960) and then comparing the forecast to the actual data (e.g. estimated 1970 vs. actual 1970).

In Table 3 are shown the coefficients of the population variables used in the final versions of the EMPIRIC population equations for each region. A fair degree of consistency is found here, though there are some obvious discrepancies both in sign and magnitude of these coefficients. Note that the coefficients

Name of Region	Number of Zones	Time Period	Test Type	Lower Income	Lower Middle	Upper Middle	Upper Income
Atlanta	183	1961-70	Calib.	0.558	0.792	0.812	0.770
	290	1961-70	Reliab.	0.540	0.670	0.810	0.830
Boston	104	1950-60	Reliab.	0.990	0.950	0.915	0.946
	453	1950-60	Reliab.	0.951	0.906	0.793	0.826
Denver	183	1960-70	Calib.	0.647	0.841	0.655	0.839
			Reliab.	0.938	0.890	0.702	0.694
Puget Sound	244	1961-70	Calib.	0.573	0.719	0.900	0.850
			Reliab.	0.880	0.816	0.822	0.855
Twin-Cities	108	1960-70	Calib,	0.702	0.708	0.812	0.715
			Reliab.	0.919	0.940	0.880	0.827
Washington, D.C.	110	1960-68	Calib.	0.698	0.770	0.844	0.750
			Reliab.	0.947	0.917	0.877	0.886

Table 2: Fitting and Reliability Results - R^2 for Several EMPIRIC Applications for the Four Population Classes

Table 3: POPULATION COEFFICIENTS IN EMPIRIC MODELS

Dependent Variable	Study Area	Population by Income (Independent Variable)							
		Change in Share				Base Year Share			
		Lower	Lower Middle	Upper Middel	Upper	Low	Lower Middle	Upper Middle	Upper
Change in Share Low Income Population	Atlanta	-.119	+.558	-.367					
	Denver		+.129			-.392	+.337		
	Washington		+.229		-.281	-.199		+.258	
	Twin Cities		+.40		-.39	-.42	+.36		
	Puget Sound		+.352			-.314	+.294		
	Boston		+.637	-.295		+.133		-.109	
Change in Share Low-Middle Income Population	Atlanta	+.512		+.480					
	Denver	+.201		+.307			-.353	-.334	
	Washington	+.194		+.781			-.279	-.182	
	Twin Cities	+.28		+.45				+.10	
	Puget Sound			+.531			-.054*		
	Boston	+.53		+.337			-.101		
Change in Share Upper-Middle Income Population	Atlanta		+.439		+.338				
	Denver		+.612		+.25			-.27	
	Washington		+.658		+.399				-.155
	Twin Cities	-.14	+.45		+.26	-.16			
	Puget Sound		+.434		+.43			-.219	+.113
	Boston	-.125	+.637		+.294			-.224	
Change in Share Upper Income Population	Atlanta		+.512						-.447
	Denver		+.685						-.481
	Washington	-.507		+.504					
	Twin Cities	-.42		+.83					
	Puget Sound			+.657				+.219	-.437
	Boston		-.282	+.603					-.278

*Base Year Share
Total Household

shown are those which were statistically significant, as those which were not significant are not published in the PMM reports.

An interesting pattern shows in Table 3. For each population class, the change in share of a region's total population found in each zone, moves with the change in share of the adjacent population class, viz; Lower Income moves with change in share of Lower Middle Income, Lower Middle Income moves with change in shares of Lower Income and Upper Middle Income, and so on. Further, for each population class, change in share moves in opposition to (i.e. the signs of the coefficients are negative) its own concentration in the base year and moves with (though the pattern is weaker) concentrations of the next higher income group. Stated in other words, changes in share by zone of each income group move 1) with changes in shares of the next higher and next lower income income group, and 2) away from concentrations of their own income group towards concentrations of the next higher income group.

The patterns found in these coefficients of the population variables are quite consistent with hypotheses regarding peoples desires for increased socio-economic status, as well as with hypotheses regarding peoples unwillingness to live among groups very different from themselves. The patterns of coefficients of other variables in the population equations as well as those of the variables in the employment equations do not exhibit a similar degree of uniformity, and consequently are not tabulated here, though the specific case of the Twin-Cities application is discussed in more detail below.

In the other portions of these EMPIRIC model equations the sense and sensibility of the variables used, and their coefficients is another matter. There are a number of instances of contraintuitive coefficient signs and many constructed variables whose real meaning is somewhat obscure. An harsh critic could assert that the equations derived all their correlations from the unavoidable implicit correlations between activities in urban areas. Thus from the causal point of view the model results could be called fortuitous and/or spurious. A more reasonable position would be that the equation sets depend, to a significant degree, upon these strictly associative relationships, but that they will probably produce reasonably good near term forecasts, taken all together. Another view of these equations is that they are the reduced form of structural equations (in the econometric sense) which are unknown. If this view is correct, as it well may be, the use of these equations for forecasting requires that both the structure and the parameters of the unknown structural equations remain constant over the forecast period. Problems arise, as will be discussed later in this report, when policy tests with this model are attempted by means of changing specific variables. In the absence of a known, or even of an assumed structural form, it is likely that changing variables in the reduced form equations will produce peculiar results. That this concern is justified will be amply demonstrated in the discussion of sensitivity tests of EMPIRIC in a later chapter of this report.

As part of this project the three EMPIRIC applications for which data were available were all run several times, to the end of becoming more familiar with their operation. Of these three, Boston, Washington, D.C., and the Twin-Cities, recalibration runs were made for the Boston and Twin-Cities data sets. For the Twin-Cities data set the equations presented in the PMM final report were rerun using both ordinary least squares (OLS) regression and two stage least squares (TSLS), regression.[1] The differences between the OLS and TSLS

---

[1] Peat, Marwick, Mitchell and Co. (1971). "Calibration and Application of an EMPIRIC Activities Allocation Model for the Twin-Cities Metropolitan Area", prepared for the Metropolitan Council, St. Paul, Minneasota.

calibration reruns were minor, as were all but one of the differences between the PMM calibration and these calibration reruns. The reason for the one larger difference is neither known nor important in the context of this project. The differences in coefficients were also minor in all cases The variable definitions for this EMPIRIC application are shown in Table 4. The statistically significant coefficients of the equations for the TSLS calibration rerun are given in Table 5.

The great number of constructed variables used in the EMPIRIC equations make it rather difficult to interpret the results of the parameter estimations. There are few consistencies to be found in this parameter set. There are many peculiarities to be mused over. Why is change in a zone's share of population in the low income quartile positively related to change in local government and educational employment and negatively related to change in the product of highway accessibility to employment and used land area? Why is change in a zone's share of population in the upper middle income quartile not related to any employment or access variable? Why is change in a zone's share of population in the high income quartile positively related to the base year industrial employment as proportion of total employment in the zone; and not related to any other employment or access measure? More generally why aren't the EMPIRIC variables described as relative values rather than shares, thus avoiding the need to interpret what a zone's share of the percentage of something in the zone implies?

In the absence of an explicit theory or an attempt at structural equations, there can be few expectations regarding signs and magnitudes of coefficients. Consequently there is little point in discussing the EMPIRIC calibration results at length. Suffice it to say, the parameters of EMPIRIC model can be calibrated to yield relatively close fits to the data. The only consistency in the parameters

Table 4: VARIABLES DEFINITIONS - TWIN CITIES EMPIRIC

Note: Shares means regional share of variable X to be found in zone

Δ indicates "change-in-share" variables; all others are base year shares.
LIQ = Households in lowest income quartile
LMIQ = Households in lower-middle income quartile
UMIQ = Households in upper-middle income quartile
HIQ = Households in highest income quartile
MISC = Construction and other miscellaneous employment
MFGW = Manufacturing and wholesale employment
TCU = Transportation, communications, utilities employment
RET = Retail employment
SVCFIR = Service, finance, insurance, real estate employment
LGOVED = Local government and education employment
HAHU = Highway accessibility to households
TAHU = Transit accessibility to households
AHU = Composite (sum of highway and transit) accessibility to households
HAEMP = Highway accessibility to employment
AEMP = Composite accessibility to employment
SEWER = Percent of district "sewered"
NCA = Net commercial area
NIA = Net industrial area
NPA = Net public and semi-public area
USEDAC = Used area = NCA + NIA + NPA + net residential area
VACAC = Vacant or agricultural area
DEVAC = Developable area = USEDAC + VACAC
TOTAC = Total area of the district
TOTHU = Total housing units
TOTEMP = Total employment
NRA = Net residential area

Table 5: EMPIRIC EQUATIONS FOR TWIN CITIES - U. OF P. TWO STAGE LEAST SQUARES ESTIMATES

Note: Variables are 1960 share or (for Δ variables) change in share 1960-1970. R^2 for these equations are given in Table 6

$$\begin{aligned}\Delta\mathrm{LIQ} = {} & 0.407\Delta\mathrm{LMIQ} - 0.377\Delta\mathrm{HIQ} + 0.106\Delta\mathrm{LGOVED} - 0.415\mathrm{LIQ} + 0.357\mathrm{LMIQ} - 0.890\Delta(\mathrm{HAEMP} \ast \mathrm{USEDAC}) \\ & + 0.269\Delta\mathrm{SEWER} + 0.060(\mathrm{SEWER} \ast \mathrm{VACAC}) + 0.112\ (\mathrm{TOTEMP}/\mathrm{TOTHU})\end{aligned}$$

$$\Delta\mathrm{LMIQ} = 0.299\Delta\mathrm{LIQ} + 0.425\Delta\mathrm{UMIQ} + 0.092\mathrm{UMIQ} - 0.109(\mathrm{AEMP} \ast \mathrm{USEDAC}) + 0.300\Delta(\mathrm{HAEMP} \ast \mathrm{USEDAC})$$

$$\Delta\mathrm{UMIQ} = -0.144\Delta\mathrm{LIQ} + 0.415\Delta\mathrm{LMIQ} + 0.261\Delta\mathrm{HIQ} - 0.163\mathrm{LIQ} + 0.058\ (\mathrm{SEWER} \ast \mathrm{TOTAC}) + 0.104\ (\mathrm{UMIQ}/\mathrm{TOTHU})$$

$$\Delta\mathrm{HIQ} = -0.416\Delta\mathrm{LIQ} + 0.0\Delta\mathrm{LMIQ} + .830\Delta\mathrm{UMIQ} + .248\Delta\mathrm{SEWER} - .260(\mathrm{HIQ}/\mathrm{TOTHU}) + .274(\mathrm{INDUS}/\mathrm{TOTEMP})$$

$$\begin{aligned}\Delta\mathrm{MISC} = {} & .44\Delta\mathrm{RET} + .20\Delta\mathrm{SVCFIR} - .026(\mathrm{TOTEMP}/\mathrm{TOTHU}) + .112(\Delta\mathrm{SEWER} \ast \mathrm{TOTAC}) - .256\mathrm{MISC} \\ & -.096\mathrm{SVCFIR} + .109(\mathrm{NIA} \ast \mathrm{VACAC}/(\mathrm{USEDAC} + \mathrm{VACAC}) + .094\mathrm{TAHU})\end{aligned}$$

$$\begin{aligned}\Delta\mathrm{MFGW} = {} & .013\Delta\mathrm{SVCFIR} + .190(\mathrm{SEWER} \ast \mathrm{TOTAC}) + .254\mathrm{SVCFIR} - .189\mathrm{MFGW} - .268\mathrm{NCA} \\ & - .531(\mathrm{USEDAC}/(\mathrm{USEDAC} + \mathrm{VACAC})) - .248\Delta\mathrm{HAHU} \ast \mathrm{USEDAC} + .52\mathrm{HAHU}\end{aligned}$$

$$\begin{aligned}\Delta\mathrm{TCU} = {} & .737\Delta\mathrm{RET} + .919\Delta\mathrm{SEWER} + .249\mathrm{NIA} \ast \mathrm{VACAC}/(\mathrm{USEDAC} + \mathrm{VACAC}) - .352\mathrm{MFGW} \\ & + .60\mathrm{USEDAC}/(\mathrm{USEDAC} + \mathrm{VACAC}) + .1827\mathrm{CU} - .53(\mathrm{TOTEMP})/(\mathrm{NIA} + \mathrm{NCA} + \mathrm{NPA}) + .31(\mathrm{TOTEMP}/\mathrm{TOTAC}) \\ & - .423(\mathrm{NCA} \ast \mathrm{VACAC}/(\mathrm{USEDAC} + \mathrm{VACAC}))\end{aligned}$$

$$\begin{aligned}\Delta\mathrm{RET} = {} & .473\mathrm{SVCFIR} + .518\Delta\mathrm{LMIQ} + .077\mathrm{NCA} \ast \mathrm{VACAC}/(\mathrm{USEDAC} + \mathrm{VACAC}) - .32\mathrm{RET} \\ & + .291\mathrm{AHU} \ast \mathrm{USEDAC}\end{aligned}$$

$$\Delta\mathrm{SVCFIR} = .169\Delta\mathrm{UMIQ} + .202\mathrm{MFCW} + .344\mathrm{RET} - .154\mathrm{GOVED} - .228\mathrm{SVCFIR} + .236\mathrm{RET}$$

$$\Delta\mathrm{LGOVED} = .291\Delta\mathrm{LIQ} + .313\mathrm{TAHU} + .214\mathrm{NCA} - .539\mathrm{LGOVED}$$

Table 6: COMPARISON OF CALIBRATIONS OF EMPIRIC: TWIN-CITIES DATA

Dependent Variable	1 PMM-R^2	2 UoP-TSLS-R^2	3 UoP-OLS-R^2
ΔLIQ	0.702	0.703	0.706
ΔLMIQ	0.708	0.714	0.720
ΔUMIQ	0.812	0.816	0.824
ΔHIQ	0.715	0.715	0.724
ΔMISC EMP	0.750	0.746	0.761
ΔMFG	0.718	0.708	0.714
ΔTRANSP	0.504	0.464	0.464
ΔRET	0.790	0.790	0.793
ΔSERV+FIRE	0.755	0.754	0.758
ΔLOGOV+ED	0.545	0.545	0.546

Column 1 - Resulting R^2 from PMM calibrations

Column 2 - Resulting R^2 from this project's recalibration using Two Stage Least Squares regression.

Column 3 - Resulting R^2 from this project's recalibration using Ordinary Least Squares regression.

Identical dependent and independent variables were used in all three calibrations.

from one application to the next appears in the population group-to-population group relationships. The parameters for other variables and other equations are catch as catch can, and raise questions as to the simultaneity alluded to in the general descriptions of the model which accompany each application. Overall, attempts to use these models for any but short term, no policy, forecasts should be viewed with considerable skepticism.

Calibration of DRAM

During the initial stages of this project the decision was taken to compare the EMPIRIC model to a package containing a version of IPLUM and a simple basic employment model. There was no intention of performing any model development work for this project. This was all well and good as the project proceeded on through its early stages. It was when work began on the calibration of IMPLUM, that trouble became apparent. In fact, the entire history of application of Lowry and "Lowry derivative" type models in the U.S. is fraught with tales of calibration difficulties. Further investigation yielded the unpleasantly interesting fact that in U.S. practice, with but one exception, no Lowry type model had ever been successfully calibrated (in a statistical sense). Partial calibrations, in some cases of the p_{ij} function, in others of a multivariate measure of A_i, had been accomplished, but no complete estimations of a model's parameters had been done. There had, however, been a number of successful calibrations of Lowry type models in British practice. Consequently an effort was undertaken to determine whether IPLUM could be calibrated by the procedures used in the British work.

The British calibrations draw upon a reformulation of Lowry type models according to the Wilson maximum-entropy approach.[1] This approach was also

---

[1] Wilson, A. (1970). Entropy in Urban and Regional Modelling, Pion Ltd., London.

used in the one U.S. exception mentioned above, the Voorhees U.S.M.[1] When so reformulated, the path to calibration of Lowry type models becomes quite clear, though it does require use of mathematical search techniques for non-linear equations, rather than the better known multiple regression techniques used in EMPIRIC. Consequently it became necessary to recast IPLUM in the entropy maximizing form. As a part of this effort several very desirable improvements to the model became not only possible, but, in a sense, inevitable. In particular, the population sector of the model, formerly considered as one homogenous group, was disaggregated into four sectors defined in terms of income. Further the need for many of the arbitrary correction factors used in the later portions of the IPLUM model was eliminated. This new formulation of the model eventually became sufficiently different from its progenitor to warrant a new name - Disaggregated Residential Allocation Model (DRAM).

The mathematical development of DRAM and its calibration requirements are described in Appendix I to this report. The resulting equation for household location is as follows:

$$N^k_{it} = \sum_j E^k_{jt} \left[\frac{W^k_{it}\, f(C_{ijt})}{\sum_i W^k_{it}\, f(C_{ijt})} \right]$$

where

N^k_{it} = number of type k households located in zone i at time t

E^k_{jt} = number of type k employees working in zone j at time t

C_{ijt} = travel time between zone i and zone j at time t

W^k_{it} = multi-attribute measure of the attractiveness of zone i at time t to households of type k

---

[1] A. M. Voorhees and Associates (1972). "Application of the Urban Systems Model (U.S.M.) to a Region-North Central Texas", prepared for North Central Texas Council of Government, Dallas - Fort Worth, Texas.

The definition of the attractiveness measure is of crucial importance, and so will be discussed here. The zonal attractiveness measure consists of two principal parts. One part is the actual amount of land in the zone which is available for development, perhaps adjusted by its developability or existing level of development. The second part is the desirability, of the land in the zone as viewed by potential locators, apart from that solely due to its spatial location.

Many measures of the intrinsic attractiveness of residential zones have been proposed. These have included property value, quality of school systems, housing mix, degree of land use mix (e.g. other uses beside residential), and population mix. There was evidence in prior work by this author as well as by others that household incomes would be a good overall (or perhaps surrogate) measure of zonal attractiveness. Thus it was decided that the percentage composition of household types (in income quartiles) would serve as the attractiveness measure.

The amount of land available in the zone was measured in terms of vacant developable land. In order to adjust the attractiveness of vacant land to represent the presence or lack of infrastructure (e.g. sewers, water, electricity, etc.) vacant land is weighted by the percentage of developable land in the zone already developed. Finally, as a surrogate for the type of residential development, residential density was included.

Similarly, it was necessary to define a trip probability function. It is well known that most empirical trip distributions take the form of a normal curve considerably skewed to the left. It has been hypothesized that this distribution arises from the product of exponentially increasing numbers of opportunities (for trip satisfaction) encountered with increasing distance travelled, and exponentially decreasing propensity to travel each additional unit of distance. This is shown graphically in Figure 1. The trip probability function which results has the form

$$P_{ij} = (D_{ij}^{\alpha} \ \exp \ \beta D_{ij})$$

as described at the start of this chapter.

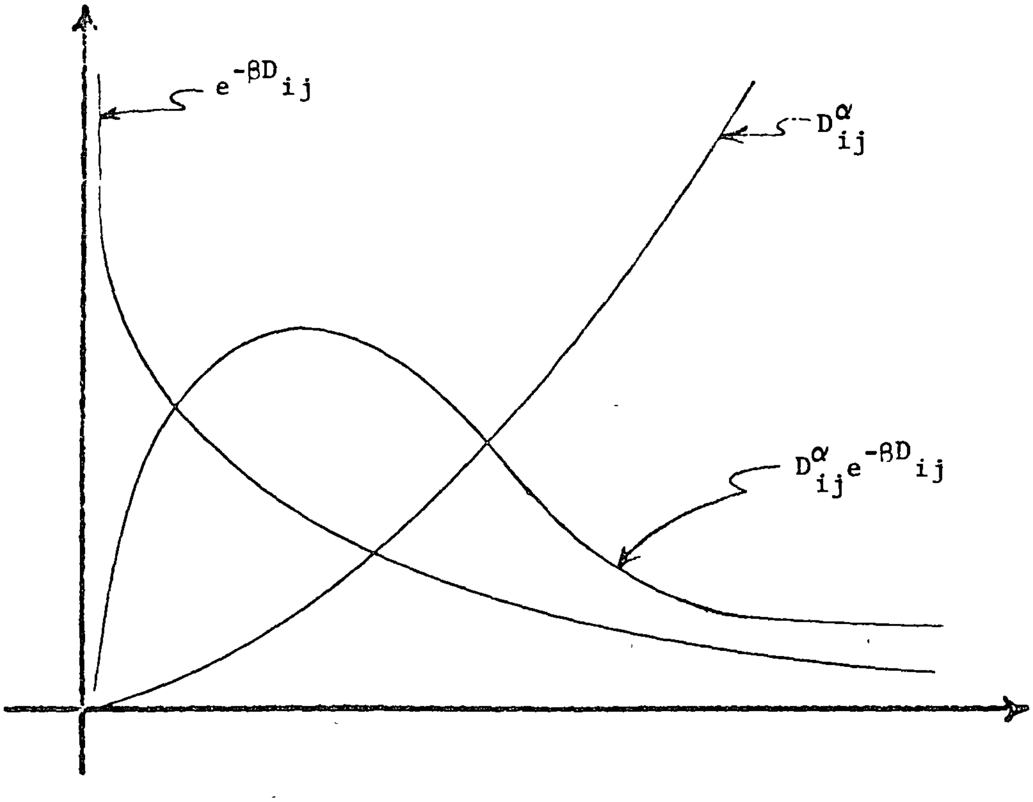


Figure 1: Trip Function Equation Form

The underlying hypothesis of the zonal attractiveness measure, developable vacant land weighted by its intrinsic attractiveness and its desirability, results in the use of a product function form. Thus the measure is written as,

$$W_i^k = (n_i^1)^{\delta 1} \ (n_i^2)^{\delta 2} \ (n_i^3)^{\delta 3} \ (n_i^4)^{\delta 4} \ R_i^{\delta 5} \ Q_i^{\delta 6} \ V_i^{\delta 7}$$

where

n_i^1 = percentage of zone i households which are in the lowest income quartile

n_i^2 = percentage of zone i households which are in the low middle income quartile

n_i^3 = percentage of zone i households which are in the upper middle income quartile

n_i^4 = percentage of zone i households which are in the upper income quartile

R_i = residential land area of zone i

Q_i = percentage of developable land in zone i which has been developed

V_i = vacant land in zone i

δn = a set of n parameters to be estimated

Thus, taking the trip probability function along with the zonal attractiveness measure, there are nine parameters to be estimated for each of the four household types.

Some model efforts have estimated parameters for a model's trip probability function, but have assigned values of 0.0 or 1.0 to any δ parameters (or their equivalents) in the attractiveness measure. Other model efforts have estimated values of the δ parameters by multiple linear regression, independent of the trip probability function. For an urban area with an existing spatial distribution of activities the parameters of these functions should not be separately estimated.

The reality which we are attempting to describe, like a solution of sugar and water, cannot be separated by mechanical means. The shape of the trip probability function is due, in part, to the spatial distribution of zonal attractiveness. Similarly, the attractiveness of a zone results in part from the households located therein. These have so located, in part, responding to the work trips which are an implicit aspect of living in that particular zone. Simultaneous estimation of the attractiveness parameters and the work trip parameters is thus required as a consequence of the inseparability of these phenomena.

A glance at the DRAM equations shown above makes it clear that standard parameter estimating procedures such as regression techniques are inadequate. At the other extreme, brute force trial and error methods may not yield useful results at reasonable cost. What is needed is a sophisticated n-dimensional search technique that doesn't make the assumptions necessary in regression, but is much more efficient than trial and error. Two candidate methods are pattern search and gradient search.

Holding, momentarily, the question of a proper criterion function, it may be assumed that one exists. In pattern search, successive explorations are made, as to the change in the criterion which results from a change in each of the parameters. Then, based on the information gleaned from these explorations a step is taken in all n-dimensions at once. In gradient search the gradient (an n-dimensional vector orthogonal to the mathematical surface, whose projection on that surface points in the direction of its steepest ascent) is evaluated at a given point and an n-dimensional step is taken in the direction indicated by the gradient projection. The gradient may be found using numerical approximations or by analytically solving the function's partial derivatives with respect to the criterion. It was not the purpose of this project to develop new calibration techniques. An efficient, operational, gradient search program was available and was used to estimate the parameters of DRAM. At some future date an investigation of alternative search procedures will be made.

Our inexperience with these techniques led to the aggregation of the San Francisco data set to 30 zones in order to lower the cost of our inevitable mistakes. A conservative approach was taken, with much attention paid to initial estimates (starting points for the search procedure) of the parameters. Experience with the technique has shown these concerns to have been unwarranted, the parameter estimates being easily done at reasonable computer expense. These preliminary efforts are further described in Appendix I.

Having once established the feasibility of the technique, the 108 zone Minneapolis-St. Paul data was approached. Again, no difficulties were encountered and the parameter estimates were readily obtained. The parameters and the associated criterion value, R^2, are shown in Table 7. No t- or F-values are provided, as these statistics, unfortunately do not apply to the non-linear equations of DRAM.

It is worth noting that the criterion function used in these estimations was the least-squares (sum of squared differences between observed and estimated data points) criterion. It was observed in the parameter estimations that the criterion surface tended to be rather flat in the neighborhood of the

Table 7: POPULATION COEFFICIENTS FOR DRAM - TWIN CITIES

	Distance		Household Composition				Land Conditions			
Variables	Opp'ty	Decay	L.I.	L.M.I.	U.M.I.	U.I.	Res.	% Dev.	Vacant	
Household Sector	α	β	$\delta 1$	$\delta 2$	$\delta 3$	$\delta 4$	$\delta 5$	$\delta 6$	$\delta 7$	R^2
Low Income 0-8000	1.04	-2.18	0.765	0.142	-0.558	-0.339	0.893	0.145	-0.031	0.927
Low Middle Income 8000-12000	2.11	-1.46	0.244	0.835	-0.373	-0.152	0.899	0.177	-0.044	0.900
Upper Middle Income 12000-16500	2.81	-1.31	0.086	0.157	0.500	-0.080	0.795	0.251	-0.079	0.900
Upper Income 16500 →	2.10	-1.44	0.131	0.099	-0.191	0.776	0.752	0.292	-0.033	0.905

best fit parameter estimates. Consequently the criterion was, in some cases, insensitive to small changes in some of the parameters. Furthermore, gradient search provides no information as to the statistical properties of the parameters found. A current investigation of the use of the maximum-likelihood criterion as an alternative to the least-squares criterion appears to be leading to resolution of both these problems.

Refering to Table 7, consider first the "Distance" or trip probability function parameters. The Low Income households show the lowest propensity to travel to work i.e. the largest negative β or "Decay" parameter. The Upper Middle Income households show the highest propensity to travel to get to work. Lower Middle Income and Upper Income households are much more willing to travel a given distance to work than is a Lower Income household, and not quite as willing as an Upper Middle Income household. These results are in accord with other empirical findings as well as with theories concerning the portion of a household's budget allocable to travel expense.

The opportunities encountered or "Opp'ty" parameter shows opposite results. As trip length increases Low Income households find the fewest opportunities for trip satisfaction. Upper Middle Income households find the most rapidly increasing numbers of opportunities with increasing trip length. In DRAM, trip satisfaction is the choosing of a residential location. Taking these results together, we find that Upper Middle Income households are likely to have the longest work trips, due not to a lack of residential opportunities but rather to a great willingness to travel to the "right place". At the same time, the Lower Income households will also have longer work trips but, in this case due to fewer residential opportunities, despite a greater unwillingness or inability to travel.

Turning next to the Household Composition variables, the first conclusion is that each income group is most likely to locate (or be located) in zones where it is already concentrated. The second conclusion is that any other household type is least likely to be located with concentrations of Upper Middle Income households. No other general conclusions may be drawn from these parameters.

Finally, turning to the Land Conditions, all household classes are positively affected by amount of residential land (Res.) and by percentage of developable land developed (% Dev.). There is a slight negative response to vacant land. The need for a statistical significance measure for these parameters is clearly needed here. Without it we cannot sort the meaning out of these last sets of parameters. As mentioned above, further work on this is now being done.

The values of R^2 achieved in the DRAM calibrations are higher than those of EMPIRIC. The numbers of zones are identical for both of these calibrations to the Minneapolis - St. Paul data. The dependent variables in the EMPIRIC calibrations are, however, "change-in-share" variables, while the dependent variables in DRAM are simply "share". One expects better data fits with share than with change-in-share variables. Thus it is difficult to compare these sets of results.

In order to better compare the models' performance, both the EMPIRIC and DRAM model packages were run form a 1960 base to a 1970 projection year. The 1970 model estimates were then compared to actual 1970 Census data. Due to data incompatibilities, only the household sections of the models were comparable. The results of this test, in terms of the correlations between the model estimates and the actual data are shown below (in terms of r^2).

Household Type	EMPIRIC	DRAM
LIQ - lower income	0.918	0.750
LMIQ - lower middle	0.941	0.828
UMIQ - upper middle	0.889	0.844
HIQ - upper income	0.829	0.699

From these, more comparable, evaluations it is clear that EMPIRIC achieves somewhat better fits to the data than does DRAM. Balanced against this is the fact of DRAM's more understandable and theoretically satisfactory equation structure, along with the empirical support derived from the signs and magnitudes of the fitted parameters.

Calibration of Employment Model for DRAM

The modelling of employment differs somewhat between the models being studied here. The EMPIRIC model incorporates six types of employment directly in its equation system and produces forecasts for each type as a matter of course. The IPLUM and DRAM models, as is the case with all Lowry derivative models, do not include a procedure for estimating "basic" employment, but require such estimates as an input. In the many applications of these models basic employment estimates have been generated in a number of ways ranging from educated guesses to rather complex models.* Consequently it was necessary to add a procedure for estimating basic employment to either DRAM or IPLUM in order to compare their performance to that of EMPIRIC. This section of the report describes the development of such a procedure.

As is described above in the section on calibration of the residential models, the initial calibration work for DRAM was undertaken with data from San Francisco, it was decided to begin work on an employment estimating procedure by using the same data sets. It was not the purpose of the project to develop new models, so the first thought was to use BEMOD, a model which had been developed with the San Francisco data.** The model used a large number of variables to describe each census tract in the region. The variables used were: slope, elevation, presence of navigable waterway,

---

* See Putman, S. H. "Intraurban Employment Forecasting Models: A Review and a Suggested New Model Construct", Journal of the American Institute of Planners, Vol. XXXVIII, No. 4, July 1972.

** Nathanson, J. "Basic Employment Model: A Model for Intra-County Location of Basic Employment and Land", BATSC Technical Report 222 (Preliminary), Bay Area Transportation Study Commission, Berkeley, California. (1970)

presence of rail facilities, a general accessibility measure, density of existing development, residential land, unused land, vacant land, and the distribution of employment types in the zone. The areal unit used was the census tract, but the parameters for the census tracts in each county or group of counties (six in the study area) were estimated separately. Ten employment sectors were used. The results of the regressions, done in May of 1968, for that model are tabulated in Table 8. Despite the many variables used, these results were not very good, particularly when compared to those obtained with the much simpler formulation used in the USM model for the Dallas - Fort Worth region.* The USM used little more than a lagged variable and an access measure to obtain much better data fits. Consequently an attempt was made to develop a similar set of simpler estimating equations for San Francisco.

The ten employment types of BEMOD were first disaggregated to twelve employment types. A number of regressions were estimated, using the two-hundred ninety-one zone areal system used in IPLUM. This areal system is an aggregation of the seven hundred seventy seven census tracts used in BEMOD, and further, was not broken into separate county regressions. The results yielded poor data fits. In an attempt to improve them, the seven manufacturing sectors were aggregated to three, resulting in a total of eight employment types. The three levels of sectoral aggregation used in these analyses are shown in Table 9. The results of these analyses, while in most cases as good or slightly better than the old BEMOD results were not satisfactory, the values of R^2 ranging from 0.35 to 0.58.

---

* Voorhees, A. M. and Associates "Application of the Urban System Model (USM) to a Region-North Central Texas", Prepared for North Central Texas Council of Government. (1972)

Table 8: BEMOD REGRESSION RESULTS (R^2), MAY 1968, TEN EMPLOYMENT TYPES, CENSUS TRACT

	MFG1	MFG2	MFG3	MFG4	MFG5	TRAN	WHOL	FIN	SERV	GOVT
San Francisco	.7098	.6376	.6344	.5543	.5636	.8725	.8725	.3921	.4710	.4969
San Mateo	.3107	.6469	.7763	.6772	.7444	.6413	.8993	.0571	.4521	.7693
Santa Clara	.2584	.1336	.0607	.0674	.0272	.0281	.1716	.1754	.3243	.2502
Alameda	.1078	.2091	.1712	.4958	.6459	.0927	.1672	.0546	.0565	.1381
Contra Costa	.0697	.1292	.0762	.0953	.2357	.0665	.2680	.1889	.5008	.2680
North Bay (Marin, Solano Napa, Sonoma)	.0938	.1681	.1622	.1128	.1082	.2361	.1457	.3522	.1058	.4412

MFG1 = Manufacturing, New Technology
MFG2 = Manufacturing, Centralized Urban
MFG3 = Manufacturing, Decentralized Urban
MFG4 = Manufacturing, Metal Fab. and Machinery
MFG5 = Manufacturing, Petrochems., Primary Metals

TRAN = Transportation
WHOL = Trade
FIN = Finance and Insurance
SERV = Services
GOVT = Government

Table 9

SECTORAL DEFINITIONS IN 4 EMPLOYMENT ANALYSES

	BEMOD* Sectors		EMPMOD 12 Sector Analysis		EMPMOD 8 Sector Analysis
1.	MFG 1	1.	Ag., For. Fish.	1.	Ag., For. Fish.
2.	MFG 2	2.	Mining	2.	Durable Heavy Mfg.
3.	MFG 3	3.	New Technology	3.	Durable light Mfg.
4.	MFG 4	4.	Centralized Urban	4.	Non-durable Mfg.
5.	MFG 5	5.	Decentralized	5.	Trade
6.	TRAN	6.	Metal & Machinery	6.	Fin & Ins.
7.	WHOL	7.	Petroleum & Prim. Met.	7.	Services
8.	FIN	8.	Transp.	8.	Gov't
9.	SERV	9.	Trade		
10.	GOVT	10.	Fin & Ins.		
		11.	Services		
		12.	Gov't		

* See Table 8 for definition of BEMOD sectors

Reference again to the USM work suggested that the most important single variable in their employment equations was lagged employment i.e. employment of the same type as that being estimated, in the same area, in the prior time period. This variable was not available for San Francisco which, incidentally, also precluded estimation of the parameters for the EMPIRIC model in that region. A double-logarithmic form of equation was tried instead of the traditional additive linear form, with a modest increase in the R^2 values. The two-hundred ninety-one zone data set was aggregated to ninety-eight zones in the hope of improving the ability to estimate employment in the area. These results were somewhat improved, showing values of R^2 ranging from 0.53 to 0.80 for the eight employment sectors. At this point it was decided that further work on employment estimates with this data base would be fruitless. This work confirmed our expectations as to what could be done in this vein with strictly cross-sectional data. The multiple regressions tested while not very good, produced results which were as good or better than the early results for the BEMOD regressions.

EMPMOD: The Minneapolis - St. Paul Estimates

The data set for Minneapolis - St. Paul contains employment data for two points in time. It is this fact which allows the calibration of the EMPIRIC model on this data and which, as will be described below, yields relatively good employment estimating equations.

The EMPIRIC model uses six types of employment, as follows:*

---

* Peat, Marwick, Mitchell & Co. "Calibration and Application of an "EMPIRIC" Activities Allocation Model for the Twin Cities Metropolitan Area", Final Report, Metropolitan Council, St. Paul, Minnesota, Dec. 1971.

1. MISC - miscellaneous (S.I.C. 01-17)
2. MFG - manufacturing and wholesale (S.I.C. 19-39,50)
3. TRANSP - transport., communic., utilities (S.I.C. 40-49)
4. RET - retail (S.I.C. 52-59)
5. SERV+FIRE - finance, ins., real est., services (S.I.C. 60-89)
6. LOGV+ED - local gov't., education (S.I.C. 82, 91-94)

Each of these types of employment was estimated in EMPIRIC with an additive linear equation. These estimated equations were shown in Table 5.

Given the EMPIRIC results, work was begun on estimation of the parameters of a set of equations for employment estimates to be used as input to DRAM. The equation form tested was additive linear, and a standard multiple regression estimation procedure was used. The variable definitions and equations are given in Tables 10 and 11. The same eight sectors defined for the San Francisco analysis were used for these Minneapolis-St. Paul parameter estimates. The areal system was the same one hundred eight zone system as was used by the EMPIRIC and DRAM analyses, and represented a slightly greater degree of areal aggregation than the two-hundred ninety-one zone system for San Francisco. The regression results, in terms of values of R^2, were as follows (for comparable sectors):

	Sector	EMPMOD	Sector	EMPIRIC
1.	Ag. Forest. & Fish	0.421	MISC	0.761
2.	Durable, Heavy Mfg.	0.758		
3.	Durable, Light Mfg.	0.811	MFG	0.714
4.	Non-durable Mfg.	0.812		
5.	Trade	0.866		
6.	Finance & Ins.	0.970	SERV + FIR	0.758
7.	Services	0.901		
8.	Gov't.	0.960	LOGOV	0.546

Table 10: DEPENDENT VARIABLES: EMPMOD

E(K) = Employment in industry group K (K=1, 8) in 1970

EMP(K) = Employment of type K, in 1960

TLA = (AVL-TL(I)**2: Variance of zonal total land

TLB = $\frac{\text{MAXTL - TL(I)}}{\text{MAXTL - MINTL}}$: Normalized " " "

BL = Industrial land (1960) + Available land (1960) in (I)

IDENS = Industrial density (1960) in zone (I)

POP = Res. population in (I) (1960)

RDENS = Res. density in (I) (1960)

DGRI = % change in # of households in income quartiles (1+2)

DGR2 = % change in # of households in income quartiles (3+4)

DACIN = Change in composite accessibility to manufacturing emp.

DACCM = Change in composite accessibility to commercial emp.

DACG 1 = Change in composite accessibility to income group 1

DACG 2 = Change in composite accessibility to income group 2

DSWR = Change in Sewer system (land)

STE = % of total emp. in (I) in (1960)

SACTE = % of composite accessibility to TE. in (I) in 1960

SACHH = % of composite accessibility to HH in (I) in 1960

Table 11: ESTIMATED EMPLOYMENT EQUATION FROM EMPMOD

E1 = -0.0181 TLA + 0.0007 POP + 1331.4 DGR2 + 18629. DAGG1 - 15748. DACG2 + 8.5920

E2 = 1.1745 E1LAG + 770532. DACG1 - 1009028. DACG2 + 17931. DSWR + 28192. SACTE - 29.163

E3 = 1.3604 E3LAG + 15.848 RDENS + 23385. DGR1 + 23652. DGR2 + 109.96

E4 = 0.9491 E4LAG + 152642. DACIND - 107994. DACG2 + 166.76

E5 = 1.0515 E5LAG + 0.5846 IDENS + 0.0160 POP + 26253. DGR1 + 45314. DGR2 - 131.62

E6 = 0.8977 E6LAG - 0.1785TLA - 5.7198TLB + 1.9488 IDENS + 0.0035 POP
+ 4.6650 RDENS + 4471.2 DGR1 + 7411.3 DGR2 + 65497. DACG1 - 72701. DACG2
- 3969.4 DSWR + 4157.8 STE + 449.19

E7 = 0.6957 E7LAG - 0.4636 TLA - 14.782 TLB - 0.6809 IDENS - 0.0089 POP + 30452. DGR1
+ 26937. DGR2 + 290845 DACIND - 422560. DACCM + 85920. DACG1 + 16910. STE
+ 33841. SACHH + 1136.9

E8 = 1.5057 E8LAG - 0.7470 IDENS - 0.0062 POP - 28.378 RDENS + 13015. 13015. STE + 77.806

Note: All coefficients significant at 5% or better

There are several points to be noted here. First, as was the case in the residential models, these results are not strictly comparable as the EMPMOD dependent variables are static, while the EMPIRIC dependent variables are change in regional share. Second we note that the EMPIRIC and EMPMOD results are not strictly comparable due to different sector definitions.

There is a further point to be mentioned regarding numbers of zones. It has sometimes been asserted that for a given urban area and a particular model one should expect increasingly good fits of equations to data with decreasing areal disaggregation. In other words, ceteris paribus, fewer zones yields better fits to data. While it was not possible to specifically test this hypothesis in the study, the impression gained from working with both models on the various data sets is that this phenomenon, if it operates at all, is rather weak in its effects. In fact, while it probably operates at the high end e.g. a difference between 200 zones or 600 zones, it probably operates in reverse at the low end e.g. a difference between 100 zones and 10 zones. Further exploration of this phenomenon is planned for future work with DRAM.

Retail Employment Estimates in DRAM

Finally, it should be noted that in order to save project time the existing, uncalibrated, local serving employment model included in IPLUM was used during the sensitivity and policy tests of the DRAM package. This submodel should probably be replaced by a better "retail" model during future of the DRAM package.

Summary

Thus we have developed the two model approaches, EMPIRIC and the DRAM-

EMPMOD combination. EMPIRIC contains four household types and six employment types. DRAM-EMPMOD contains four household types and eight employment types.

All models were fit (i.e. parameters estimated) to the same 108 zone data base for Minneapolis-St. Paul, 1960 to 1970. EMPIRIC achieved a somewhat better fit to the data than DRAM-EMPMOD.

The focus of the work was on the residential equations of EMPIRIC, and on DRAM. The employment equations in EMPIRIC, and the whole of EMPMOD were simply necessary to provide inputs to the residential location estimates.

It is worth noting that a criticism often leveled at Lowry derivative models, e.g. DRAM, is that they depend on basic employment estimates as inputs. These inputs, it is contended, can never be perfect and therefore must have an adverse effect on the residential estimates. A test run of DRAM was made with EMPMOD inputs and an alternative run was made with actual employment data inputs. The subsequent two sets of residential outputs were compared to actual residential data. The differences in the correlations between the estimates and actual population were not statistically significant. Further, it should be recognized that the outputs of the employment equations of EMPIRIC, however perfect or imperfect, are inputs to the residential equations just as the outputs of EMPMOD are input to DRAM.

Finally, though EMPIRIC achieves better fits to the data, it lacks theoretical form and its estimated parameters do not agree well with theory, intuition, and other empirical findings concerning urban spatial phenomena. DRAM's parameters do agree with these, but do not result in as good a fit to base data as was found with EMPIRIC.

CHAPTER THREE: TESTS OF MODELS

Introduction

The previous chapter's comparison of the two models ends inconclusively. The EMPIRIC model achieves somewhat better fits to the data, but leaves much to be desired in the way of theoretical underpinnings. The DRAM package has stronger theoretical underpinnings, which are supported by the empirical results, but does not achieve as good fits to the data as does EMPIRIC.

In this chapter the models are tested and evaluated in a different way. By models we mean, EMPIRIC on the one hand, including all population and employment sectors. The DRAM package, on the other hand, consists of DRAM and EMPMOD as described in the previous chapter. Each of the models, again using a common data set, is subjected to a wide variety of changes in inputs. The resulting changes in output are then compared both to each other and to expectations based on theory and existing evidence as to the behavior of urban spatial systems.

Each type of test was done by making identical input changes to both models and observing the resulting changes in output. In some cases further tests were made of one model and not the other in order to further investigate a particular question.

The tests described below fall into several types:

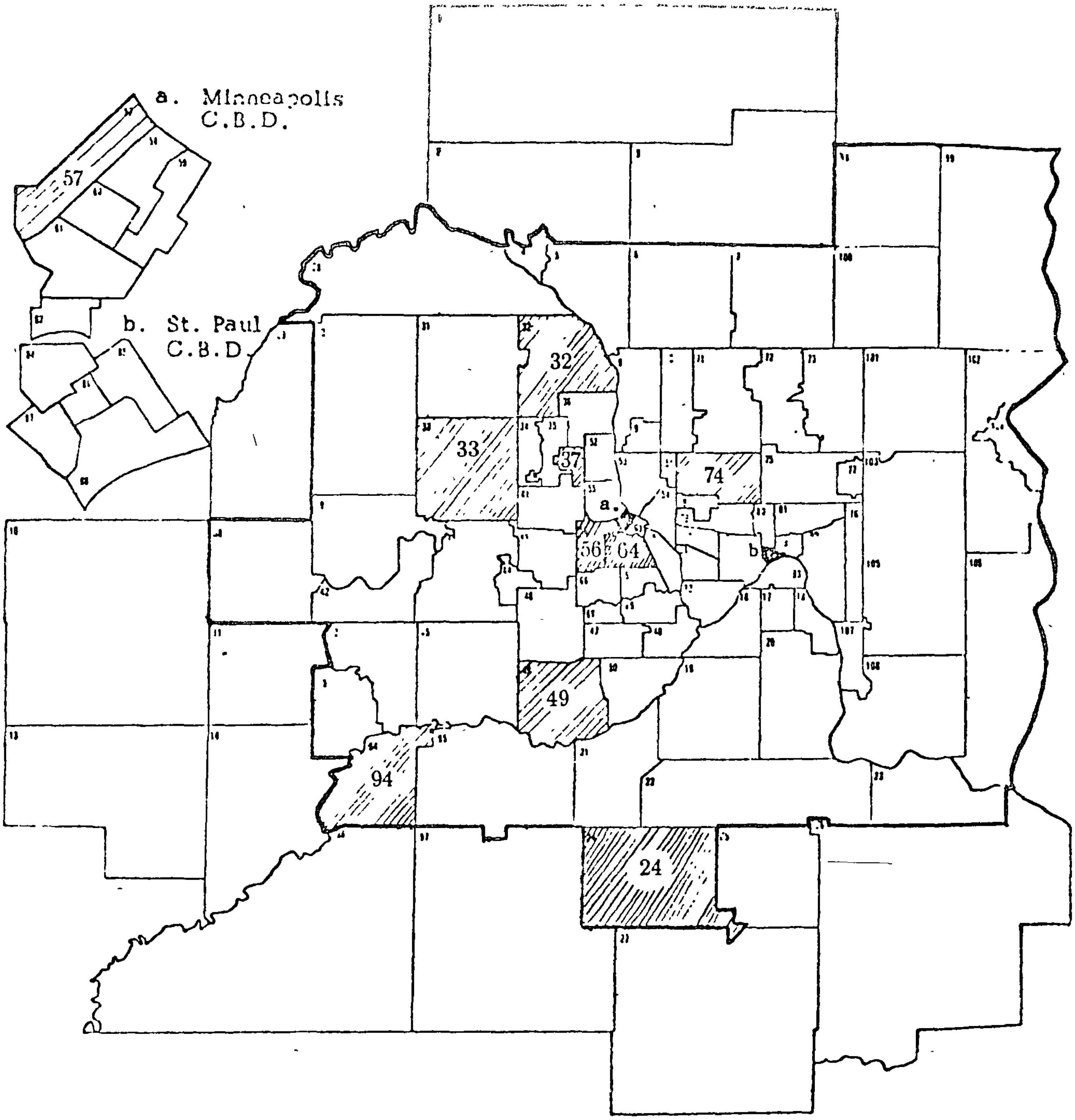
1. Changes in population
2. Changes in accessibility

3. Changes in employment
4. Changes in land use

Each of the tests involved changing input data to the models, and observing corresponding changes in outputs. All tests were made using the Minneapolis-St. Paul data base, with 1960 as the base year and 1970 as the future year. All results (i.e. test run outputs) were compared to a control run of each model (hereafter referred to as CR). The CR was also, 1960 to 1970, but no inputs were manipulated. Some tests can be construed as representing possible policy alternatives, while others are clearly just manipulations of the models. The tests described, a subset of all those done, rather clearly illustrate the model's capabilities. Many additional tests might have been run but, as the reader will find by the end of this chapter, they were unnecessary for the purposes of this project.

Changes in Population: Low Income Household Reductions

These tests were made for several different zones. One example was for Zone 57 in the Minneapolis urban core (see Map 1). In this zone a simplistic form of "Urban Renewal" was done, all low income households were removed in the base year. This amounted to 697 low income households, or 79.6% of the households in Zone 57 in the base year. The EMPIRIC base year input, being expressed in shares (as discussed in the previous chapter) was changed so that the



MAP 1: MAP OF TWIN CITIES REGION WITH ANALYSIS ZONES SHOWN

(Minneapolis and St. Paul CBD areas shown enlarged) region's share of low income households in Zone 57 was zero.

The responses of the two model packages to this input change were quite different. EMPIRIC showed virtually no response. Differences between this test run and CR were in no case more than 0.05% of CR. This result is contrary to one's intuitive expectations. Reference to the EMPIRIC equations in Table 5 of Chapter 2 will provide an explanation. Base year low income households (LIQ) appear only in the first equation, and would have the value 0.0044 (LIQ for Zone 57 divided by the region's total LIQ) in the CR. Given this value, the coefficient of -0.415, and the additive form of the equation, deletion of LIQ will produce only a very small change in ΔLIQ. This very small change appears on the right side of the equations for the other population types and government employment. In each case, the fractional coefficient and the additive form of the equation suggest that differences from CR will be minimal. In fact, they turn out to be negligible.

The DRAM response to this input change was quite a different matter. Compared to CR, the low income households in Zone 57 were down 65% in the projection year. Lower middle income households were down 42% in Zone 57. At the same time upper middle income households increased 16% while upper income households showed a 15% increase. The absolute sizes of the two upper income household class increases were, however, rather small, so the zone showed a net household decrease of 47% compared to CR. These results are very like actual metropolitan experiences with renewal attempts. Renewed center city zones often remain relatively empty for a time, with perhaps some increases (significant percentages but small absolute numbers) in upper and upper middle income households.

It is worth noting where the displaced households relocated. Of the 697 low income households removed from Zone 57 in the base year, 241 returned to that zone. Of the remainder 194 relocated in a ring of zones adjacent to Zone 57, mostly in those zones with prior large numbers of low income households. The remainder of the low income households relocated throughout the region, with emphasis on the next adjacent ring of zones and in two zones adjacent to the St. Paul urban core; in all cases, zones with relatively large numbers of low income households in the base year. (Table A1)[1]

A second DRAM run was made, differing in that the deleted low income households were not allowed to return to Zone 57. The pattern of these results was identical to those of the previous test. In all cases the results simply showed slightly larger responses. (Table A2)

Pursuing this line of experiment with DRAM, a second pair of test runs was made for Zone 49. This zone, in the CR was a rather rapidly growing upper-middle and upper income suburb. In the base year 384 low income households, 7.5% of the total households in the zone, were deleted for the test run. By the projection year there were 928 low income households, making 8.3% of the new total. In CR the corresponding figures are 1046 and 9.2%. For the projections of other household types in the test run the respective percentages were lower-middle income 19.0%, upper-middle income 34.6% and upper income was 38.2%. In CR the corresponding figures were 20.8%, 34.3%, and 35.7%. Thus, the elimination of a small low-income enclave in the base year did not prevent growth of low income households in in the zone, but did slightly retard that growth. (Table A3)

The test was rerun, with the same deletion, but precluding the return of low income households (perhaps simulating large lot zoning?). In this case there was a net decrease in the zone's population of 9.9% compared to CR. There were no low income households, a 23.0% decline in lower-middle income

---

[1] Table references at the ends of test run discussions are to tables included in Appendix II, which tabulate selected, relevant, run outputs.

households, a 0.2% increase in upper-middle income households, and an 11.3% increase in upper income households. The low income households which were prevented from locating in Zone 49 all located in the zones falling between suburban Zone 49 and the Minneapolis urban core. Apparently the exclusion of low income households from this suburban zone has the effect of preventing those households from leaving the urbanized area immediately adjacent to the city's core. (Table A4)

Having tested "urban renewal" in an urban core zone, and in a suburban zone, one more DRAM test was made, for Zone 56, midway between the core and the suburbs. In this case, 1000 low income households were removed from the base year. This was a 10.2% decrease in the low income households and 4.8% decrease in the total households in Zone 56. In both the test run and CR the zone's households increased substantially from the base year to the projection year. The test run showed fewer low income and low-middle income and more upper-middle and upper income households than did CR. These four household types showed growths of 82.9%, 56.7%, 26.1%, and 1.8% in the test run and of 102.6%, 67.6%, 25.7%, and -2.9% in CR. Thus again, a decrease in low income households in a zone in the base year retards the growth of the lower two income groups, and slightly accelearates the growth of the two upper income groups. Finally, even though the deleted low income households were not prevented from returning to the zone, most relocated in adjacent zones slightly closer to the urban core. (Table A5)

As a last test run in this series, a run was made where 1000 low income households were deleted and 1000 upper income households were added to Zone 56 in the base year. Comparison of this test run to CR yielded differences almost exactly double those found between the previous test run and CR. (Table A6)

To summarize the results of the first in this series of tests: 1) EMPIRIC shows no response in either population or employment, 2) DRAM shows excellent population response and almost no employment response. EMPIRIC's lack of response is not surprising given its equations and parameters, but is inconsistent with current theoretical and empirical findings. DRAM's response was sufficiently interesting to suggest a further set of tests. On their conclusion it appears that the population responses of DRAM are quite in keeping with our current understanding of the actual phenomena being simulated. DRAM's lack of employment response is perhaps explained by the fact that the population changes, while significant, are of small absolute magnitude and therefore do not stimulate a noticeable employment response.

Changes in Population: Low Income Household Increases

In the same way that a decrease in a zone's low income households may be used to crudely describe a simplistic form of urban renewal, an increase in a zone's low income households may be used to crudely describe a public housing project. A series of test runs was undertaken to study this phenomenon.

A zone midway between the urban core and the suburbs (Zone 37, see Map) was selected for the first of these tests. One thousand low income households were added to this zone in the base year. Once again, EMPIRIC showed virtually no response to this change in inputs. DRAM, once again, responded in a way consistent with theoretical and empirical findings by others.

In Zone 37 DRAM showed projection year increases, compared to CR, of 41.1% in low income households, 30.5% in low-middle income households, and

1.1% in upper-middle income households. Upper income households showed a 5.4% decrease compared to CR. Overall, Zone 37 showed an 11.3% increase in total househols in the test run, compared to CR. Thus the addition of the 1000 low income households in the base year (a 21.2% increase in the zone's base year total households) yields a rather strong tendency for the zone's household composition to change. In the control run, the composition is 18.0%, 22.3%, 29.2% and, 30.4% low income to high income respectively, and in the test run the composition is, 22.8%, 26.2%, 26.5%, and 24.4% for the four income groups, low income to high income respectively. Finally, we note that the household types which increased in Zone 37 were drawn from a ring of adjacent zones, and the upper income households which left Zone 37 dispersed to the ring of adjacent zones. (Table A7)

To further explore this phenomenon two further DRAM test runs were made for two suburban zones. In the first of these runs Zone 32 received an increment of 1000 low income households in the base year, and in the second run the increment was put in Zone 33 instead. The results in both these runs were virtually identical. Comparison of the test runs to CR showed increases in low income and low-middle income households and decreases in upper-middle and upper income households. For Zone 32 the percentage composition of households was 10.0%, 24.3%, 34.8%, and 30.9% (low to high income) in CR and changed to 13.9%, 31.4%, 31.0%, and 23.9% in the test. For Zone 33 the figures were 9.1%, 24.2%, 35.7%, and 31.0% for CR and 12.7%, 31.5%, 31.9%, and 23.9% for the test. Thus the introduction of low income household increments to suburban zones in the base year produced long term changes in the zones household composition. (Table A8)

A final test run of a low income household increment was made for a rural zone. Zone 94 is a rural zone which showed a net decline in population from the base year to the projection year. The bulk of this decline was in the low income households who dropped 66.4% from the base year to the projection year. During the same period the high income households increased 43.5% in the zone. The addition of 1000 low income households to this zone in the base period only partly altered its situation in the projection year. Low income households still declined, though by a somewhat smaller 47.4%. Lower-middle income households grew by 25.8% compared to a decline of 11.6% in CR. Upper-middle income was relatively unchanged, growing by 28.7% in the test run and by 24.5% in CR. Finally, high income household grew by 28.0% and 43.5% in the test run and CR respectively. (Table A9)

To summarize the results of these tests, where low income households are added to a zone, we find EMPIRIC not responding and DRAM responding as expected. Adding low income households in the base year changes the zone's projection year household distribution. The shift is in the direction of increases in low and low-middle income households and decreases in high and high-middle income households. The extent of the shift depends on the initial total population and population composition of the zone.

Changes in Population: Upper Income Increases

As a final set of DRAM population tests, two runs were made where upper and/or upper-middle income households were added in the base year. In the first of these tests 1000 upper income households were added to Zone 74 in the base year. This is a populous zone, well within the urban area, but not in the core. These new households represented a net increase of 16.7%

and a 40.0% increase in high income households. The result of this change was that the high income households in Zone 74 grew somewhat more in the test run than in CR, and all other income classes grew somewhat less. (Table A10)

The second of these runs was a test of Zone 94 (a rural zone used above for a low income increment test) in which an increment of 1000 high income plus 1000 high-middle income households was added to the zone in the base year. This produced results similar to those produced by the high income household increment, but not quite so pronounced. (Table A10)

Changes in Population: Summary

EMPIRIC shows no response to exogenous changes in base year population (households) DRAM shows responses consistent with both theoretical and empirical descriptions of urban phenomena. This is amply demonstrated by an extensive series of DRAM tests. Briefly stated, increases in low income households in a zone produce decreases in high income households, decreases in low income households produce increases in high income households, and increases in high income households produce decreases in low income households. Ripple effects are often observed in the ring of zones adjacent to the zone in which the test was effected.

Changes in Accessibility: Regionwide

Measures of the ease or difficulty of interaction between activities are central components of virtually all urban spatial models. This interaction phenomenon is contained in the several accessibility variables found in EMPIRIC, and is an integral part of the DRAM formulation in the form of the trip probability function. The next series of model runs described were intended to evaluate the models' response to changes in this variable.

Throughout the literature on transportation and urban development one finds the observation that where transportation is readily available and consequently, access is great, development tends to be spread out. Similarly when access is poor, development tends to be more concentrated and, in the case of larger regions, subnucleated. Recent experience throughout the United States amply demonstrates the generality of this phenomenon, with virtually every transportation improvement being closely followed by further spread of activities.

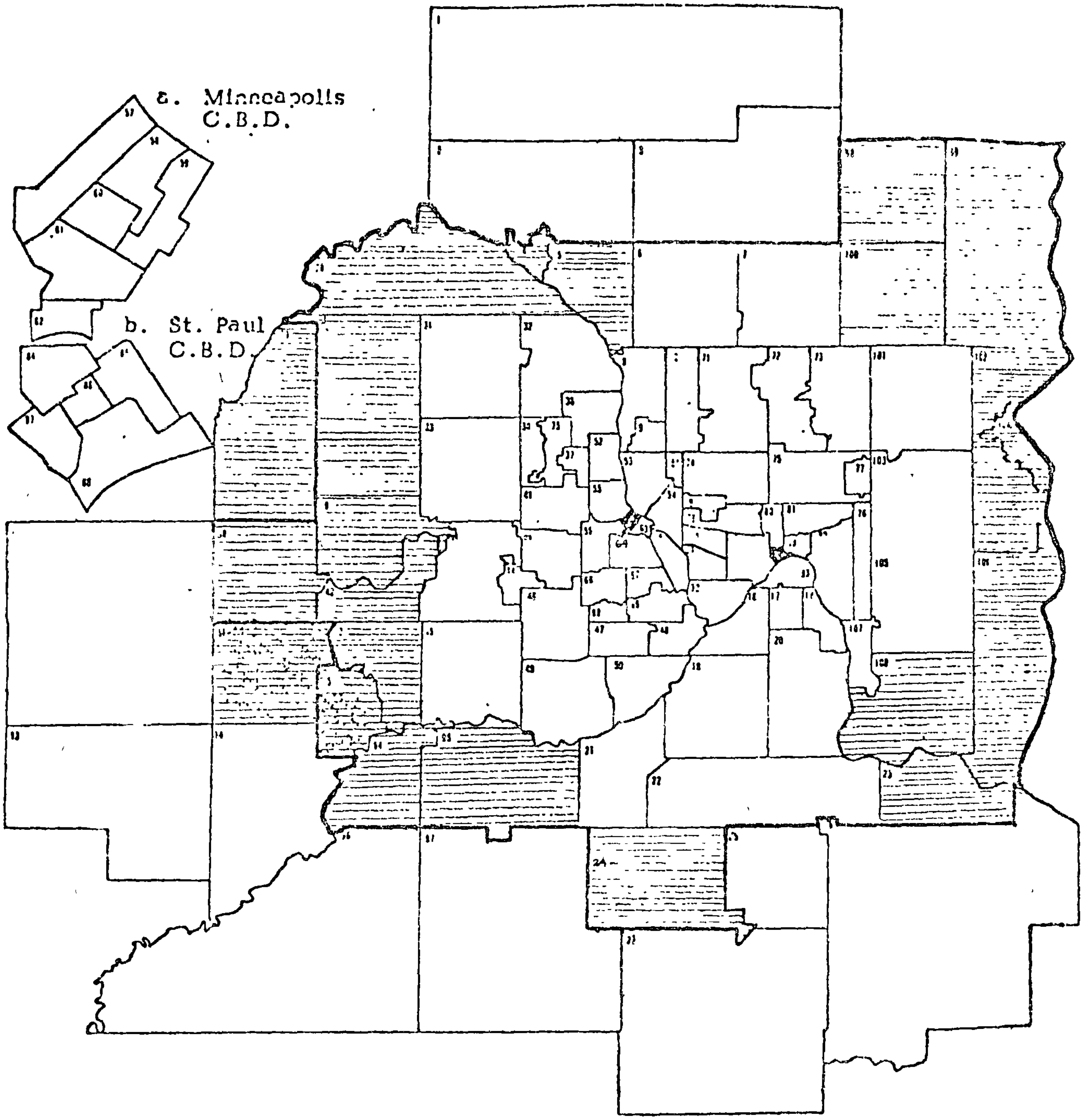
The first tests of the models' performance in response to access changes were with respect to regionwide changes. These tests might, for example, represent significant increases or decreases in fuel cost and/or availability. In the first run there was an arbitrarily imposed increase in impedance (i.e. an increase in travel time and/or cost and a subsequent reduction in accessibility). EMPIRIC showed no response to this input. Before discussing the DRAM response it may be useful to describe the mechanics of implementing these changes in the models.

In both models the initial datum is a zone-to-zone matrix of travel times or costs (or some composite figure). For the Minneapolis-St. Paul data base the interzonal travel times, estimated for an unloaded (uncongested) network were used in the models. While the congested travel times are preferable for these models, they were not available for use in this project. For use in DRAM, these matrices are simply one of the data inputs and are used directly in the running of the model. Changes in these impedances may be implemented by actually modifying the dataset or by adding the modification to DRAM's input routines, thus modifying the data as it is read in. In either case, a regionwide increase or decrease can be accomplished by multiplying by 1.1 or 0.9. Parking charges can be imposed by adding to the vector of impedances terminating in the zones where charges are to be instituted. Improvements between various zones and other zones can be imposed by multiplying the appropriate and/or columns of the impedance matrix. All in all, changes in impedance for DRAM are easy and direct.

In EMPIRIC, the procedures are more complex. The model package contains many programs for manipulations of data inputs and outputs. Modification of the inputs to EMPIRIC involves several steps of processing data through various programs. For impedances, there are many processing steps resulting in the several accessibility measures. Each of these measures is a vector of length n (n equals the number of zones). Finally, these vectors are converted to regional shares, i.e. scaled so their sum equals 1.0 exactly. It is precisely this scaling that results in EMPIRIC's nil response to regionwide changes in impedance, zone specific changes do produce responses as will be discussed below.

DRAM showed a consistent response to the first run of this set, which involved a regionwide 10% increase in impedances i.e. highway times and/or costs. The low income household response to this was mixed, with some fringe areas showing declines compared to the CR, and some showing increases. The other three income classes were, however, uniform in their relative decreases in the urban fringe zones. The implication here, are which certainly needs further study, is that ceteris paribus, low income households are the least sensitive to travel costs. It may be not so much a matter of insensitivity to changes in travel cost as a matter of inability to respond to those changes, due to other factors such as housing discrimination or limited employment opportunities. This, of course, ties in with the interpretation of the distance parameters discussed in the previous chapter. The net effect on all households is that 15% of the regions zones, all located at the urban-suburban fringe, showed relative declines of 10% or more compared to CR. At the same time, while less marked, employment showed some degree of centralization and a good deal of churning in the core and near core. Map 2 shows the zones with 10% or more decline in total households, the declines were absorbed (i.e. matched by increases) in the urban cores.

The second run of this set involved a regionwide 10% decrease in impedances. Again, DRAM's low income household response was mixed, with equal numbers of fringe zones showing gains or declines compared to CR. The lower-middle income households showed a strong tendency towards decentralization, with gains of 10% or better in a ring of fringe zones completely surrounding the metropolitan area. Upper-middle income households and high income households also showed strong decentralization response to this regionwide impedance decrease. At the same time, Basic 1 employment declined in the urban core, Basic 2 employment showed signs of beginning suburbanization, and Non-basic employment declined in the urban core. The net effect of these impedance decreases is a substantial decentralization of population and the beginnings of decentralization of



Map 2: ZONES SHOWING GREATER THAN 10% DECLINE IN ANY HOUSEHOLD SECTOR DUE TO REGIONWIDE IMPEDANCE INCREASE

employment, as compared to CR. Map 3 shows zones with a 10% or more increase in total households, the corresponding declines were all in and adjacent to the urban core. In short, relative decreases in transportation costs encourage further urban sprawl, while relative increases in transportation costs discourage sprawl and perhaps even encourage re-centralization.

Changes in Accessibility: Zone Specific

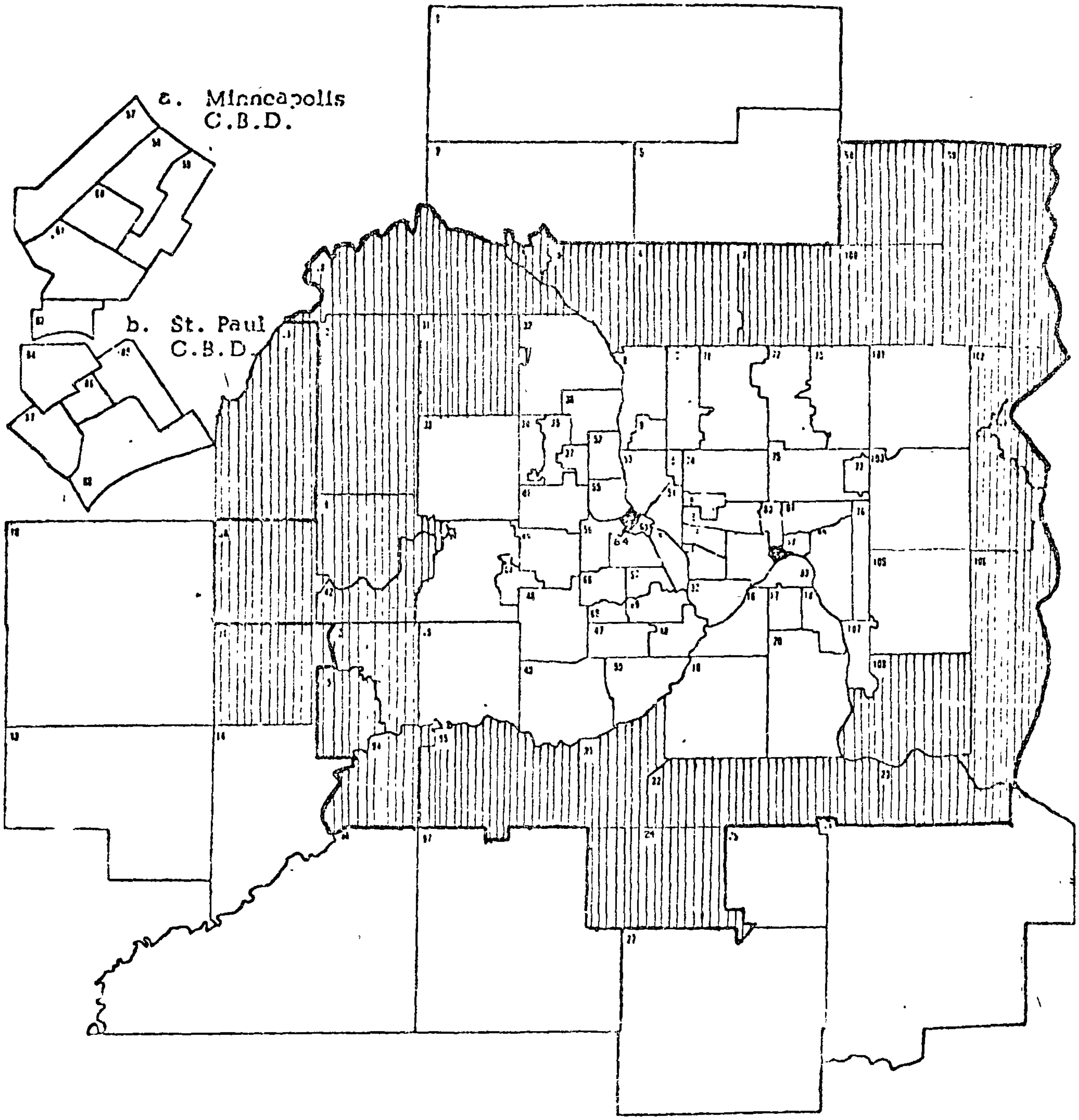
The second group of access tests involved changes in impedance between specific zones and the rest of the region. The first in this group of tests involved a small improvement in the accessibility, to the rest of the region, of a zone on the suburban-rural fringe (Zone 24). This improvement was in the form of a 5% reduction in impedance from that used in the CR. Theoretically one would expect modest increases in population and employment in the affected zone. Since their impedances (access) to the region would remain unchanged, there would be little or no spillover effect to surrounding zones. We note that an actual transport improvement to one zone would simultaneously affect others because the network would connect many zones to each other via the improved link(s). In an integrated transportation and land use model this could easily be simulated as the transportation system is described and used in link-by-link form.[1] This was not possible in the present project due to the use of impedance data in lieu of the actual networks.

The EMPIRIC response to this test was minimal. No population or employment sector changed as much as 1% from the CR.

The DRAM run showed responses more in accord with expectations. All household types showed increases between 10% and 19%, with total households

---

[1] Putman, S. H. (1974) "Preliminary Results from an Integrated Transportation and Land Use Models Package", Transportation, Vol. 3, pp. 193-224.



Map 3: ZONES SHOWING GREATER THAN 10% INCREASE IN ANY HOUSEHOLD SECTOR DUE TO REGIONWIDE IMPEDANCE DECREASE

in Zone 24 increasing by 13.3% compared to CR. Total employment in the zone showed a slight increase of 1.1% compared to CR. The non-basic employment showed the greatest increase, 9.4% compared to CR. There were virtually no effects in surrounding zones, as the absolute change in households in the zone was only 221, with only 14 more employees. (Table A11)

The second test of this pair involved a 20% improvement in accessibility of the same zone (Zone 24) to the rest of the region. The expectation was that the results should be simply an amplified version of those from the previous test.

The EMPIRIC results were still minimal. The population changes were still less than 1%, though the total employment for the zone did show a 3.7% increase. The DRAM results showed a 67.9% increase in total households and an 8.0% increase in total employment. Though these changes seem large, it must be remembered that a 20% improvement in accessibility of an individual zone to the entire region is a phenomenal increase in accessibility; almost equivalent to replacing an unpaved road with an expressway. (Table A12)

The same pair of tests was then repeated for an urban core zone (Zone 64). The first of these runs was for a 5% accessibility improvement for the zone. The EMPIRIC response to this change consisted of very slight declines in the lower two income classes and modest increases in the upper two income classes. Total employment increased by approximately 2% in the zone. This is the first EMPIRIC run to show any noticeable response. The results are in accord with what one expects from looking at the equation coefficients, but not quite what might be expected from other theoretical and empirical findings. It is not clear why improved accessibility of an urban core zone should cause a decline in its largest population groups. Typically when expressways have connected to urban core areas the areas have experienced declines in upper income groups and increases in lower income groups.

The DRAM response to this change in accessibility was a 4.4% decrease in upper-middle income households, increases of approximately 2% in low-middle income and high income households, and a substantial increase of 20% in low income households. These results are more in accord with our expectations. Employment in the zone shows a total increase of 10.3% compared to CR. (Table A13)

Rerunning these tests with a 20% accessibility improvement produces similar, but stronger responses in all cases. We note that in the EMPIRIC run the decline in low income households was taken up by adjacent urban zones and similar zones in the St. Paul urban core. The increase in high income households was at the expense of suburban zones running northwest from the Minneapolis urban core. In the DRAM runs the low income increment was drawn from adjacent urbanized zones, while the upper-middle income decline was taken up by several suburban zones. (Table A14)

In summary, DRAM is again more responsive than EMPIRIC, in this case to changes in accessibility. EMPIRIC shows no response whatever to regionwide changes in accessibility. DRAM shows increased urban sprawl or decentralization with regional improvements in access, and decreased sprawl or centralization with regional access decreases. When access to a specific zone is increased, DRAM shows increases in population and employment. If the improvement was for a suburban zone, all types of population increased. When the improvement was for an urban area the principal increase was in low income households, with a decline in upper-middle income households. All in all, even though this was the first set of tests to evince any response from EMPIRIC, the response did not seem to be correct. DRAM, again gave the proper response though we have some minor reservation as to whether it may have been an over response.

Changes in Employment: Basic

A third set of tests of the models was run in the form of changes in base year employment in a zone. In each case an arbitrary increase was added to the base year employment in a particular zone. The regional forecasts for the projection year were not changed, so that any projection year increase, compared to CR, in one zone was at the expense of some other zone. A large number of test runs of this sort were made, but only some of them, enough to describe the model's responses, are discussed here.

Again, an urban and a rural zone are described, the same zones (Zones 64 and 24 respectively) as were described in the accessibility tests.

Taking first a 10% increase in basic (the sum of BASIC 1 plus BASIC 2) employment for Zone 24, EMPIRIC shows virtually no response. DRAM shows virtually no population response, and a total employment increase of 3.3% for the zone. Thus in both cases the 10% basic employment increment was dispersed throughout the region with no significant effect. (Table A15)

The second set of runs had a 30% increase in basic employment in the base year for Zone 24. EMPIRIC again shows virtually no response. DRAM shows a total population response of less than 1% increase above CR. There is a net employment increase of 10%. This looks small until realizing that the 30% increase in the base year was 84 BASIC 1 employees and 5 BASIC 2 employees. This yields an increase of 113 BASIC 1 and 8 BASIC 2 employees over CR for the projection year. This is a multiplier of 1.35 and 1.60

for each of these employment types. These changes were so small as to have negligible effects on adjacent zones. (Table A16)

A similar pair of test runs was made for the urbanized Zone 64. The DRAM run showed increases of 13.6% and 12.4% for BASIC 1 and BASIC 2 respectively in response to the 10% basic employment increment. But as basic employment is only 25% of the zone's employment, the total employment in the zone shows a net increase of 2.3% compared to CR. Population therefore shows virtually no change. (Table A17)

The EMPIRIC run of the 10% test showed no population change and an employment change of 14.4% increment to BASIC 2. The EMPIRIC run of the 30% test showed no population change and a 43% increase in BASIC 2. (Table A17)

The DRAM response to the 30% test showed small changes, less than 1%, in population and 41% and 37% increases in BASIC 1 and BASIC 2 respectively. The net employment increase for the zone was almost 7%. There were modest increases in low income households in all the zones adjacent to the Zone 64 test zone, at the expense of urban core zones in both Minneapolis and St. Paul. (Table A18)

Changes in Employment: Non-Basic

A similar set of runs were made with changes in non-basic employment in specific zones in the base year 1. These results were similar to the Basic employment runs. DRAM showed small (almost negligible) changes in population, and employment changes mostly made up of the exogenous change. EMPIRIC showed no population response and conflicting patterns of employment changes. (Tables A19-A22)

To summarize the results of changing base year employment, neither model shows much response to significant percentage increases in employment if they are not significant absolute increases as well. DRAM does demonstrate an employment multiplier effect, in that a base year increase of X% basic employment yields a projection year increase of $(1+\alpha)X\%$. EMPIRIC shows no response when suburban or rural zones are tested. In urban areas only BASIC 2 employment changes in the projection year even though both BASIC 1 and BASIC 2 were changed in the base year. EMPIRIC in no test shows any population response. In DRAM significant population response resulted only from large employment changes and was principally a matter of low income household increases in zones where there had been large employment increases.

Changes in Land Availability

The last group of runs to be discussed here involved tests of several degrees of land conservation policy. Each of these runs adopted a different policy as to the definition of open space for preservation. In effect, each of these runs removed different amounts of land from the available developable land in each zone.

The first set of runs deleted floodplain areas from available developable land. These areas were, of course located along the various rivers and streams that flow through the area.

The EMPIRIC run of this test showed moderate increases of low income households in older, but not core, urban areas. There was no significant movement of lower middle income households, while there were large declines in the upper two income groups for those same urban, but not core, areas. The low income increases were drawn from throughout the region, and the two upper income decreases were up in zones throughout the region. The employment response was mostly a matter of modest churning movements throughout the region.

The DRAM run of this test showed modest decreases in all income classes, throughout the region, taken up by large increases in zones adjacent to the urban core areas. The lowest and highest income households were least affected, with only modest changes. The two middle income household groups showed more substantial changes. Employment showed modest churning throughout the region.

The second set of runs deleted both floodplains and aquifer recharge areas. The sum of these began to be a substantial amount of land. Both the EMPIRIC run of this test and the DRAM run of this test showed considerable churning of both households and employment. A detailed, zone-by-zone, analysis of these runs is the only way to properly describe the results since the policy protected acres do not conform to the more traditional urban vs. suburban or rural sorts of descriptions. Suffice it to say that constraints on the use of land of the magnitudes involved here, produce rather substantial locational effects throughout the region.

In the next set of runs the land policy was even more restrictive, prohibiting use of floodplain, aquifer recharge areas, and wetlands. As would be expected both EMPIRIC and DRAM responded to this policy with even more churning of households and employment. It is interesting to note that their responses were very different, one from the other, even to the extent of being almost opposite. Closer investigation revealed that while the response was smaller than would be expected, the EMPIRIC responses were in the proper direction. The DRAM responses were, in some zones, backwards.

This was traced to discrepancies in the policy descriptions which attempted to preserve more land than was available. Consequently it was not possible to fully evaluate the DRAM responses to these policies in this study. Subsequent work with the model has corrected these problems so that these policy tests will eventually be properly examined.

CHAPTER FOUR: CONCLUSION

When, in the early 1960's, the first urban computer simulation models were being developed, one of the principal goals was to develop the capability of assessing the consequences of various urban renewal plans on the spatial distribution of activities. It was hoped that different public policies capable of altering the mix of activities in a zone could enter the models in various forms. The arrival or departure of an employment facility would induce significant effects in the model outputs. The arrival of a number of households of a particular income class might well result in changes in location of other households and perhaps of some employees too. Similarly the departure of a group of households would probably further, induce changes in a zone's activity mix.

Further, it was hoped that the density and degree, or extent, of development in a zone would also be affected by policy inputs. Clearance of certain types of structure would change density as would the erection of new structures. The construction of large new development, say of single family residential homes, or at a different density -- of apartments, would change both the zone's density as well as its extent of development. These changes would induce other changes, both in employment and in population location. In a related way, changes in the amount of land available in a zone should affect future location of activities in a zone. More stringent land use controls, having the effect of reducing available land, will change the pattern of activities locating in a zone. Similarly holding back land from development should also result in changed location patterns.

Finally, the spatial separation of activities from each other was expected to be a key variable in these models. This variable is usually expressed in terms of travel times and/or travel costs between zones and activities. Thus any substantial change in the transportation facilities should result in a change in activity distributions.

Many modelling projects were begun, with very few being successfully completed. It was a chaotic time for urban modelling. Each model had its proponents who claimed that their's was "the way". Not many of these models have survived, though there are occasional uses of one-time-only models or newly developed ones. The majority of recent model applications have been of either EMPIRIC or Lowry derivative models, with basic research efforts being performed independent of ongoing applications. Thus it seemed to be a good time to assess these two most-used models and to subsequently provide some guidance as to future applications of existing models as well as to directions for future research efforts.

The results of this project are quite clear. EMPIRIC achieves good fits to base data, but is not adequately sensitive to changes in input variables. This is probably due to its lack of an explicit theoretical form. The model has, however, been very useful for shorter term urban projections and it should be remembered that at first, even its authors claimed associative validity, rather than any genuine theoretical validity.[1]

The best of the Lowry derivative models in current U.S. use would not have compared especially well to EMPIRIC. Its theoretical structure is rudimentary, its disaggregation of population types is accomplished

---

[1] Hill, D.M. (1965) op.cit.

independent of the location procedure, and it relies on several exogenously defined constraint mechanisms to achieve its relatively good fits to base data. Finally, there was no standardized procedure for a statistically valid estimation of its parameters. Taking a cue from current British modelling practice, the model was reformulated in a more theoretically correct form so as to 1) allow for explicit disaggregation of the population as part of the location process, 2) eliminate the need for the exogenous constraint procedures, and 3) define the proper method for estimating its parameters. At the same time, the model differed from current British practice by making use of a multivariate zonal attractiveness measure, a particular feature of many of the U.S. developed Lowry derivatives.

This new model, called DRAM, did not fit the base data quite as well as EMPIRIC. However, its response to changes in input variables was excellent. This, coupled with the empirical confirmation of the model's form by its parameter estimates, indicates that it would give more accurate forecasts than EMPIRIC. This is expecially the case when the forecasts are of responses to policy inputs.

Current research with DRAM is proceeding in several directions. First, attempts are underway to routinize its parameter estimation procedure. This procedure, which utilizes mathematical search procedures, is no more complex than multiple regression, but is less well known and thus may cause some apprehension in potential users of the model. It is hoped that several case study applications of the model will help to ameliorate these problems. Second, further testing and improvement of the model itself is underway, including its fitting to as many different data sets as possible in order to test its consistency for different urban areas. Finally, DRAM has been incorporated in the Integrated Transportation and Land Use Package - ITLUP as a part of an ongoing research effort.[1]

It is perhaps only a little presumptuous to suggest that this work be used to mark the end of an era. For policy sensitive forecasting applications it would seem to be difficult to justify using anything other than a Lowry derivative model, perhaps DRAM, and calibrating it by the procedures discussed in this report. For future research efforts, it seems reasonable to suggest attempts to extend the theoretical structure of this model in the direction of bridging the gap to micro-economic theories of behavior on the one hand and in the direction of fuller integration of models with planning processes on the other. Were these suggestions to be followed, applied models would reflect the most advanced techniques practical for planning purposes at the same time that their results would provide feedback in the form of empirical results which could influence ongoing research in urban spatial dynamics.

---

[1] Putman, S. H. (1974) op.cit.

Appendix J

DRAM paper

Calibrating a Disaggregated Residential Allocation Model - DRAM

Stephen H. Putman
Associate Professor of City and Regional Planning
University of Pennsylvania
Philadelphia, Pa. 19174

to be presented at:
Regional Science Association - British Section
Eighth Annual Conference
London, Sept. 4-5, 1975
(Revised Dec. 1975)

ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation, and in part by the U.S. Department of Transportation, via Grant APR 73-07840-A02, "Development of an Improved Transportation and Land Use Model Package."

A large portion of the computer work necessary to produce these results was done by Fred Ducca, a graduate student in the Department. Earlier work on fitting portions of the models was done by Chuck Sawyer and Rob Mathie, both graduate students in the Department.

Introduction

As part of a National Science Foundation sponsored effort to compare the performances of two different land use models calibrated on the same data base, several fundamental problems in urban land use modelling have been encountered and partially resolved. In particular, the fact that no Lowry derivative land use model had ever been properly calibrated in U.S. practice became abundantly clear. In order then, to accomplish the desired comparison of different models on a common data base, it became necessary to develop a calibration procedure for these models. The development of this calibration procedure in turn suggested a reformulation of the model which appears to be much superior to the original and which is sufficiently different to justify a new name, Disaggregated Residential Allocation Model - DRAM, which will differentiate it from its predecessor. A rather unique characteristic of this model, cast in entropy maximizing form, is its multivariate attractiveness measure.

Background

The development of the Lowry model of land use distribution (Lowry, 1964) along with that of numerous derivatives of its basic model structure has been described elsewhere (Goldner, 1971; Putman, 1975). Some years after development of these models had begun in the U.S., substantial further development of them was undertaken in Great Britain (Batty, 1972). Interestingly, despite the model's originating in the U.S., some of the most fruitful work in extending the concept has been done in recent years in Great Britain. Further, and of critical importance to applications of the model, the question of estimation of the model's parameters has to the knowledge of this author, never, with perhaps one exception (Voorhees, 1972) been properly settled in any U.S. work. In contradistinction, it appears that the British work has produced rather conclusive evidence as to the means by which these models may be calibrated (Batty, 1970; Batty and Mackie, 1972).

A modified version of the Incremental Projective Land Use Model (IPLUM) was used in the Integrated Transportation and Land Use Package (ITLUP). This ITLUP version of IPLUM is fully described elsewhere (Putman, 1973). In brief, the residential portion of this model allocates increments of residential locators to their places of residence in response to increments in basic employment and changes in the transportation facilities. This response is determined by a probability function which describes the distribution of work trips, and by a measure of residential attractiveness for each potential location zone. The purpose of this paper is to outline the steps thought to be necessary for a proper calibration of the ITLUP-IPLUM, and which ultimately led to the development of DRAM.

Virtually all Lowry derivative models used in the U.S. have as their residential allocation function some form of the following expression.

$$N_i = g \sum_j P_{ij} E_j \qquad (1)$$

where,

N_i = number of residential locators locating in area i

P_{ij} = the probability of living in area i and working in area j

E_j = the number of employees in area j

g = a scaling factor such that the sum of the N_i over all i equals an exogenous control total there are often other scaling or multiplier factors to convert from employees to households and to assure internal consistencies of various types.

The p_{ij} is most important component of equation (1). In the original Lowry model, the function used was:

$$p_{ij} = (D_{ij})^{-1.33R} \qquad (2)$$

where,

D_{ij} = airline distance between the centroids of area i and area j

R = number of zones in an annulus D_{ij} miles from the origin

In various of the Lowry derivative models, p_{ij} is modified to include measures of the attractiveness of area i. In particular in the ITLUP form of IPLUM,

$$p_{ij} = f(D_{ij})O_i \qquad (3)$$

where,

$$f(D_{ij}) = (\beta/D_{ij}^2) \quad \exp(\alpha - \beta/D_{ij}) \qquad (4)$$

O_i = a measure of residential "opportunities" in i

D_{ij} = travel time between centroids of zones i and j

α,β = empirically derived parameters

The measure of opportunities is basically an adjusted measure of residential holding capacity (previous level of residential density times amount of available land). The adjustment Q_i is a logistic curve function of, the proportion of the developable land in zone i which has been developed by the end of the base time period.

This is:

$$O_i = a_i^v (h_i / a_i^r)(Q_i) \qquad (5)$$

where

a_i^v = vacant acreage in zone i

h_i = housing units in zone i

a_i^r = residential acreage in zone i

Q_i = development level factor

and, where

$$Q_i = 1 - \frac{\gamma}{(1-\gamma)\exp(\delta x_i^2)} \qquad (6)$$

where

γ, δ = parameters

x_i = the percentage of developable land area in zone i which has been developed

The parameters of the trip function were estimated by fitting the equation to observed work-trip distributions from the San Francisco area. The parameters of the development level function Q_i have not been statistically estimated nor has the complete p_{ij} function been fit to any actual data. It was precisely this fitting of the complete p_{ij} function which was necessary, but which had never been done (excepting the Voorhees attempt) during U.S. work with Lowry model derivatives.

Reformulation of the Model

In all of these models the essence of the residential allocations is either the work-trip (home-to-work or work-to-home) or a combination of the work-trip with measures of attractiveness of the potential residential locations. Implicit therefore, in any of these models' estimates of residential locations, is a set of work trip estimates as well. Very little use has been made of this fact in U.S. practice. Yet, it is precisely the fact of these implicit trip matrices that leads to a more satisfactory method of estimating these models' parameters. The use of IPLUM in the ITLUP package is a particular exception to the usual ignoring of these implicit trips. In this case these implicit work trips are made explicit by extraction of the trips from the model directly. These trips are later used to load the transport network (Putman, 1973).

It is a virtue (and perhaps in the first instance was the source) of the Wilson entropy maximizing approach to analysis of these models that the question of these trips is made explicit (Wilson, 1967). For example, the Lowry model may be rewritten, based on this approach as (Wilson, 1970),

$$T_{ij} = E_j f(C_{ij}) \tag{7}$$

where

T_{ij} = number of persons working in zone j and residing in zone i

E_j = number of persons working in zone j

C_{ij} = impedence (usually travel time or travel cost) between centroids of zone i and zone j.

An important problem of this formulation is that there is no constraint on the sums of trips. Without the constraint there is no reason to expect that,

$$\sum_i T_{ij} = E_j \tag{8}$$

This implies that the number of employees in zone j will not necessarily equal the sum of the employees residing in all zones i who claim to work in zone j.

A simple residential location model may be derived from entropy maximizing concepts as follows,

$$T_{ij} = A_i B_j O_i E_j f(C_{ij}) \tag{9}$$

where

T_{ij} = trips between zones i and j or, number of persons living in zone i and working in zone j

O_i = trip origins or, employed persons living in zone i

E_j = trip destinations or, employees employed in zone j

A_i = balancing factor for trip origins

B_j = balancing factor for trip destinations

(C_{ij}) = impedance function

It is possible to replace the trip origins O_i by a measure of attractiveness of the origin zone, W_i.. This eliminates the need for the origins balancing factor A_i thus giving

$$T_{ij} = B_j W_i E_j f(C_{ij}) \tag{10}$$

In order for the constraint on the sums of trip destinations, equation (8), to be met, we have

$$B_j = \frac{1}{\sum_i W_i f(C_{ij})} \tag{11}$$

It is informative to substitute this expression back into the original equation (Senior, 1973), which yields

$$T_{ij} = E_j \left[\frac{W_i f(C_{ij})}{\sum_i W_i f(C_{ij})} \right] \tag{12}$$

If the term $W_i f(C_{ij})$ is called an "accessibility attractiveness" measure, then the fraction in equation (12) is a relative measure of the accessibility-attractiveness of zone i to zone j compared to all other zones i. Further, it is clear that the total number of employed residents residing in zone i is

$$N_i = \sum_j T_{ij} \tag{13}$$

and, substituting

$$N_i = \sum_j \left\{ E_j \left[\frac{W_i f(C_{ij})}{\sum_i W_i f(C_{ij})} \right] \right\} \tag{14}$$

If one is willing to assert that,

$$p_{ij} = W_i f(C_{ij}) / \sum_i W_i f(C_{ij})$$

then equation (14) is equivalent to saying

$$N_i = \sum_j E_j p_{ij} \tag{15}$$

which is the same function as the Lowry model, described in equation (1).

Thus it can be seen that the IPLUM allocation procedure may be considered, in the context of the entropy maximizing formulation, as a simple residential location model. However, IPLUM is a dynamic model in that it estimates changes in the number of residential locators, as follows:

$$\Delta N_i = \sum_j (\Delta E_j)\ p_{ij} \tag{16}$$

where

ΔN_i = change in the number of employed residents of zone i from time to time t+1

ΔE_j = change in the number of employees in zone j from time t to time t+1

P_{ij} = probability that a person will live in zone i and work in zone j, at time t+1

A question arises here as to whether ΔP_{ij} might be more appropriate in the new formulation than P_{ij}? Resolution of this question leads unfortunately to the question, among others, of location of in-migrants versus location of intra-metropolitan movers. In-migrants probably make their location decisions somewhat differently than the intra-metropolitan movers. None of Lowry class of models deals properly with this question. The TOMM models (Crecine, 1964, 1969) do so in a very superficial way by means of the "stable-household" functions. It was not possible to resolve this problem in the current work, so the existing practice of using P_{ij} has been maintained for the present. Further, as will be discussed below, ultimately it was the static form of the model which was estimated.

Calibration: Initial Discussion

To date, virtually all U.S. attempts to calibrate these models have involved assorted procedures, no one of which achieved any more than a partial calibration of the allocation function. Some procedures have fitted an $f(D_{ij})$ as in equation (2) or equation (4) to observed trip data, without taking into account the effects of the characteristics of the origin zone or destination zone. Other calibration attempts have fit a function with N_i as the dependent variable and various characteristics of zone i as independent variables, thus ignoring any explicit consideration of the trip distribution. Neither of these two procedures nor any of their many variations is capable of properly estimating the parameters of such a model.

For a model expressed in the form of equation (9), the only parameter(s) to be estimated is/are the parameter(s) which may be included in $f(C_{ij})$. It has been shown that in the fitting of parameters for such a model, statistics summarizing the goodness of fit of the work trip distributions were much more sensitive to changes in model parameters than statistics summarizing the goodness of fit of the activity distributions (Batty, 1970). This result argues for the use when possible, of work trip statistics as criteria for model calibration. Other work has derived several summary statistics of the work trip distributions, each of which is appropriate for particular functional forms of $f(C_{ij})$ (Hyman, 1969).

A problem posed by the form of the model shown in equation (10) is that W_i, the attractiveness measure, is not a directly observable or measurable variable. In one model effort, number of dwellings in zone i or population in zone i were proposed as proxy measures of W_i (Cripps and Foot, 1969). Population was finally selected and produced quite acceptable calibration results. In another model effort, usable land area in zone i were suggested as proxy measures of W_i (Barras, et. al., 1971). In both of these cases, by using a single proxy variable for W_i, calibration of the model remains as a matter of estimating the parameter(s) of $f(C_{ij})$.

In these cases, as well as those using the original form of the model in equation (9), the calibration process involves; a) selecting starting values of the parameters, b) estimating the trip distribution, c) comparing the estimated trip distribution to the actual trip distribution, d) revising the parameter values, and e) iterating to find the best fit parameter values. Work has been done on efficient means of doing this (Hyman, 1969; Batty and Mackie, 1972).

At this point, regretfully, it becomes necessary to introduce a troublesome consideration, the need to disaggregate the residential locators into types. First we acknowledge that this disaggregation may easily be described in terms of the entropy maximizing approach, by considering T_{ij}^{kw} to be the number of employees of type w who work in zone j and live in type k housing in zone i. An appropriate set of equations and constraints can be developed to cover this situation as well as several others (Wilson, 1970). Solving such a model involves an endogenous procedure for estimating the housing stock by zone. This is not a welcome prospect for our current research efforts though clearly it is a consideration for the future. What is necessary then is a model of the form of equation (10), but disaggregated only by type of locator. This may be written

$$T_{ij}^{k} = B_j^k E_j W_i^k f_k (C_{ij}) \tag{17}$$

then

$$B_j^k = \frac{1}{\sum_i W_i^k f_k(C_{ij})} \tag{18}$$

Finally, it seemed desireable to investigate the use of a multivariate attractiveness measure. There is empirical evidence that the attractiveness of zone i is a function of, among other vaiables, the distribution of household types living in zone i (Putman, 1973). This evidence suggests that the attractiveness of a zone to a particular household type is a function of the zone's percentage composition of household types. Further, the amount of developable land in a zone seems to be a determining factor in residential location, as does a developability factor which appears to act as a proxy variable for the extent of the available urban infra-structure. Thus a W_i^k may be defined as follows:

$$W_i^k = \left[\sum_g a_g^k (N_{ig} / \sum_g N_{ig})\right] r_i V_i Q_i \qquad (19)$$

where

a_g^k = parameters to be estimated

N_{ig} = number of households of type g in zone i, note the g household types correspond directly to the k household types

r_i = residential density - households/acre in zone i

V_i = available, developable, vacant land in zone i

Q_i = development level factor - see equation (6)

The parameters in the expression for Q_i may be estimated independent of the rest of the model. The parameters a_g^k need to be estimated within the structure of the model. In addition, the parameter(s) of the $f_k(C_{ij})$ must also be estimated within the structure of the model.

The precise form of the model desired would be, as per all the previous discussion, dynamic rather than static,

$$\Delta T_{ij}^k = B_j^k W_i^k (\Delta E_j^k) f_k (C_{ij}) \qquad (20)$$

To do this it would be necessary to have data for ΔT_{ij}^k and ΔE_j^k. At the time when this work was being done, these data were not available, making it impossible to estimate any but the static model.

In order to specify data requirements it will be helpful to write out the model in full.

$$T_{ijt}^k = B_j^k W_{it}^k E_{jt}^k f_k(C_{ijt}) \qquad (21)$$

Substituting in for B_j^k and W_i^k

$$T^k_{ijt} = E^k_{jt} \left[\frac{\sum_g a^k_g [N_{igt} / \sum_g N_{igt}] r_{it} V_{it} Q_{it} f_k(C_{ijt})}{\sum_i \left\{ \sum_g a^k_g [N_{igt} / \sum_g N_{igt}] \right\} r_{it} V_{it} Q_{it} f_k(C_{ijt})} \right] \tag{28}$$

Thus the required data are

T^k_{ijt} = the number of persons of type k employed in area j and living in area i at time t

E^k_{jt} = the number of persons of type k employed in zone j at time t

N_{igt} = the number of households of type g living in zone i at time t

r_{it} = residential density (households/acre) in zone i at time t

V_{it} = vacant developable land in zone i at time t

Q_{it} = development index, as described above, for zone i

C_{ijt} = travel cost (impedance) between the centroids of zones i and j at time t

Before discussing the calibration results, the perspicacious reader may have noticed a further problem, which exists with the definitions of T^k_{ijt}, E^k_{jt}, and N_{igt}. The E^k_{jt} are defined as number of persons of type k working in zone j at time t, and the T^k_{ijt} are number of persons of type k employed in area j and living in area i at time t. The N_{igt} however are number of households of type g living in zone i at time t. Clearly a conversion from employees to households is necessary at some point in the process. In order to simplify conversion of the T^k_{ijt} to vehicle trips for use in the network model, it will be most convenient to make the conversion at the residence end of the trip. Thus a matrix for converting households of type g to employees of type k must be developed from regional data for the regions to which the model is being fit. This was done for both the San Francisco and the Minneapolis - St. Paul applications, but the use of these regional conversion rates accross the board, makes it necessary to keep careful track of this conversion throughout the calibration process.

Calibration Results: Partial Estimates

It was initially intended that before the complete model equation was fit, preliminary estimates of its parameters would be developed by partial estimation of them by least squares regression. This was later found to be unnecessary, but some of the results related to the independent fitting of the trip distribution are of some interest.

It will be recalled that in equations (2) and (4) above, the distance functions used in the Lowry Model and in PLUM were given. These are but two of a vast number of functions which could be fitted to tripmaking data. To test several of these, a tabulation of the first work trips from the San Francisco Home Interview data file was prepared. These trips were tabulated according to the household and employment classes enumerated above for the 291 zone areal system. The distributions were then normalized and the resulting distributions were fit, using a non-linear least squares procedure, to several different functions. The work trip distributions took the familiar form shown in Figure 1.

Of the various functions investigated, the several varieties of gamma distribution seemed to produce the best fits. The general form of this

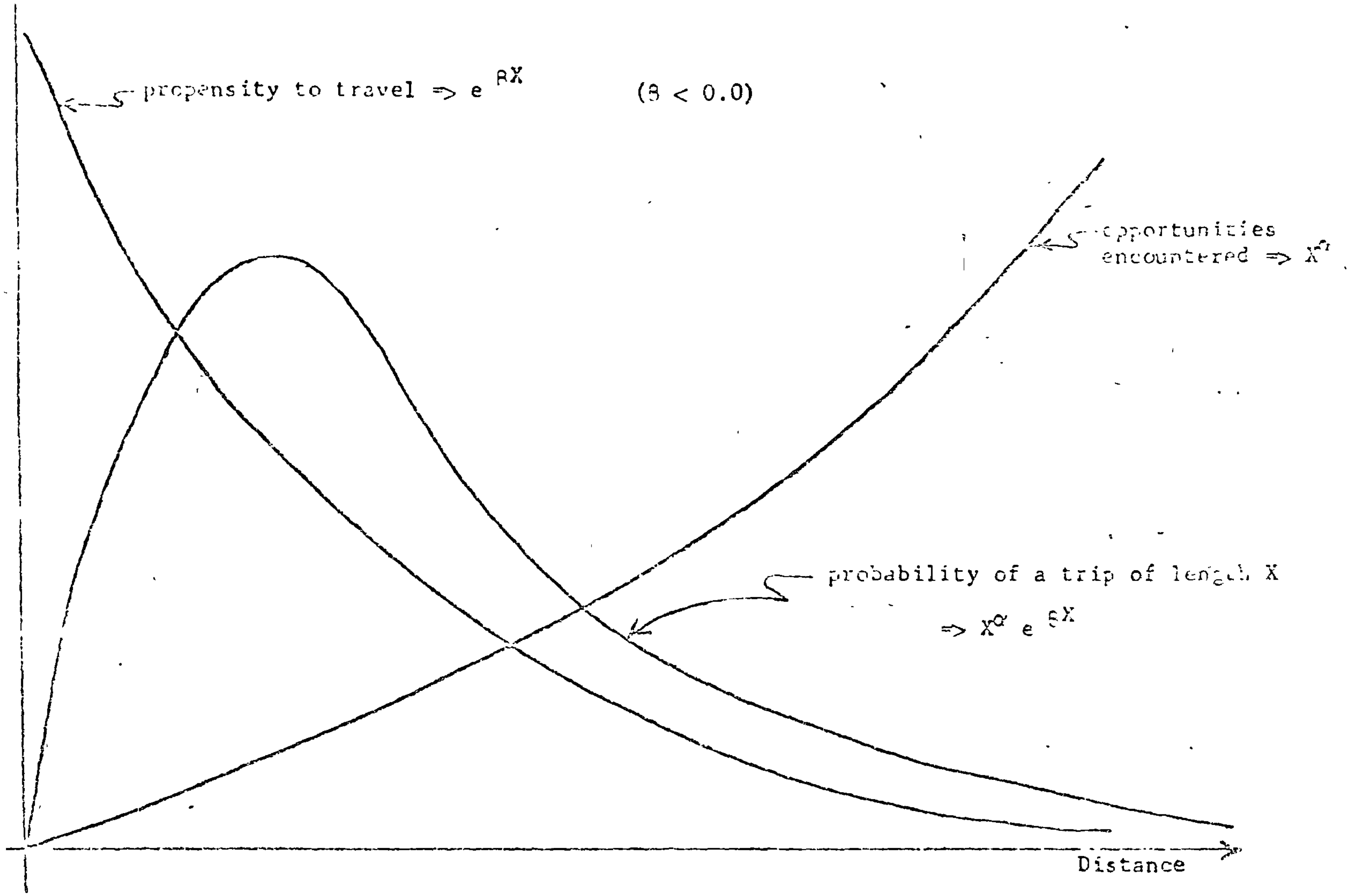


Figure 1: Trip Distribution Formulation

distribution is:

$$y = x^{\alpha} \exp\left\{ f(x) \right\} \qquad (23)$$

where

y = number of trips, or trip frequency

x = trip time or cost

The specific functions which best fit the data were sometimes best in one household income class and sometimes best in another. No one function was best for all four income classes. The function selected for further work on this prototype effort was

$$y = x^{\alpha} \exp(-\beta x) \qquad (24)$$

This function, known as Tanner's function, had been used in this type model elsewhere (Cripps and Foot, 1969). The best fit parameters for the 291 zone system in San Francisco are shown in Table 1. These parameters in Tanner's function do yield the skewed, peaked, curve shown above.

In the calibrations described below, the San Francisco data were aggregated to a 30 zone system, thus increasing to greater than eight minutes the three minute average travel time between adjacent zones of the 291 zone system. At that scale all the values of α become negative and Tanner's function takes on the appearance of a simple declining exponential function. For the Minneapolis - St. Paul calibration (about 100 zones) the level of disaggregation is sufficient for the α to be positive again and for the skewed peaked curve to reappear. All of this reenforces the

Table 1: San Francisco - Work Trip Function Parameters

Income Class	α	β
$ 0 - 4999	0.383	0.900
5000 - 9999	0.750	0.963
10000 - 14999	0.849	0.992
> 15000	0.784	0.990

proposition that the level of spatial aggregation or disaggregation has noticeable effects on the apparent functional forms of these models.

Calibration Results: Complete Estimates for San Francisco

The preliminary estimates of parts of the model were of very little use except that they indicated that a product formulation for W would probably yield better fits than the sum form initially proposed. Consequently, equation (22) was rewritten. First, let

$$W^k_{it} = \left[\pi_g \left(N^g_{it}/\Sigma_g N^g_{it}\right)^{a^k_g}\right]\left[\pi_m \left(X^m_{it}\right)^{a^k_m}\right] \qquad (25)$$

where

N^g_i = the number of type g households in area i

X^m_i = a measure of attribute m of zone i

It was hoped that the attractiveness measure would continue to include intrinsic neighborhood attractiveness as indicated by the household types located there; a measure of "capacity" for development; and a measure of developability in terms of infrastructure. Various attributes were tested, including: residential density, vacant developable land, percentage of developable land developed, and percent industrial (basic) land.

The variables ultimately selected were:

n^k_i = the percentage of the total households in zone i which are of type k

V_i = available developable land in zone i

P_i = percentage of developable land in zone i which has been developed

r_i = residential density (households/acre) in zone i

Thus the form of W used in the final calibrations was (using four household types)

$$W^k_{i,t} = \left[\prod_{g=1}^{4} n^g_{i}\cdot(\exp a^k_g)\right] V_i^{a_5}\; P_i^{a_6}\; r_i^{a_7} \qquad (26)$$

Note that based on the preliminary estimates it was decided to replace the development level factor Q_i by a simple measure of existing level of development, P_i.

Then, rewriting equation (23) we get

$$T^k_{ijt} = E^k_{jt} \left[\frac{W^k_{it}\, f(C_{ijt})}{\sum_i W^k_{it}\, f(C_{ijt})} \right] \qquad (27)$$

Now there are two ways in which the parameters may be estimated. First, the simplest case, is by looking at the activities distribution(s). In this case, by definition:

$$\sum_j T^k_{ijt} = \text{number of households of class k living in i}$$

and thus

$$N^k_{it} = \sum_j E^k_{jt} \left[\frac{W^k_{it}\, f(C_{ijt})}{\sum_i W^k_{it}\, f(C_{ijt})} \right] \qquad (28)$$

Consequently it is possible to estimate the parameters in the W and $f(C)$ functions and this may be called calibration of the aggregated form of the model.

Various authors have, however, asserted that there are disadvantages to calibration of the aggregated form of the model. Their remedy for these problems involves calibration of the disaggregated form of the model given in equation (27). It is an unfortunate fact that in order to calibrate the disaggregated form of the model it is necessary to have a good data source for the T's. In the work described here there were questions as to the quality of these data. If, at some later date, these questions can be satisfactorily resolved along with the development of an acceptable expansion of the San Francisco "sample" to an estimate of the "population", then a calibration of the disaggregated form may be undertaken. In the meantime, calibration has been undertaken for the aggregated form of the model only.

It is immediately obvious that equation (28) cannot be fit to a data set by use of the traditional procedures of linear or even nonlinear multiple regression. In fact the only procedures available are those which, by some hopefully efficient procedure, search for the parameters which produce the best fit of the model to the data. One such procedure is that of gradient search. The use of gradient search involves the following steps:

a) definition of a criterion function to be maximized or minimized
b) definition of the partial derivatives of the criterion function with respect to each of the parameters
c) selection of a starting point (parameters) and calculation of the criterion and the derivatives, hence the gradient, at that point
d) alteration of the parameters as a function of the calculated derivatives and gradient, and iteration through steps c) and d) until a minimum or maximum has been reached

While this may sound like a rather lengthy and difficult undertaking, this is not actually the case. The computer software is somewhat difficult but is available from a variety of sources, including the University of Pennsylvania. It does, at this stage in its development, require experienced staff for its proper use. Nevertheless, once set up, the procedure is rather straightforward and results may be quickly obtained.

The San Francisco data were aggregated to a thirty zone areal system primarily for operating economy in the face of no prior experience as to the costs and difficulties of performing such calibrations. It was felt that the thirty zone system would take less computer time to calibrate while still providing useful information about both the model and the calibration process in general.

The model to be fit is given in equation (28). The distance function is that of equation (24). The variables in the attractiveness measure are the same as were used in equation (25). The calibration was achieved with surprisingly little difficulty. Once the programs were operating correctly there were no significant problems encountered. An interesting point is that a broad, flat ridge in n-space was found where the search program's criterion value, R^2, was somewhat insensitive to parameter variations. This was an expected occurrance, as suggested above (Batty, 1970), nonetheless, with patience, a maximum was reached. The parameters found are shown in Table 2.

There are a number observations to be made regarding these parameter values. Principally, before leaping to unwarranted conclusions, it must be remembered that the household data used in these runs is from the 1960 census, while the land use and employment data are from surveys conducted in San Francisco in 1965. Thus the time subscripts for these variables are not correct for the formulation of the model. The purpose of this particular effort was to explore the problems of calibration of Lowry derivative models via the Wilson entropy approach. That this is a practical procedure has been amply demonstrated.

Table 2: BEST FIT PARAMETERS (EXPONENTS) - DRAM - SAN FRANCISCO (30 ZONES)

	Household Composition				Land Development			Distance		
Household type	a_1^k	a_2^k	a_3^k	a_4^k	a_5	a_6	a_7	α	β	r^2
$ < 5000	1.90	0.40	-0.50	0.33	0.18	-0.73	-0.26	-2.06	0.57	0.91
$5000 - 10000	0.06	1.65	-1.22	0.48	0.27	-1.50	-0.07	-1.75	0.72	0.87
$10000 - 15000	0.14	1.09	-0.26	0.76	0.24	-1.34	-0.14	-1.76	0.76	0.90
$ > 15000	0.72	1.00	-0.34	1.50	0.23	-1.48	-0.04	-1.64	0.48	0.93

Calibration Results: Complete Estimates for Minneapolis - St. Paul

The DRAM model was also calibrated for an available data base for the Minneapolis - St. Paul metropolitan area. This area was divided into 108 zones. The equation form used was also that of equation (27) with the distance function of equation (23). The household income classes differed from those of San Francisco in that they were income quartiles. One of the attractiveness measures, r^i - residential density was replaced by R_i - residential land, which produced better fits. The results of these estimates are shown in Table 3.

The data used in this case are all from approximately 1970, thus resulting in parameter estimates for a static form of the model. It is interesting to note that the scaling or control total procedures, typically used in these models after the allocations are completed, have moved, with the DRAM reformulation, deeper into the workings of the model. Referring to equation (27) it may be seen that the term in brackets on the right-hand side is a proportion. Consequently each E^k_{jt} is simply allocated over all i zones. Consequently the sum of the N^k_i will be equal to the sum of the E^k_j. It was mentioned above that it was necessary to convert the E^k_j from employees of type k to heads of households of type k. If it is assumed that the E^k_j sum to a prespecified regional employment total (or are forced to do so) then the N^k_i can be forced to sum to a regional population total as part of the employee to head of household conversion. This, while still arbitrary, is not so arbitrary as the various forms of scaling procedure typically used in these models, which often involve altering sophisticated model estimates with rather crude prorating procedures and thus vitiating the model results.

Table 3: BEST FIT PARAMETERS (EXPONENTS) - DRAM - MINNEAPOLIS - ST. PAUL (108 ZONES)

	Household Composition				Land Development			Distance		
Household Type	a_1^k	a_2^k	a_3^k	a_4^k	a_5	a_6	a_7	α	β	r^2
First Quartile	0.77	0.14	-0.56	-0.34	-0.03	0.15	0.89	1.04	2.18	0.93
Second Quartile	0.24	0.84	-0.37	-0.15	-0.04	0.18	0.90	2.11	1.46	0.90
Third Quartile	0.09	0.16	0.50	-0.08	-0.08	0.25	0.80	2.81	1.31	0.90
Fourth Quartile	0.13	0.10	-0.19	0.78	-0.03	0.29	0.75	2.10	1.44	0.91

Discussion - Problems of Calibrations

The work described here was originally undertaken simply for the purpose of exploring the possibility of calibrating a Lowry-derivative model with a multivariate attractiveness measure via the Wilson entropy formulation. That this is possible has been amply demonstrated. Nonetheless, problems with the available data, particularly with respect to their time indices, makes interpretations of the substance of the results somewhat chancy.

The general question of parameter interpretation in models of this form is worth discussing. First note that the scale of any of the variables is immaterial since the effect of the balancing factor (equation 11) will be to normalize each variable in all cases. Thus parameters may be interpreted in terms of a variable which ranges from zero to one. Care must be taken to avoid having a variable reach zero if its exponent is negative and checks should be incorporated in both parameter estimation and forecasting programs to, at the least, alert the user to this situation if it should arise.

In Figure 2 several members of the family of curves of the form $y=x^{\alpha}$ are plotted for different values of α. The range in which we are particularly interested is from $x=0$ to $x=1$. Taking first the case of $\alpha \geq 0$, we see that for any α the value of y is < 1. Thus any variable x_i, for which the estimated α is < 1, will have an attenuating effect on the region's share of households in area i. This attenuation gradually diminishes as the value of x_i increases from zero to one. It is important to note that the intuitive expectation of a variable with a positive exponent being an amplifying variable is not quite correct here. In the case of a variable whose range is zero to one, a positive exponent implies decreasing attenuation with increases in the magnitude of the variable to its limit of one.

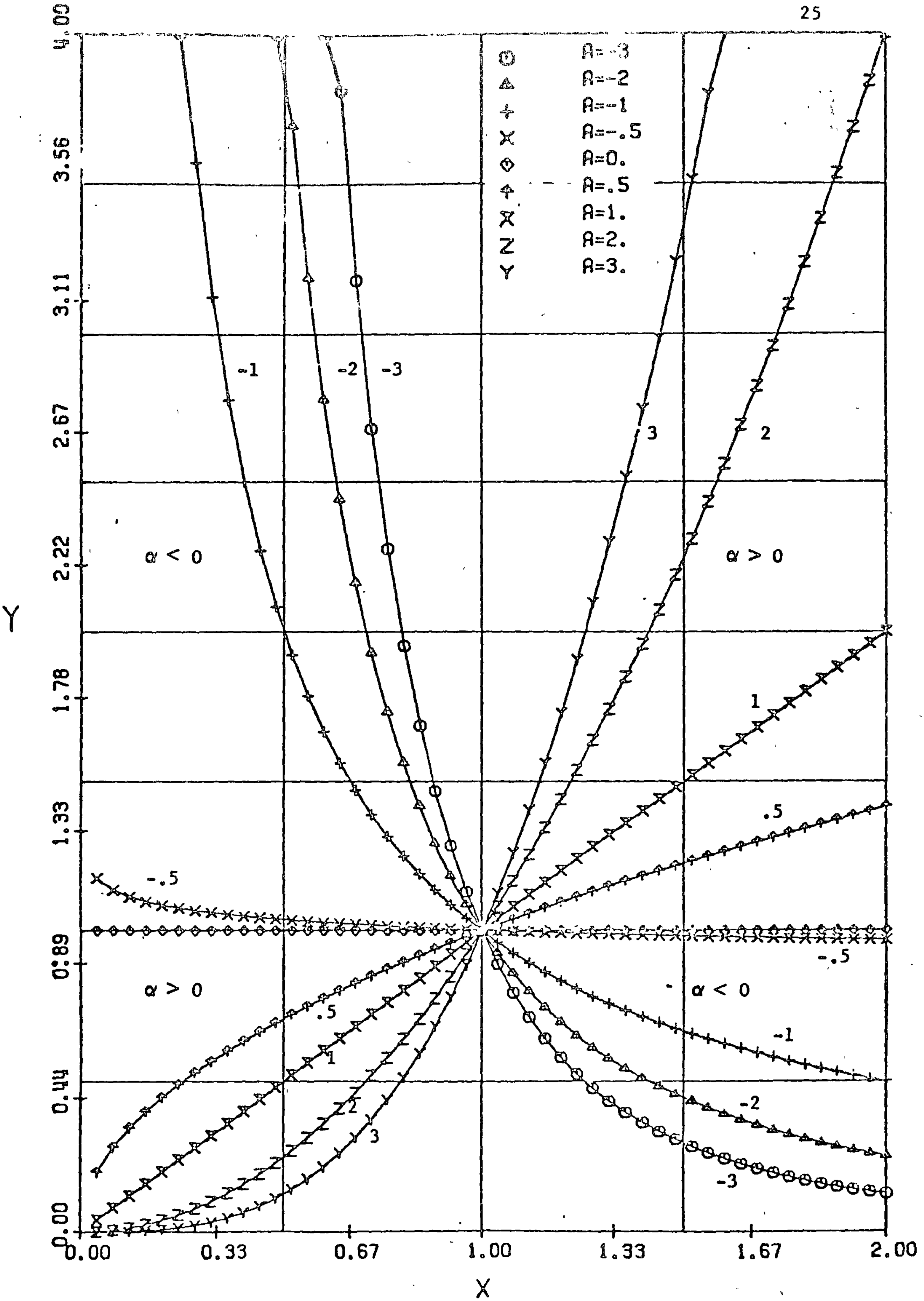


Figure 1: PLOTS OF Y=X expα FOR 0<X<2 AND α= -3,-2,-1,-.5,0,.5,1,2,3

The case of $\alpha < 0$ produces considerable amplification for very small values of x, with the amount of amplification decreasing as x increases to its limit of one. Again, the intuitive notion of a variable with a negative exponent being an attenuating variable is not quite correct. For the case of variables whose range is zero to one, a negative exponent implies decreasin amplification as the variable goes from zero to its limit of one.

In the static situation, thinking about each zone vis-a-vis all other zones makes more sense. For a variable with a positive value of α, all other variables being equal, one would expect greater values of the dependent variable to be found with greater values of the independent variable with a positive α. Similarly lesser values of the dependent variable would be expected to be found with greater values of the independent variable with a negative α. This reasoning also holds for the situation of increases or decreases in the particular independent variable. Nonetheless, it must be remembered that interpretation of the model's parameters does involve the notions of decreasing attenuation producing increases and decreasing amplification producing decreases, and that this is, to a certain degree counter-intuitive.

In this same connection the use of the exponential product form of the model caused some operating difficulties. These are when one or another of the independent variables approaches zero. It may easily be seen in Figure 1 that near zero the function $y = x \exp \alpha$ becomes rather volatile for all non-zero values of α. Consequently the Minneapolis - St. Paul data were rerun with all the independent variables with ranges from 0.0 to 1.0 shifted to the range 1.0 to 2.0 by simply replacing, say P_i, by $(1.0 + P_i)$. These results are shown in Table 4. While there are some noticeable changes in the coefficients compared to the results in Table 3, the overall patterns

Table 4: REVISED BEST FIT PARAMETERS (EXPONENTS) - DRAM - MINNEAPOLIS - ST. PAUL (108 ZONES)

	Household Composition			Land Development				Distance		
Household Type	a_1^k	a_2^k	a_3^k	a_4^k	a_5	a_6	a_7	α	β	r^2
First Quartile	2.92	0.62	-1.71	-1.82	-0.10	0.55	0.83	0.92	2.14	0.89
Second Quartile	1.51	2.04	-1.36	-1.57	-0.06	0.65	0.85	2.24	1.36	0.88
Third Quartile	0.03	0.45	1.06	-0.64	-0.09	0.60	0.87	2.84	1.32	0.89
Fourth Quartile	-0.54	-0.55	-0.06	1.33	-0.07	0.63	0.88	2.48	1.52	0.86

of coefficients are virtually identical. In this form both the problems of instability as the variables approach zero, and of the counter-intuitive operation of the exponents are remedied.

It is very difficult to refrain from speculation as to the substantive implications of the parameters obtained in these estimations. Nonetheless, this would be the wisest policy at this time. We cannot, however, resist the temptation to call attention to the household composition variables and the interesting speculations which the reader may wish to draw therefrom. Two questions are posed here which should be explored during further work with the model. First, with regard to these parameters of the household composition in each zone, is there an apparent preference amongst household types for "equals" or "betters" i.e. higher income classes? Further, if this preference appears is it a preference for the amenities with with which they are associated? Second, having seen how a change in the size of the areal unit changes the shape of the travel function, one wonders at the effect of such a change on the attractiveness portion of the model. To the extent that the household compositions are representative measures of a complex of variables, their meaning may be lost on large areas. The representation, for example, of neighborhood which may show up at a small area level may disappear when the areas are aggregated to larger zones.

Another set of questions which must be resolved during further work with this model has to do with the interaction between the "travel parameters and the "attractiveness" parameters. In these experiments one might first constrain the attractiveness parameters to zero and observe the fit of data to the travel function only, within the construct of the model. Then the reverse could be explored by constraining the travel parameters to zero and observing the fit of data to the attractiveness function only. This information might have been obtained from the independent fitting of the two parts of the model formulation as described above. However, the functions used were not quite correct, nor were the data.

In retrospect it seems that the earlier independent estimation of portions of the model done for the San Francisco data was unnecessary in terms of estimating starting values of parameters for the complete model estimation. The knowledge obtained about the appropriate functional forms to be used in the complete model was, however, a worthwhile output. In future calibration work with this model it will probably be more efficient to begin with the complete model form, while perhaps omitting some of the attractiveness variables or at least constraining their parameter values for the first few runs while initial values of the other parameters are determined. This procedure seemed to work reasonably well for the Minneapolis - St. Paul data. Finally, it should be noted that the use of r^2 as the criterion for parameter fitting is not clearly the best criterion for functional forms like DRAM. The use of maximum likelihood criteria is being investigated for future work.

In conclusion, the initial tests of this model formulation are quite promising. The model appears to be capable of providing direct spatial allocations of households, by several types, without the need for complex input variables or involved sets of constraints and adjustments which are usually found at the tail-end of land use models. At the time of this writing an effort is underway to reevaluate much of the work described here and to produce a more final and definitive form and calibration of DRAM.

REFERENCES

1. Barras, R., et al. (1971) "An Operational Urban Development Model of Cheshire", <u>Environment and Planning</u>, Vol. 3, pp. 115-234.

2. Batty, M. (1970) "Some Problems of Calibrating the Lowry Model", <u>Environment and Planning</u>, Vol. 2, No. 1, pp. 95-114.

3. _____ (1972) "Recent Developments in Land-Use Modelling: A Review of British Research", <u>Urban Studies</u>, Vol. 9, No. 2, pp. 151-177.

4. _____ and S. Mackie (1972) "The Calibration of Gravity, Entropy, and Related Models of Spatial Interaction", <u>Environment and Planning</u>, Vol. 4, pp. 205-233.

5. Crecine, J. (1964) "A Time Oriented Metropolitan Model for Spatial Location", Department of City Planning, Pittsburgh, Pennsylvania.

6. _____ (1969) "Spatial Location Decisions and Urban Structure: A Time-Oriented Model", Discussion Paper No. 4, Institute of Public Policy Studies, University of Michigan, Ann Arbor, Michigan.

7. Cripps, E. and D. Foot (1969) "The Empirical Development of an Elementary Residential Location Model for Use in Sub-Regional Planning", <u>Environment and Planning</u>, Vol. 1, pp. 81-90.

8. Goldner, W. (1971) "The Lowry Model Heritage", Journal of the <u>American Institute of Planners</u>, Vol. 37, pp. 100-110.

9. Hyman, G. (1969) "The Calibration of Trip Distribution Models", <u>Environment and Planning</u>, Vol. 1, pp. 105-112.

10. Lowry, I. (1964) "A Model of Metropolis", RM-4035-RC, RAND Corporation, Santa Monica, California.

11. March, L. (1971) "Urban Systems: A Generalized Distribution Function" in A. G. Wilson, ed. <u>Urban and Regional Planning: London Papers in Regional Science 2</u>, Pion Ltd. London.

12. Putman, S. et. al. (1973) "The Interrelationships of Transportation Development and Land Development", (in two volumes) Department of City and Regional Planning, University of Pennsylvania, Philadelphia, Pa.

13. _____ (1975) "Urban Land Use and Transportation Models: A State-of-the-Art Summary", <u>Transportation Research</u>, Vol. 9, pp. 187-202.

14. Senior, M. (1973) "Approaches to Residential Location Modelling 1: Urban Ecological and Spatial Interaction Models (A Review)", Environment and Planning, Vol. 5, pp. 165-197.

15. Voorhees, A. M. and Associates (1972) "Application of the Urban Systems Model (USM) to a Region-North Central Texas", Prepared for North Central Texas Council of Government.

16. Wilson, A. (1967) "A Statistical Theory of Spatial Distribution Models", Transportation Research, Vol. 1, pp. 253-269.

17. _____ (1970) "Disaggregating Elementary Residential Location Models", Papers of the Regional Science Association, Vol. 14, pp. 103-125.

Appendix II

Tables of Model Outputs

Table A1: TEST RESULTS OF DELETING BASE YEAR LOW INCOME HOUSEHOLDS FROM ZONE (#57) (ALLOWING THEIR RETURN)

	Results of DRAM Package Runs			Results of EMPIRIC Package Runs		
Sector	Test Run	Control Run	Percent Difference	Test Run	Control Run	Percent Difference
Low Income	14	1123	- 98.8	368	368	+0.0
Low-Middle Income	52	138	- 62.3	0	0	+0.0
High-Middle Income	165	66	+150.0	0	0	+0.0
High Income	52	12	+333.3	0	0	+0.0
Total Population	283	1339	- 78.9	368	368	+0.0
Basic 1	4153	4153	+ 0.0	708	708	+0.0
Basic 2	1315	1315	+ 0.0	3627	3628	+0.0
Non-Basic	21714	12714	+ 0.0	3609	3610	+0.0
Total Employment	18183	18182	+ 0.0	7944	7946	+0.0

Table A2: TEST RESULTS OF DELETING BASE YEAR LOW INCOME HOUSEHOLDS FROM ZONE (#57) (NOT ALLOWING RETURN)

	Results of DRAM Package Runs		
Sector	Test Run	Control Run	Percent Difference
Low Income	0	682	** *
Low-Middle Income	98	213	- 54.0
High-Middle Income	112	87	+ 28.7
High Income	72	54	+ 33.3
Total Population	282	1036	- 72.8
Basic 1	4153	4153	0.0
Basic 2	1315	1315	0.0
Non-Basic	12725	12725	0.0
Total Employment	18193	18193	0.0

Table A3: TEST RESULTS OF DELETING BASE YEAR LOW INCOME HOUSEHOLDS FROM ZONE (#49) (ALLOWING THEIR RETURN)

	Results of DRAM Package Runs		
Sector	Test Run	Control Run	Percent Difference
Low Income	928	1046	- 11.3
Low-Middle Income	2123	2377	- 10.7
High-Middle Income	3866	3908	- 1.1
High Income	4267	4071	+ 4.8
Total Population	11184	11402	- 1.9
Basic 1	3258	3258	0.0
Basic 2	667	667	0.0
Non-Basic	1981	1983	- 0.1
Total Employment	5906	5908	0.0

Table A4: TEST RESULTS OF DELETING BASE YEAR LOW INCOME HOUSEHOLDS FROM ZONE (#49)
(NOT ALLOWING RETURN)

	Results of DRAM Package Runs		
Sector	Test Run	Control Run	Percent Difference
Low Income	0	1046	- **
Low-Middle Income	1831	2377	- 22.9
High-Middle Income	3915	3908	- 0.2
High Income	4529	4071	+ 10.1
Total Population	10275	11402	- 9.9
Basic 1	3258	3258	0.0
Basic 2	667	667	0.0
Non-Basic	1974	1983	- 0.5
Total Employment	5899	5908	- 0.2

Table A5: TEST RESULTS OF DELETING BASE YEAR LOW INCOME HOUSEHOLDS FROM ZONE (#56) (ALLOWING THEIR RETURN)

	Results of DRAM Package Runs		
Sector	Test Run	Control Run	Percent Difference
Low Income	8884	9836	- 9.7
Low-Middle Income	4994	5342	- 6.5
High-Middle Income	3089	3080	+ 0.3
High Income	2852	2771	+ 2.9
Total Population	30887	30737	+ 0.5
Basic 1	1822	1822	0.0
Basic 2	1237	1237	0.0
Non-Basic	28543	28564	- 0.1
Total Employment	31602	31623	- 0.1

Table A6: TEST RESULTS OF DELETING BASE YEAR LOW INCOME HOUSEHOLDS AND ADDING 1000 UPPER INCOME HOUSEHOLDS TO ZONE (#56)

	Results of DRAM Package Runs		
Sector	Test Run	Control Run	Percent Difference
Low Income	8121	9836	- 17.4
Low-Middle Income	4617	5342	- 13.6
High-Middle Income	3040	3080	- 1.3
High Income	2978	2771	+ 7.5
Total Population	18756	21029	- 10.8
Basic 1	1822	1822	0.0
Basic 2	1237	1237	0.0
Non-Basic	28528	28564	- 0.1
Total Employment	31587	31623	- 0.1

Table A7: TEST RESULTS OF ADDING 1000 LOW INCOME HOUSEHOLDS TO ZONE 27 IN THE BASE YEAR

	Results of DRAM Package Runs			Results of EMPIRIC Package Runs		
Sector	Test Run	Control Run	Percent Difference	Test Run	Control Run	Percent Difference
Low Income	2174	1179	+ 84.4	795	797	+ 0.0
Low-Middle Income	2583	2058	+ 25.5	1756	1759	- 0.0
High-Middle Income	3030	3113	- 2.7	2210	2212	- 0.0
High Income	3175	3339	- 4.9	1254	1253	+ 0.0
Total Population	10962	9689	+ 13.1	6017	6019	- 0.0
Basic 1	687	687	+ 0.0	467	467	+ 0.0
Basic 2	228	228	+ 0.0	166	164	+ 1.2
Non-Basic	2696	2687	+ 0.3	3435	3436	- 0.0
Total Employment	3611	3602	+ 0.2	4068	4067	+ 0.0

Table A8: TEST RESULTS OF ADDING 1000 LOW INCOME HOUSEHOLDS TO ZONES 33 AND 32 IN THE BASE YEAR

	Results of DRAM Package Runs (Zone 33)			Results of DRAM Package Runs (Zone 32)		
Sector	Test Run	Control Run	Percent Difference	Test Run	Control Run	Percent Difference
Low Income	1038	654	+ 58.7	816	518	+ 57.5
Low-Middle Income	2333	1594	+ 46.4	2019	1381	+ 46.2
High-Middle Income	2302	2284	+ 0.8	2048	2041	+ 0.3
High Income	1746	2028	- 13.9	1536	1774	- 13.4
Total Population	7419	6560	+ 13.1	6419	5714	+ 12.3
Basic 1	968	968	0.0	1379	1379	0.0
Basic 2	213	213	0.0	421	421	0.0
Non-Basic	1728	1719	+ 0.5	671	668	+ 0.4
Total Employment	2909	2900	+ 0.3	2471	2468	+ 0.1

Table A9: TEST RESULTS OF ADDING 1000 LOW INCOME HOUSEHOLDS TO ZONE 94 IN THE BASE YEAR

	Results of DRAM Package Runs		
Sector	Test Run	Control Run	Percent Difference
Low Income	326	134	+143.0
Low-Middle Income	472	354	+ 33.3
High-Middle Income	560	616	- 9.1
High Income	172	212	- 18.9
Total Population	1530	1316	+ 16.3
Basic 1	840	840	+ 0.0
Basic 2	160	160	+ 0.0
Non-Basic	374	373	+ 0.3
Total Employment	1374	1373	+ 0.1

Table A10: TEST RESULTS OF ADDING 1000 HIGH INCOME HOUSEHOLDS TO ZONE 74 AND 1000 HIGH INCOME PLUS 1000 UPPER MIDDLE INCOME HOUSEHOLDS TO ZONE 94 IN THE BASE YEAR

	Results of DRAM Package Runs (Zone 74)			Results of DRAM Package Runs (Zone 94)		
Sector	Test Run	Control Run	Percent Difference	Test Run	Control Run	Percent Difference
Low Income	3596	3839	- 6.3	77	133	- 42.1
Low-Middle Income	3367	3605	- 6.6	222	405	- 45.2
High-Middle Income	4002	4168	- 4.0	508	551	- 7.8
High Income	6756	6407	+ 5.4	384	297	+ 29.3
Total Population	17721	18019	- 1.7	1191	1386	- 14.1
Basic 1	5789	5789	0.0	840	840	0.0
Basic 2	5180	5180	0.0	160	160	0.0
Non-Basic	2023	2022	0.0	374	376	- 0.5
Total Employment	18071	18075	0.0	1374	1376	- 0.1

Table A11: TEST RESULTS OF A 5% ACCESS IMPROVEMENT FOR AN URBAN FRINGE ZONE (#24)

	Results of DRAM Package Runs			Results of EMPIRIC Package Runs		
Sector	Test Run	Control Run	Percent Difference	Test Run	Control Run	Percent Difference
Low Income	160	143	+ 11.9	385	386	- 0.3
Low-Middle Income	626	547	+ 14.4	600	601	- 0.2
High-Middle Income	739	672	+ 10.0	698	698	0.0
High Income	352	294	+ 19.7	650	649	0.2
Total Population	1877	1656	+ 13.3	2333	2334	0.0
Basic 1	1042	993	+ 4.9	272	271	+ 0.4
Basic 2	58	106	- 45.3	0	0	0.0
Non-Basic	152	139	+ 9.4	1072	1066	+ 0.6
Total Employment	1252	1238	+ 1.1	1344	1337	+ 0.5

Table A12: TEST RESULTS OF A 20% ACCESS IMPROVEMENT FOR AN URBAN FRINGE ZONE (#24)

	Results of DRAM Package Runs			Results of EMPIRIC Package Runs		
Sector	Test Run	Control Run	Percent Difference	Test Run	Control Run	Percent Difference
Low Income	323	143	+125.9	383	386	- 0.8
Low-Middle Income	942	547	+ 72.2	599	601	- 0.3
High-Middle Income	928	672	+ 38.1	698	698	+ 0.0
High Income	588	294	+100.0	651	649	+ 0.3
Total Population	2781	1656	+ 67.9	2331	2334	- 0.1
Basic 1	1020	993	+ 2.7	273	271	+ 0.7
Basic 2	115	106	+ 8.5	23	0	+ ***
Non-Basic	202	139	+ 45.3	1091	1066	+ 2.3
Total Employment	1337	1238	+ 8.0	1387	1337	+ 3.7

Table A13: TEST RESULTS OF A 5% ACCESS IMPROVEMENT FOR AN URBAN CORE ZONE (#64)

	Results of DRAM Package Runs			Results of EMPIRIC Package Runs		
Sector	Test Run	Control Run	Percent Difference	Test Run	Control Run	Percent Difference
Low Income	22726	18806	+ 20.8	10107	10307	- 1.9
Low-Middle Income	12113	11874	+ 2.0	9530	9545	- 0.2
High-Middle Income	5244	5485	- 4.4	5148	5035	+ 2.2
High Income	4127	4063	+ 1.6	2855	2682	+ 6.5
Total Population	44210	40228	+ 9.9	27640	27569	+ 0.3
Basic 1	1762	1090	+ 61.7	1936	1907	+ 1.5
Basic 2	3090	3685	- 16.1	4797	4578	+ 4.8
Non-Basic	24450	21798	+ 12.2	22502	22251	+ 1.1
Total Employment	29302	26573	+ 10.3	29235	28736	+ 1.7

Table A14: TEST RESULTS OF A 20% ACCESS IMPROVEMENT FOR AN URBAN CORE ZONE (#64)

	Results of DRAM Package Runs			Results of EMPIRIC Package Runs		
Sector	Test Run	Control Run	Percent Difference	Test Run	Control Run	Percent Difference
Low Income	39741	18806	+111.3	9502	10307	- 7.8
Low-Middle Income	12419	11874	+ 4.6	9486	9545	- 0.6
High-Middle Income	4305	5485	- 21.5	5488	5035	+ 9.0
High Income	4130	4063	+ 1.6	3368	2682	+ 25.6
Total Population	60595	40228	+ 50.6	3474	27569	+ 1.0
Basic 1	3298	1090	+202.6	2026	1907	+ 6.2
Basic 2	1615	3685	- 56.2	5430	4578	+ 18.6
Non-Basic	30454	21798	+ 39.7	23256	22251	+ 4.5
Total Employment	35367	26573	+ 33.1	30712	28736	+ 6.9

Table A15: TEST RESULTS OF A 10% BASIC EMPLOYMENT INCREASE IN AN URBAN FRINGE ZONE (#24)

	Results of DRAM Package Runs			Results of EMPIRIC Package Runs		
Sector	Test Run	Control Run	Percent Difference	Test Run	Control Run	Percent Difference
Low Income	145	143	+ 1.4	386	386	+ 0.0
Low-Middle Income	547	547	+ 0.0	601	601	+ 0.0
High-Middle Income	672	672	+ 0.0	698	698	+ 0.0
High Income	294	294	+ 0.0	649	649	+ 0.0
Total Population	1658	1656	+ 0.1	2334	2334	+ 0.0
Basic 1	1031	993	+ 3.8	271	271	+ 0.0
Basic 2	109	106	+ 2.8	0.0	0.0	+ 0.0
Non-Basic	139	139	+ 0.0	1064.	1066.	- 0.2
Total Employment	1279	1238	+ 3.3	1335	1337	- 0.1

Table A16: TEST RESULTS OF A 30% BASIC EMPLOYMENT INCREASE IN AN URBAN FRINGE ZONE (#24)

	Results of DRAM Package Runs			Results of EMPIRIC Package Runs		
Sector	Test Run	Control Run	Percent Difference	Test Run	Control Run	Percent Difference
Low Income	151	143	+ 5.6	386	386	+ 0.0
Low-Middle Income	549	547	+ 0.4	601	601	+ 0.0
High-Middle Income	672	672	+ 0.0	698	698	+ 0.0
High Income	294	294	+ 0.0	649	649	+ 0.0
Total Population	1666	1656	+ 0.6	2334	2334	+ 0.0
Basic 1	1106	993	+ 11.4	271	271	+ 0.0
Basic 2	114	106	+ 7.5	17	0	+ ***
Non-Basic	139	139	+ 0.0	1060	1066	- 0.6
Total Employment	1359	1238	+ 10.0	1348	1337	+ 0.8

Table A17: TEST RESULTS OF A 10% BASIC EMPLOYMENT INCREASE IN AN URBAN CORE ZONE (#64)

	Results of DRAM Package Runs			Results of EMPIRIC Package Runs		
Sector	Test Run	Control Run	Percent Difference	Test Run	Control Run	Percent Difference
Low Income	18831	18806	+ 0.1	10307	10307	+ 0.0
Low-Middle Income	11873	11874	+ 0.0	9545	9545	+ 0.0
High-Middle Income	5481	5485	- 0.1	5035	5035	+ 0.0
High Income	4062	4063	+ 0.0	2682	2682	+ 0.0
Total Population	40247	40228	+ 0.0	27569	27569	+ 0.0
Basic 1	1238	1090	+ 13.6	1903	1907	- 0.2
Basic 2	4141	3685	+ 12.4	5236	4578	+ 14.4
Non-Basic	21809	21798	+ 0.1	22186	22251	- 0.3
Total Employment	27188	26573	+ 2.3	29325	28736	+ 2.0

Table A18: TEST RESULTS OF A 30% BASIC EMPLOYMENT INCREASE IN AN URBAN CORE ZONE (#64)

	Results of DRAM Package Runs			Results of EMPIRIC Package Runs		
Sector	Test Run	Control Run	Percent Difference	Test Run	Control Run	Percent Difference
Low Income	18880	18806	+ 0.4	10307	10307	+ 0.0
Low-Middle Income	11870	11874	+ 0.0	9545	9545	+ 0.0
High-Middle Income	5473	5485	- 0.2	5035	5035	+ 0.0
High Income	4061	4063	- 0.0	2682	2682	+ 0.0
Total Population	40284	40228	+ 0.1	27569	27569	+ 0.0
Basic 1	1539	1090	+ 41.2	1896	1907	- 0.6
Basic 2	5041	3685	+ 36.8	6538	4578	+ 42.8
Non-Basic	21830	21798	+ 0.1	22059	22251	- 0.9
Total Employment	28410	26573	+ 6.9	30493	28736	+ 6.1

Table A19: TEST RESULTS OF A 10% NON-BASIC EMPLOYMENT INCREASE IN AN URBAN FRINGE ZONE (#24)

	Results of DRAM Package Runs			Results of EMPIRIC Package Runs		
Sector	Test Run	Control Run	Percent Difference	Test Run	Control Run	Percent Difference
Low Income	145	143	+ 1.4	386	386	+ 0.0
Low-Middle Income	548	547	+ 0.2	601	601	+ 0.0
High-Middle Income	672	672	+ 0.0	698	698	+ 0.0
High Income	294	294	+ 0.0	649	649	+ 0.0
Total Population	1659	1656	+ 0.2	2334	2334	+ 0.0
Basic 1	993	993	+ 0.0	275	271	+ 1.5
Basic 2	106	106	+ 0.0	0	0	+ 0.0
Non-Basic	152	139	+ 9.4	835	1066	- 21.7
Total Employment	1251	1238	+ 1.1	1110	1337	- 17.0

Table A20: TEST RESULTS OF A 30% NON-BASIC EMPLOYMENT INCREASE IN AN URBAN FRINGE ZONE (#24)

	Results of DRAM Package Runs			Results of EMPIRIC Package Runs		
Sector	Test Run	Control Run	Percent Difference	Test Run	Control Run	Percent Difference
Low Income	150	143	+ 4.9	386	386	+ 0.0
Low-Middle Income	549	547	+ 0.4	601	601	+ 0.0
High-Middle Income	671	672	- 0.0	698	698	+ 0.0
High Income	294	294	+ 0.0	649	649	+ 0.0
Total Population	1664	1656	+ 0.4	2334	2334	+ 0.0
Basic 1	993	993	+ 0.0	270	271	- 0.4
Basic 2	106	106	+ 0.0	12	0	***
Non-Basic	180	139	+ 29.5	1184	1066	+ 11.1
Total Employment	1279	1238	+ 3.3	1466	1337	+ 9.6

Table A21: TEST RESULTS OF A 10% NON-BASIC EMPLOYMENT INCREASE IN AN URBAN CORE ZONE (#64)

	Results of DRAM Package Runs			Results of EMPIRIC Package Runs		
Sector	Test Run	Control Run	Percent Difference	Test Run	Control Run	Percent Difference
Low Income	18958	18806	+ 0.8	10307	10307	+ 0.0
Low-Middle Income	11867	11874	- 0.1	9545	9545	+ 0.0
High-Middle Income	5472	5485	- 0.2	5035	5035	+ 0.0
High Income	4056	4063	- 0.2	2682	2682	+ 0.0
Total Population	40353	40228	+ 0.3	27569	27569	+ 0.0
Basic 1	1090	1090	+ 0.0	1991	1907	+ 4.4
Basic 2	3685	3685	+ 0.0	1684	4578	- 63.2
Non-Basic	23866	21798	+ 9.5	16074	22251	- 27.8
Total Employment	28641	26573	+ 7.8	19749	28736	- 31.3

Table A22: TEST RESULTS OF A 30% NON-BASIC EMPLOYMENT INCREASE IN AN URBAN CORE ZONE (#64)

	Results of DRAM Package Runs			Results of EMPIRIC Package Runs		
Sector	Test Run	Control Run	Percent Difference	Test Run	Control Run	Percent Difference
Low Income	19258	18806	+ 2.4	10307	10307	+ 0.0
Low-Middle Income	11853	11874	- 0.2	9545	9545	+ 0.0
High-Middle Income	5445	5485	- 0.7	5035	5035	+ 0.0
High Income	4043	4063	- 0.5	2682	2682	+ 0.0
Total Population	40599	40228	+ 0.9	27569	27569	.0
Basic 1	1090	1090	+ 0.0	1864	1907	- 2.3
Basic 2	3685	3685	+ 0.0	5971	4578	+ 30.4
Non-Basic	27943	21798	+ 28.2	25289	22251	+ 13.7
Total Employment	32718	26573	+ 23.1	33124	28736	+ 15.3

Environment and Planning A, 1977, volume 9, pages 449-460

Calibrating a residential location model for nineteenth-century Philadelphia

S H Putman
University of Pennsylvania, Philadelphia, Pa 19174, USA
Received 30 September 1976 in revised form 10 November 1976

Abstract The Disaggregated Residential Allocation Model (DRAM) is a modified Lowry derivative model which follows the Wilson entropy formulation. The model is a component of the Integrated Transportation and Land Use Package—ITLUP As a part of ongoing work with DRAM, it has been calibrated for several US cities by use of conventional late twentieth-century data sets Because of the work of the Philadelphia Social History Project a nineteenth-century data set for Philadelphia became available for use This provided a unique opportunity to try calibrations of DRAM on this century-old data and thus attempt to validate the general structure of the model The results of these calibrations suggest a rather high degree of descriptive validity and consistency with macro behavioral theories of spatial allocation

Introduction

Data inadequacies are a perennial problem in urban simulation studies Data sets that are adequate for estimating a model's parameters are often difficult to assemble It is virtually never possible to assemble a data set adequate for the task of validating a model's performance. The data used in the work described here are a unique resource and are an exception to the rule They are the result of years of effort by the Philadelphia Social History Project at the University of Pennsylvania (Hershberg, 1976) The resulting files contain a relatively consistent and comprehensive data base for the years 1850, 1860, 1870, and 1880 for the city of Philadelphia.

The model and parameter estimation procedures used in this work are part of the program library of the Planning Sciences Simulation Laboratory of the Department of City and Regional Planning at the University of Pennsylvania. They, too, represent the results of years of effort on model development, collection, documentation, and use

The conjunction of these two resources to estimate the parameters of an urban simulation model for nineteenth-century Philadelphia has produced striking evidence of the generality and descriptive validity of the model

A description of the model

The model used in this research is a version of the author's Disaggregated Residential Allocation Model—DRAM (Putman, 1977). The DRAM model is a sophisticated variation on the basic Lowry model theme (Lowry, 1964) The hypotheses of the Lowry model assert that, given a spatial distribution of employment and a description of zone-to-zone travel times (or costs), it is possible to estimate the location of the employees' residences This location is taken to be a result of trip length probabilities and, in the more complex variants of the model, of residence zone characteristics This may be written in equation form as

$$P_i = \sum_j E_j p_{ij} W_i \, , \qquad (1)$$

P_i is the population of zone i
E_j is employment in zone j,
p_{ij} is the probability of a work-trip as long as the travel time (or cost) between zones i and j,
W_i is a measure of the residential attractiveness of zone i

The important differences between the Lowry variants result from different functional forms to generate p_{ij} from travel times (or costs) and from different ways of measuring W_i. Further differences stem from structuring the model in static or dynamic form

The Lowry model may be rewritten as

$$T_{ij} = E_j \, f(c_{ij}) \, , \tag{2}$$

where

T_{ij} is the number of persons working in zone j and residing in zone i (the ij work trips),

c_{ij} is the impedance (usually travel time or travel cost) between centroids of zone i and zone j,

f is the impedance function.

An important problem of this formulation is that there is no constraint on the sums of trips Without the constraint there is no reason to expect that

$$\sum_i T_{ij} = E_j \tag{3}$$

This implies that the number of employees in zone j will not equal the sum of the employees residing in all zones i who claim to work in zone j

A simple and consistent residential location model may be derived from entropy-maximizing concepts (Wilson, 1970) as follows

$$T_{ij} = A_i B_j O_i E_j \, f(c_{ij}) \tag{4}$$

where

O_i are trip origins or employed persons living in zone i.

E_j is employment in zone j as represented by trip destinations

A_i is a balancing factor for trip origins,

B_j is a balancing factor for trip destinations.

If the trip origins O_i are replaced by a measure of attractiveness of the origin zone, W_i, the origins' balancing factor A_i is eliminated to give

$$T_{ij} = B_j W_i E_j \, f(c_{ij}) \tag{5}$$

In order for the constraint on the sums of trip destinations, equation (3), to be met, we have

$$B_j = \left[\sum_i W_i \, f(c_{ij}) \right]^{-1} \tag{6}$$

It is informative to substitute this expression back into the original equation, which yields

$$T_{ij} = E_j \left[\frac{W_i \, f(c_{ij})}{\sum_i W_i \, f(c_{ij})} \right] \tag{7}$$

If the term $W_i \, f(c_{ij})$ is called an 'accessibility-attractiveness' measure, then the fraction in equation (7) is a relative measure of the accessibility-attractiveness of zone i to zone j compared to all other zones i Further, it is clear that the total number of employed residents, N, residing in zone i is

$$N_i = \sum_j T_{ij} \, , \tag{8}$$

and, by substituting from equation (7),

$$N_i = \sum_j E_j \left[\frac{W_i \, f(c_{ij})}{\sum_i W_i \, f(c_{ij})} \right] \, . \tag{9}$$

Prior to the work described here DRAM had been fitted to several 1960–1970 data sets for US cities In these applications the N_i were disaggregated to four types, by income quartile The nineteenth-century Philadelphia data were not so readily disaggregated Thus, for these tests, the model was used in the homogeneous population form of equation (9) The 1960–1970 forms of DRAM used a multivariate attractiveness measure Typically this measure contained the household composition of the zone along with land availability and degree of development Thus the attractiveness measure actually contained the product of a surrogate measure of the zone's amenities, in the form of the household distribution, and the quantity of land available for development This was, in turn, weighted by its developability and/or its existing level of development. In this first work with the nineteenth-century data the attractiveness measure was reduced to, simply, the zonal area

In the 1960–1970 forms of DRAM a modified gamma function, known as Tanner's function, is used as the trip probability function This distribution may be said to represent the product of (1) exponentially increasing numbers of opportunities (for trip satisfaction, that is, residence location) encountered with increasing distance travelled and (2) exponentially decreasing propensity to travel each additional unit of distance Thus the function takes the form

$$p_{ij} = D_{ij}^{\alpha} \exp(\beta D_{ij}) , \qquad (10)$$

where

p_{ij} is the probability of a trip of length D_{ij},

D_{ij} is the travel time between zone i and zone j,

α, β are empirically derived parameters

This same form of trip function was used in these nineteenth-century analyses

Thus the form of DRAM used in this research was

$$N_i = \sum_j L_j \left[\frac{L_i^{\gamma} D_{ij}^{\alpha} \exp(\beta D_{ij})}{\sum\limits_i L_i^{\gamma} D_{ij}^{\alpha} \exp(\beta D_{ij})} \right] , \qquad (11)$$

where all terms have been previously defined except L_i, which is the area of zone i and γ, which is an empirically derived parameter Having selected a form of model with which to work, the next task was to create the appropriate data files to run the tests

Data base description

The data used for this project were constructed from the files of the Philadelphia Social History Project This data base is derived from several sources The population data are from the manuscript schedules of the US Population Census for 1850, 1860, 1870, and 1880 These are augmented by special surveys of the Pennsylvania Abolition Society and the Society of Friends manuscript census schedules of 1838, 1847, and 1856 Employment data came from the US Census of Manufacturing manuscript schedules for 1850, 1860, 1870, and 1880 Further employment information was derived from Philadelphia Business Directories of the era There is also a file of data regarding the transportation systems of the era which was not available at the time of the work reported here

All these data files were keyed to a grid system which divides Philadelphia (defined in terms of its present boundaries) into 7100 grid cells each containing 1 25 city blocks. For the calibration (parameter estimations) reported here this areal system was aggregated to a system of 105 zones These zones were designed to try to develop more or less homogeneous population densities of the grid cells included and always to involve integral aggregation (that is, no splitting) of grids Once the 105-zone system was devised the area of each zone was calculated and $x-y$ coordinates were assigned to the approximate location of the zone centroids By means of analytic geometry the 105 × 105-zone distance matrix was then calculated using the straight line distances between zone centroids. The completed 105-zone data files were then merged into a single master file, plus the distances matrix, which was used for all the calibrations described below It should be noted that this was the simplest data set sufficient for these calibrations. It is hoped that at some future date various disaggregations can be made and the model recalibrated and tested further In the work that is described next only total population, total males employed in manufacturing, zonal area, and the distance matrix were used for each of the four time points

Results of parameter estimation

Estimation of parameters for models such as this has become a relatively straightforward procedure. A mathematical search technique is used to find the set of numerical values of γ, α, and β from equation (11) which minimize the discrepancy between the observed N_i data and the model's estimates of $\hat{N}_i$

In this research the search procedure used was that of gradient search In gradient search the gradient (an n-dimensional vector orthogonal to the mathematical surface, whose projection on that surface points in the direction of steepest ascent) is evaluated at a given point, and an n-dimensional step is taken in the direction indicated by the projection of the gradient The gradient may be calculated by use of numerical approximations or by analytically solving the function's partial derivatives with respect to the criterion being minimized For these parameter estimations a maximum likelihood criterion was used The value of R^2 was also calculated and is presented here as a more readily interpreted measure of the model's data fits.

It is worth noting that in previous work with DRAM it was found that the use of a maximum-likelihood criterion rather than an R^2 criterion for the parameter estimation yielded a more sharply defined optimum This considerably ameliorated the problem of the criterion being insensitive to parameter changes when in the neighborhood of the optimum or best fit to the data The use of the maximum-likelihood criterion is also desirable on theoretical grounds (Batty and Mackie, 1972) Further, it should be noted that experiments were run with different starting points for the search procedure. The results of these runs indicate the absence of local optima in that, regardless of starting points, all runs with a given data set converged on the identical parameter estimates.

The parameter estimation (calibration) results are shown in table 1 It is in the interpretation of the α and β parameters of the travel function that the most interesting aspect of this work is to be found The role of the γ parameter is simply to make the zonal area L_i account, in part, for the number of employed residents in the zone.

Table 1 Parameter estimates nineteenth-century DRAM

Date	1850	1860	1870	1880
γ, land area	0·4644	0·8900	0·7200	0·9705
α, opportunities	−2·2779	3·0877	1·0552	1·9201
β, travel propensity	0·0576	−5·0533	−2·7692	−3·4201
Distance at trip function maximum	na	0·61	0·38	0·56
Distance at 0·5 cumulative trip probability	<0·05	0·74	0·62	0·75
R^2, measure of fit, but not the criterion	0·5572	0·8168	0·7315	0·7303

It is instructive to recall that the first term of the trip probability function [equation (10)] accounts for opportunities as a function of distance travelled (or generalized travel cost incurred) from the trip origin. Thus a positive value of α implies that the opportunities encountered will increase with increasing distance (cost), whereas a negative value of α implies that the opportunities encountered will decrease with increasing distance from the trip origin. In previous calibration work with DRAM, values of α have been estimated for population (disaggregated to income quartiles) in Minneapolis, San Francisco, and Washington, DC In all cases, at the level of areal detail used here, α has been positive.

For 1850 we find a negative value of α. This suggests that the developed area of the city and of the small towns scattered about the area, that now comprises the city were rather compact. Consequently a trip originating in a developed zone would in fact, encounter rapidly declining opportunities within a small distance from that origin. A spatial pattern of this type is, in essence, a property of the urban form which urban historians have termed the 'walking' city.

For 1860, a decade later, the data yield a rather large positive value of α This suggests that large numbers of opportunities would be encountered near, but not at the trip origin. The inference here is that each of the dense clusters of development has produced a ring of further development just slightly removed from the core The trips made would, however, remain relatively short, probably not very different in length from those of 1850

In 1870 the value of α is again positive, but smaller than it was in 1860 This suggests that a significant spread of opportunities, beyond the 1860 ring, may have taken place by 1870 Finally, in 1880 the value of α is positive and approaching its 1860 value This may represent the development of a second ring of opportunities It may also indicate a 'filling in' of the more dispersed but less dense pattern of 1870 Alternatively the sequence of α values could be taken to represent ripples of development moving outward from the core of each of the developed clusters in the region The shapes of the opportunities function for each of the four points in time are shown in figure 1

The second term of the trip probability function [equation (10)] accounts for the individual's propensity to travel varying distances from the trip origin Thus a positive value of β would imply a preference for increasingly long trips, whereas a negative value of β would imply a decreasing preference for increasingly longer trips In prior

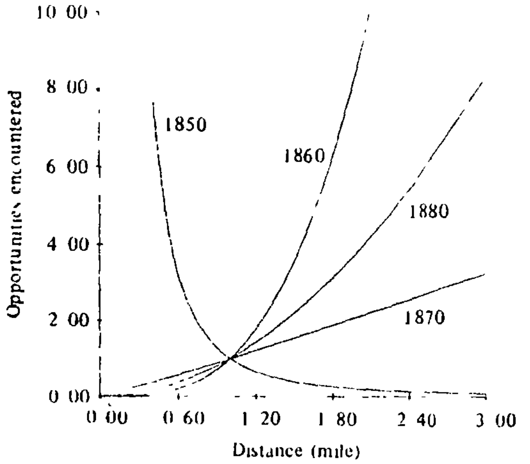


Figure 1. Opportunity function Philadelphia 1850-1880

calibration work with DRAM, values of β have also been estimated for populations, disaggregated by income quartiles, in Minneapolis, San Francisco, and Washington, DC In all cases, at every level of areal detail tested, the values found for β have been negative

The data for Philadelphia in 1850 yield a value of +0 0576 for β This result suggests that in 1850 the tripmakers (note that we are concerned primarily with the work trip in these discussions) were virtually indifferent to trip length This seems to be a rather unrealistic finding until the notion of the 'walking city' is recalled It is likely, as will be discussed in the next section of this paper that even the longer work trips were short enough for tripmakers to be indifferent to whatever minor variations in trip length there may have been.

By 1860, just one decade later, there must have been the possibility of taking longer work trips. For this data set, a value of $-5{\cdot}0533$ was found for β This value yields a steeply declining exponential curve of travel propensity as a function of increasing distance. Such a curve suggests that by 1860 longer work trips were possible than were possible in 1850. Stated in other words, 1860 must have seen the beginning of a trend towards increasing separation of workplace from residence At the same time, the curve also suggests that most tripmakers were strongly disinclined to make trips noticeably longer than the ones being made a decade earlier

In 1870 and again in 1880 we find negative values of β, yielding steeply declining exponential curves of travel propensity. Interestingly enough, the curve for 1870 declines less sharply than either that for 1860 or for 1880 This result suggests that tripmakers were somewhat more willing to take longer trips in 1870 than they were either in 1860 or in 1880. An alternative explanation, as will be discussed in the next section of this paper, is that there was a significant discontinuity of some sort which took place in Philadelphia between 1860 and 1870. This discontinuity may have shifted the travel propensity function much the same way as a demand curve (in economic analysis) may sometimes be shifted by some market or historical discontinuity The 1880 travel propensity function might then be seen as a movement back towards the 1860 curve from the 1870 displacement The travel propensity curves for each of the four points in time may be seen in figure 2

The trip probability function, the product of the opportunities term and the travel propensity term, can assume, under normal circumstances, either one of two shapes

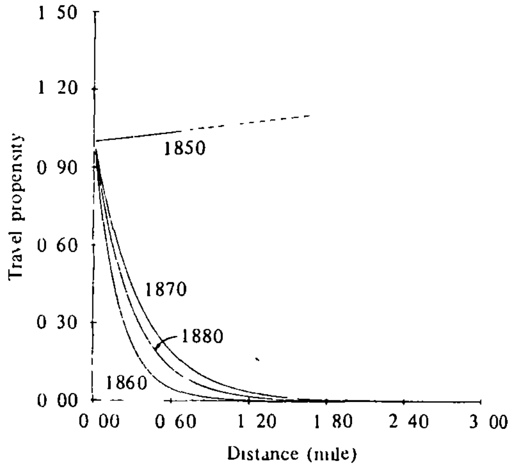


Figure 2. Travel function Philadelphia 1850-1880

If the values of α and β are both negative, the trip function takes the form of a declining exponential function. If the value of α is positive and that of β is negative, the function looks like a normal distribution skewed sharply to the left.

In prior work with DRAM, when 1960 or 1970 data sets were used, the value of β has always been found to be negative. This is consistent both with theory and with intuition, saying that as far as work trips are concerned, virtually everyone prefers the shortest possible trip. The value of α has been negative in only one instance, where the nine-county San Francisco Bay area was analyzed at a thirty-zone level of detail. At this coarse level of detail the opportunities phenomenon was apparently swamped by the travel propensity, thus forcing the negative value of α necessary for the trip probability function to have the negative exponential shape. When this same data set was disaggregated to one-hundred, and later to three-hundred, zones the value of α became positive. In defining the areal system used in the present analyses, care was taken to ensure that the zones were small enough to capture, rather than overwhelm, the opportunities factor in the trip probability function.

The trip probability functions for each of the four time points of this analysis are shown in figures 3, 4, 5, and 6. It should be noticed that these figures have different vertical scales. This is immaterial in the model, since reference to equation (11) will

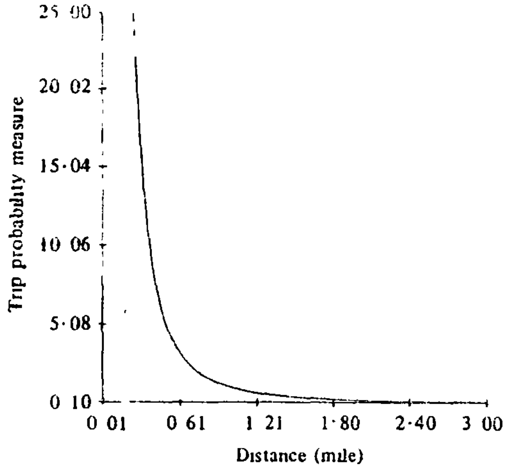


Figure 3. Trip function: Philadelphia 1850

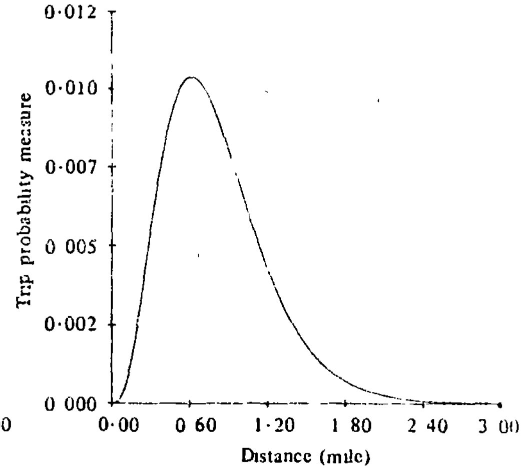


Figure 4. Trip function Philadelphia 1860

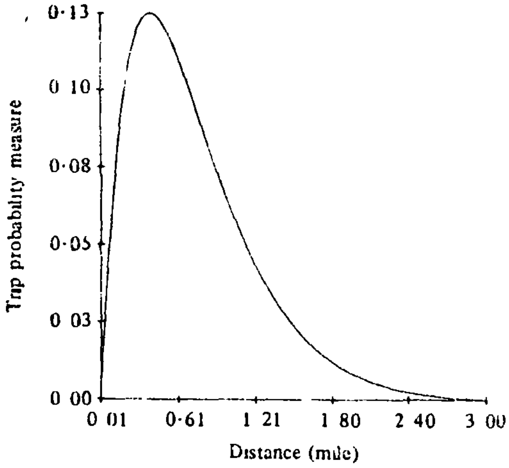


Figure 5. Trip function. Philadelphia 1870.

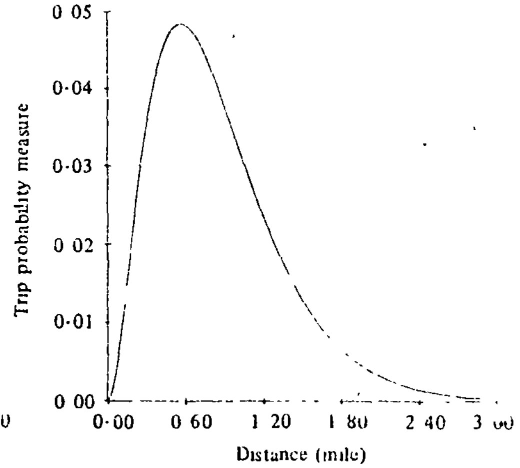


Figure 6. Trip function Philadelphia 1880.

show that any value in the numerator of the right-hand side of the equation will be normalized as an integral part of the model calculations Consequently the most important aspect of these graphs is the shape of the functions and their positions along the distance axis.

The graph for 1850 shows, despite the use of a proper, fine grain, areal system, that the trip probability function is a negative power function This is no doubt due to the tight clustering of activities, both workplace and residence, in the villages, towns and cities of the day. Apparently the majority of the population in 1850 lived well within one-half mile of their place of work

By 1860 the trip probability function assumes its customary form, having opportunities increasing with trip length and with people preferring to travel as little as possible The maximum of this function occurs at a distance of approximately 0 6 mile While the distribution is skewed slightly to the left, more than half its area is to the left of 0 75 mile The shape of this function suggests that by 1860 there must have been a beginning of a trend towards greater separation of residence and workplace Even so, the vast majority of the work trips would have been less than 1·5 miles At an estimated walking speed of 15 minutes per mile, this puts most work trips under 22·5 minutes.

The trip probability function for 1870 is shown in figure 5, and has the same general form as the 1860 function. In 1870 the trip function is more sharply skewed to the left. The peak occurs at just under 0 4 mile, with half the trips being less than 0 6 mile, though virtually all the work trips (92%) remain at less than 1 5 miles

Finally the trip probability function for 1880 is shown in figure 6, again having the same general form as in 1860 and 1870. By this time, however, there has been a clear shift towards longer work trips and the implied dispersion of urban activities In 1880 the trip probability function peaks at just under 0 6 mile, half the trips are now longer than 0·7 mile, and the 95% point has moved out to 2 0 miles

Thus we see in the trip probability functions the odd discontinuity that the 1870 results make in an otherwise regular progression of trip probability function parameters from 1850 to 1880. In most contemporary analyses of these functions and models it would be necessary to write a graceful conclusion to the article at about this point, which would include some further speculation as to what caused these results In this case, however, we have the benefit of one-hundred years of hindsight We know, for the most part, the important phenomena which were acting in and on Philadelphia a century ago These phenomena correlate perfectly with the results of our parameter estimation, and are described in the following section of the paper

Historical context

After the parameter estimates for the four time points were completed an attempt was made to place the results in an historical context In particular it was hoped that something would be found to have taken place between 1860 and 1870 to explain the 'jog' in the 1870 parameter estimates

It has been hypothesized that there is a strong relationship between transport development and building activity both at regional and at local level (Isard 1942) This study identified six associated building cycles and transport development surges. At the national level the third of these building cycles is dated 1862–1878 and is associated with a surge of railroad construction of such magnitude that US railroad mileage doubled during the period 1865–1873. This, along with a backlog of housing demand caused by the Civil War, produced a nationwide surge in housing construction after the war's end

---

In Philadelphia, beginning in the 1830s, a form of public transportation known as the omnibus, was developed These vehicles were large, enclosed carriages drawn by horses through the city's cobblestone streets By 1857 there were 322 of these omnibuses in operation. So far as can be determined from available evidence, these vehicles were little used for the journey to or from work (Hershberg et al, 1977) Virtually everyone walked to work, as was surmised above from 1850 DRAM parameter estimates.

In 1857 construction of a new form of transportation, the horse-drawn streetcar, or street railroad, was begun. This new system consisted of iron rails laid in the city streets, thus allowing the size (passenger-carrying capacity) of horse-drawn vehicles to be tripled and quadrupled. By 1858 there were 50 miles of track, over which an annual total of approximately forty-six thousand passengers were carried By 1859 the number of horse-drawn omnibuses had declined to 56, and a year later, in 1860 there were 339 streetcars in operation, whereas the omnibuses had virtually ceased to operate. By 1870 there were just under 500 streetcars in operation and by 1880 fifteen companies were operating 800 streetcars over 239 miles of track, carrying ninety-nine million passengers per annum. This transport system, in addition to considerable trackage in the developed areas of the city, radiated out to rural land and adjoining towns via six spurs, and was operational in advance, for the most part, of the city's emerging housing boom

The population of Philadelphia for the four time points was

1850	1860	1870	1880
408000	565000	674000	841000.

The nearly 40% increase between 1850 and 1860 produced a rather crowded situation with one in four persons living in the most densely populated parts of the city[1] Along with the further 20% increment between 1860 and 1870, it is clear that there was a considerable pent-up demand for housing during the 1860s and 1870s This situation was further exacerbated by the decline in construction activity during the Civil War years of 1861-1865.

After the Civil War and into the 1870s there was a tremendous construction boom in Philadelphia. There was a substantial change in the character of the construction industry. Prior to this period, housing was built by carpenters and small builders at the request of the home purchaser. During the construction boom large contractors developed great parcels of land at the fringe of the existing built-up portions of the city. By 1880, one-third of the existing buildings had been constructed in the prior ten-year period.

In the 1840s and 1850s the most common building types were three and four stories high. The new housing of the 1860s and 1870s were the two-story row houses. These could be built cheaply and sold for twice their building cost, and were within the financial reach of a larger proportion of the population. During the overall period 1850 to 1880 the inhabited area doubled, the population more than doubled, and the average population density declined by almost 17%. The proportion of persons living in the city's densest areas declined from one-in-four in 1860 to one-in-eleven in 1880. Thus the period saw a tremendous extension of the city's developed land, with more than double the 1850 population living in numerous smaller houses with fewer persons per household.

Along with all this activity, as both cause and effect, a major social reorganization of the city was taking place. The population was coming to be divided into two

---

[1] Most of these statistics are derived from various unpublished manuscripts and research proposals of the Philadelphia Social History Project.

groups, reflecting the difference between manual workers and an increasingly different white-collar group. The manual workers, composing the great majority of the labor force, had their residence location closely tied to their place of work The second, and generally wealthier, group tended to move to newly emerging suburbs, commuting to work via the streetcars and recently constructed commuter passenger railroads In fact, in the period 1850 to 1880 the central area of the city experienced a 23% decline in resident population and a 60% increase in the number of business firms This phenomenon, coupled with increasing numbers of much larger scale, land intensive, industrial enterprises, was the beginning of the eventual division of the city into separate residential, commercial, and industrial areas

An obvious consequence of all these goings on was a set of substantial changes in work-trip patterns, which tie directly back to the results of the parameter estimates presented in the previous section of this paper. From an empirical examination of the work trips of 4000 persons between 1850 and 1880, it was concluded that "The vast majority of the 4,000 persons whose journey-to-work was reconstructed lived within six-tenths of a mile from where they worked in 1850 and within one mile in 1880" (Hershberg et al, 1974). These results compare well with estimates of one-half mile in 1850 and nine-tenths of a mile in 1880 which result from the DRAM calibrations

It seems quite clear that the odd jog of the 1870 parameter estimates results from the sharp change in transport technology Introduced in 1857–1858, this new system would not have had much effect on the 1860 patterns of residence and workplace However, this change operating in the 1860s, coupled with the housing boom, would have had a noticeable effect on the city's spatial patterns by 1870 The question remains as to the way in which these patterns are reflected by the 1870 DRAM parameters? It may be that the 1870 spatial distributions did indeed yield a trip probability function shifted to the left of the 1860 function, followed by a more normal shift to the right (resuming the trend) for the 1880 trip probability function It could have been that the crowded city of the early 1860s resulted in a significant increase in peoples' willingness to travel if, by so doing, they could escape the crowding. The construction boom which began after the Civil War, being largely in the hands of developers, would have produced a pattern of widespread but 'spotty' development in 1870. In the following decade the vacant areas within this pattern, for example areas between the 'spokes' of the new radial transportation system, would have filled in. Thus the pattern of 1880 would not be very much more widespread, but rather more densely developed and largely within the confines of the 1870 pattern.

Alternatively, the 1870 shift may result from the fact that distance was used in these initial calibrations in lieu of travel time If the travel times were used in the calibrations, the 1870 parameters might not show this discontinuity. In this case one might find the trip probability functions for both 1870 and 1880 shifted far enough to the right to form a regular progression from 1850 to 1880. It is interesting to note that an estimated 37% of the street-railway trips taken in 1880 were work trips This estimate is extended to suggest that approximately 17% of employed persons regularly used this mode of travel to work. Although this is not a very large percentage, it probably accounts for a large portion of the work trips of the white-collar work force. This work force itself comprised the bulk of all workers living at any distance from their workplace. Consequently these street railroads, while not carrying a very large percentage of the work force, were carrying many of those who were most residentially mobile and who would be leading the emerging trend towards suburbanization. Future work with this data will attempt to investigate this question

As a further illustration of these hypotheses, the cumulative probabilities of the trip functions for the four time points are shown in figure 7. It is quite clear that in 1850 there was very little separation of place of work from place of residence. Historically the evidence suggests that many persons worked in one room of their residence. By 1860 there was obviously a well-established trend towards separation of place of work from place of residence, with the cumulative probability curve being not as steep as the one for 1850, but steeper than either 1870 or 1880. Notice that the 1870 and 1880 curves run parallel to each other over a substantial part of their lengths. This set of cumulative probability curves illustrates the change from 1850 to 1860 and the close similarity between 1870 and 1880. It still cannot be resolved as to whether (a) substituting travel times (including the presence of the transit system) for zone-to-zone distance would shift the 1870 and 1880 curves, together, to the right of 1860, or (b) 1870 and 1880 represent a different type of urban spatial phenomenon compared to 1850 and 1860, which would yield further parallel curves for 1890 and 1900, etc forming an orderly progression to the right along the axis of the graph, but which are in their proper relationship to 1860 as now drawn

The degree of dispersion from 1850 to 1880 varied considerably amongst various types of employment. Most types showed some dispersion, but some showed much more than others. A further attempt to disaggregate the population in estimating the DRAM parameters should yield some interesting evidence in this regard.

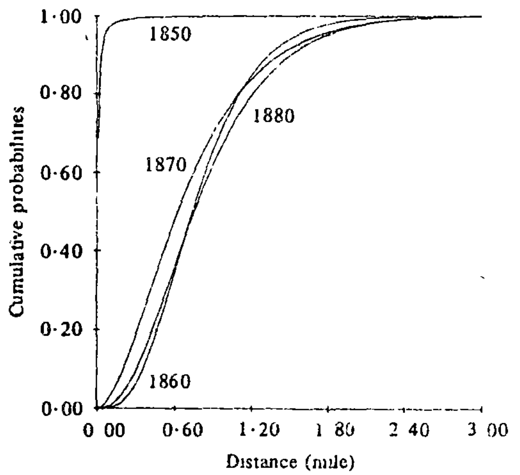


Figure 7. Cumulative trip function Philadelphia 1850-1880

Conclusion

The substantive conclusions have really been made in the previous sections of this paper. There are some general points to be made regarding methods. First, there has been some reluctance among planning practitioners in the US to undertake the use of Wilson-type Lowry models. This may be ascribed in part to the perceived difficulties of parameter estimation. It is worth noting that once the data files were prepared (a task which is much the same regardless of whether formal or informal calibration procedures are contemplated), standard search routines allowed the calibration of all four time points to be accomplished in two days at a computer cost of about $250 for the lot.

Second, and perhaps of greater importance, is the question of the appropriateness of the Tanner function as a trip probability function, and the value of systematic calibration of these models. For the first part it may now be said that the Tanner function has given results consistent with macrobehavioral theory for several different US cities, at various levels of spatial detail, both for nineteenth- and for twentieth-century data sets. This evidence becomes more difficult to repudiate than any we have had to date For the second part, it is difficult to believe that any of the ad hoc parameter estimation procedures still in use would have been anywhere near as sensitive to the changing structure of Philadelphia over a period of forty years during the nineteenth century. This certainly suggests a greater degree of validity for these types of calibration than we could have expected from any less formal parameter estimating procedures.

Acknowledgements. Theodore Hershberg, Director of the Philadelphia Social History Project, made his project literature and data resources available to us and was generally helpful in this endeavor

Julian Grauer, a graduate student in the Department collected the material from the Social History Project and did much of the work necessary to convert it to our use, including the development of the 105-zone areal system.

Fred Ducca and Dick Greenfield, both graduate research assistants, aided in the calibration programs and data manipulations respectively.

This research was partially supported by National Science Foundation Grant APR 73-07840-A02 "Development of an Improved Transportation and Land Use Model Package", Principal Investigator Stephen H Putman. The Grant is financed in part by the US Department of Transportation

References

Batty M, Mackie S, 1972 "The calibration of gravity, entropy, and related models of spatial interaction" *Environment and Planning* 4 205-233

Hershberg T 1976 "The Philadelphia social history project an overview and progress report" *Review of Public Data Use* 4(1) 29-32

Hershberg T, Cox H, Light D Jr, 1977 "The 'journey-to-work' an empirical investigation of work, residence, and transportation, Philadelphia 1850 and 1880" in *Toward a "New" Urban History Urbanization and Industrialization in Nineteenth Century Philadelphia* Ed T Hershberg (Oxford University Press, London) forthcoming

Isard W, 1942 "Transport development and building cycles" *Quarterly Journal of Economics* 57 90-112

Lowry I S, 1964 "A model of metropolis" RM-4125-RC, RAND Corporation, Santa Monica, Calif

Putman S H, 1977 "Calibrating a disaggregated residential allocation model—DRAM" in *London Papers in Regional Science 7 Alternative Frameworks for Analysis* Eds D B Massey, P W J Batey (Pion, London) pp 108-124

Wilson A G, 1970 "Entropy in urban and regional modelling" (Pion, London)

PRELIMINARY DRAFT

CALIBRATING URBAN RESIDENTIAL MODELS 1:
PROCEDURES AND STRATEGIES

Stephen H. Putman

Frederick W. Ducca

September, 1977

Urban Simulation Laboratory
Department of City and Regional Planning
University of Pennsylvania

Abstract

Calibration of urban residential models has long been an area of confusion and apprehension for prospective model users. With the evolution of entropy maximizing model formulations there has been new progress made toward resolving the calibration problem. Most of this progress has assumed the availability of spatial interaction data (i.e. the T_{ij} matrix). Many situations arise where a singly constrained model is to be calibrated, e.g. residential or retail location models, in the absence of this data. This paper describes a method which has been successfully used for calibrating urban residential models of that type.

Acknowledgement

The work described herein was supported by National Science Foundation Grant APR73-07840-A04, "Development of an Improved Transportation and Land Use Model Package", Principal Investigator, Stephen H. Putman. This project was financed in part by the U.S. Department of Transportation.

INTRODUCTION

With the development of metropolitan residential location models via the entropy maximizing forms, a new era in the calibration of such models began. Prior to this development it was difficult even to locate the parameters in the model equations, much more, to develop procedures for their estimation. Since this development, while the parameter estimation problem remains difficult, much progress has been made.

A difficulty which has arisen is that there are many strategies and procedures for calibrating these models, and it is not always clear as to the way in which these relate, one to the other. Thus, this first section of the paper is devoted to a brief discussion of some of the ways suggested as being useful in calibrating spatial interaction models.[1]

"Spatial interaction model", in this context, means the singly constrained type model used as a location model. This is but one of a set of types of such models. At one extreme is the unconstrained model, of the form:

$$T_{ij} = k \, O_i \, D_j \, \mathcal{F}(C_{ij}) \qquad (1)$$

where

T_{ij} = trips from zone i to zone j
O_i = trips originating at zone i
D_j = trips terminating at zone j
C_{ij} = "cost" of a trip between zones i and j
k = an empirically determined constant

---

[1] A more detailed survey of many of these may be found in (Batty and Mackie, 1972) and an even more complete discussion is given in ch. 6 and 7 of (Batty, 1976). Other recent discussions may be found in (Baxter and Williams, 1975), (Openshaw, 1976), and (Williams, 1976).

This form of model is intrinsically linear, and for usual forms of $\mathcal{F}(C_{ij})$ can be linearized; k can then be found by linear regression techniques.

At the other extreme is the doubly constrained model of the form:

$$T_{ij} = A_i B_j O_i D_j \mathcal{F}(C_{ij}) \tag{2}$$

where there are trip end constraints of the form:

$$\sum_j T_{ij} = O_i \ ; \ \sum_i T_{ij} = D_j \tag{3,4}$$

and where

$$\sum_i \sum_j T_{ij} = \sum_i O_j = \sum_j D_j = T \tag{5}$$

The A_i and B_j which replace the k of equation (1) have the form:

$$A_i = \left[\sum_j B_j D_j \mathcal{F}(C_{ij})\right]^{-1} \tag{6}$$

$$B_j = \left[\sum_i A_i O_i \mathcal{F}(C_{ij})\right]^{-1} \tag{7}$$

Between the two extremes are two singly constrained models. The production constrained model is:

$$T_{ij} = A_i O_i D_j \mathcal{F}(C_{ij}) \tag{8}$$

subject to

$$\sum_j T_{ij} = O_i \tag{9}$$

thus yielding

$$A_i = \left[\sum_j D_j \mathcal{F}(C_{ij})\right]^{-1} \tag{10}$$

The attraction constrained model is:

$$T_{ij} = B_j O_i D_j \mathcal{F}(C_{ij}) \qquad (11)$$

subject to

$$\sum_i T_{ij} = D_j \qquad (12)$$

thus yielding

$$B_j = \left[\sum_i O_i \mathcal{F}(C_{ij})\right]^{-1} \qquad (13)$$

None of these models have any parameters to estimate other than those which may appear in the $\mathcal{F}(C_{ij})$ term. One of the simplest $\mathcal{F}(C_{ij})$ in common usage is:

$$\mathcal{F}(C_{ij}) = e^{-\beta C_{ij}} \qquad (14)$$

Our specific concern here is with calibrating residential location models in the absence of interaction data. Thus, we formulate the following simple location (attraction constrained) model:

$$\hat{N}_i = \sum_j B_j O_i D_j e^{-\beta C_{ij}} \qquad (15)$$

where

$$B_j = \left[\sum_i O_i e^{-\beta C_{ij}}\right]^{-1} \qquad (16)$$

If we use employment in zone j as a measure of D_j and the the "attractiveness" of zone i as a measure of O_i, we then get:

$$\hat{N}_i = \sum_j E_j \left[\frac{W_i \exp(-\beta C_{ij})}{\sum_i W_i \exp(-\beta C_{ij})}\right] \qquad (17)$$

where

$\hat{N}_i$ = estimated employed residents of zone i

W_i = "attractiveness" of zone i

E_j = employment in zone j

The question now arises as to how we may estimate the numerical value of the parameter β. This requires first that a criterion be selected by which one may measure the "goodness-of-fit" of the model formulation to the data set being used for calibration. This criterion, C, will be a function of an observed variable and a set of estimates of the variable. These estimates are, in turn, a function of the parameter value. Thus:

$$C = \mathcal{F}(N_i;\ \hat{N}_i;\ \beta) \qquad (18)$$

where

N_i = observed employed residents of zone i

Two general strategies may be followed in determining that value of β which optimizes the criterion. The first of these, often followed in British practice, involves developing a set of equations using Langrangian multipliers and then partial derivatives which describe the criterion at its optimum. These equations are then solved simultaneously (usually by some iterative technique) for the value of β at the criterion's optimum. These equations are sometimes solved by search techniques, but the overall strategy remains unchanged.

The second strategy, with which we are concerned here, uses the derivative of the criterion with respect to the parameter β to guide a search technique for finding the optimum value of the criterion. If one is dealing with a model for which the derivatives are readily calculated, this calibration strategy seems to afford the advantages of simplicity of approach and computational efficiency, even for problems involving the simultaneous estimation of as many as nine parameters (the largest size problem tested to date).

Further, it may be shown that for situations where only trip end (i.e. activity location) data are available, the first strategy is not practical. Strategy One typically uses the followling likelihood function as a criterion.

$$L = \prod_i \prod_j P_{ij}^{T_{ij}} \tag{19}$$

where

T_{ij} = observed interactions (e.g. trips) between zones i and j.

and where

$$P_{ij} = \frac{W_j C_{ij}^{\beta}}{\sum_k W_k C_{ik}^{\beta}} \tag{20}$$

Strategy Two uses the following likelihood function as a criterion.

$$L = \prod_j P_j^{N_j} \tag{21}$$

where

N_j = observed activity level in zone j.

and where

$$P_j = \frac{\hat{N}_j}{\sum_j \hat{N}_j} \tag{22}$$

where

$\hat{N}_j$ = estimated activity level in zone j.

and, where

$$\hat{N}_j = \sum_i E_i \left[\frac{W_i \; C_{ij}^{\beta}}{\sum_i W_i \; C_{ij}^{\beta}} \right] \tag{23}$$

In order to make the necessary calculations feasible, in both cases the natural logarithm of the likelihood function is used. Thus we have for Strategy One, the following

$$\ln L = \sum_i \sum_j T_{ij} \ln P_{ij} \tag{24}$$

and, substituting from equation (20)

$$\ln L = \sum_i \sum_j T_{ij} \left[\left\{ \ln \left(\sum_k W_k \; C_{ik}^{\beta} \right) \right\} + \ln W_j + \beta \ln C_{ij} \right] \tag{25}$$

By use of Lagrange multipliers to incorporate the necessary constraints and partial derivatives with respect to β, the equations for the maximum of ln L may be derived. This set of equations may then be solved simultaneously for the maximizing value of β.

For Strategy Two we may also write the log-likelihood function, as follows

$$\ln L = \sum_j N_j \ln P_j \tag{26}$$

and, substituting from equation (22)

$$\ln L = \sum_j N_j \ln \frac{\hat{N}_j}{\sum_j \hat{N}_j} \tag{27}$$

Recalling the definition of $\hat{N}_j$ in equation (23) it is clear that the form of this last equation makes it difficult (perhaps impossible) to derive the necessary equations for maximizing ln L in a fashion similar to that of Strategy One. Consequently in situations where only the trip end (e.g. activity location) data are available, Strategy Two is the one to be followed

Some explorations of the nature of the criterion functions and the shapes of response surfaces of criteria as functions of parameters has been done for Strategy One, and is described in the references cited above. This paper presents explorations into similar questions associated with the second calibration strategy. We note that it has not yet been demonstrated that one will obtain identical parameter values (for a given model and data set) by either calibration strategy. This question is the subject of ongoing investigations.

In the following pages we endeavor to present results of a step-by-step exploration of the response surface for the R^2 and the likelihood (L) criteria. Variations in this surface due to problem size, problem data sets, and alternate model formulations are examined. In addition, the formulae are derived for some of the partial derivatives necessary for gradient search of these surfaces.

The R^2 Criterion Function: One Parameter Case

We begin with an exploration of the characteristics of the R^2 and ∇R^2 functions for a three zone test problem. The data for the test problem were as follows:

Employment:	Attractiveness:
$E_1 = 100$	$W_1 = 3$
$E_2 = 150$	$W_2 = 4$
$E_3 = 200$	$W_3 = 5$

Cost Matrix:

$C_{11} = 1.5$	$C_{21} = 2.5$	$C_{31} = 4.0$
$C_{21} = 2.5$	$C_{22} = 2.0$	$C_{23} = 3.5$
$C_{31} = 3.0$	$C_{32} = 3.5$	$C_{33} = 1.5$

A value of -2.0 was assumed for β.

Using these numbers, a estimate of the residents in each zone was calculated per the standard equations, as follows:

$$\hat{N}_i = \sum_j E_j \left[\frac{W_i \exp(\beta\ C_{ij})}{\sum_i W_i \exp(\beta\ C_{ij})} \right] \tag{28}$$

The results of this were

$\hat{N}_1 = 110.8718$

$\hat{N}_2 = 129.2546$

$\hat{N}_3 = 209.8734$

It was necessary to keep this level of precision as these values of $\hat{N}_i$ were, in turn used as the "observed" values of N_i in the subsequent gradient calculations. By so doing, the desired optimum of the function was known to be -2.00 exactly.

The criterion measure used for these first tests was the normalized sum of squared deviations, R^2.

The equation:

$$R^2 = 1.0 - \left[\frac{\sum_i (\hat{N}_i - N_i)^2}{\sum_i (N_i - \bar{N}_i)^2} \right] \qquad (29)$$

Where

$\hat{N}_i$ = estimated residents of zone i.

N_i = observed residents of zone i.

$\bar{N}_i$ = mean observed residents.

In all subsequent discussion, the observed residents are the estimates calculated above. That is, for all remaining calculations and discussions we use:

$N_1 = 110.8718$

$N_2 = 129.2546$

$N_3 = 209.8734$

The gradient search technique of parameter estimation requires that the derivatives of the criterion with respect to the parameter(s) of interest be calculated. In this case the criterion is R^2 and the parameter of interest is β. Consequently we must be able to calculate the derivative of R^2 with respect to β at any point. This may be done by numerical approximation, i.e. make a small change in β and calculate the resultant change in R^2. The derivatives may also be calculated exactly by developing the equation for the derivative. In general this procedure will be computationally more efficient in cases where the relationship between R^2 and β is very complex, but yields differentiable functions.

The equation is developed as follows, for the gradient, ∇, of R^2.

$$\nabla R^2 = \frac{\partial R^2}{\partial \beta} = \frac{\partial R^2}{\partial \hat{N}_i}\left(\frac{\partial \hat{N}_i}{\partial \beta}\right) \tag{30}$$

$$= -2.0 \left[\frac{\sum_i \left[(\hat{N}_i - N_i) \left(\frac{\partial \hat{N}_i}{\partial \beta} \right) \right]}{\sum_i (N_i - \bar{N}_i)^2} \right] \tag{31}$$

After a considerable number of steps, we get

$$\nabla R^2 = -2.0 \left[\frac{\sum_i \left\{ (\hat{N}_i - N_i) \left[\sum_j \left(\frac{E \quad L_{ij}}{M_j} \right) \left(C_{ij} - \frac{\sum_k C_{kj} \; L_{kj}}{M_j} \right) \right] \right\}}{\sum_i (N_i - \bar{N}_i)^2} \right] \tag{32}$$

where

$$L_{ij} = W_i \; \exp(\beta \; C_{ij}) \tag{33}$$

$$M_j = \sum_i L_{ij} = \sum_i W_i \; \exp(\beta \; C_{ij}) \tag{34}$$

We may thus calculate R^2, ∇R^2, and the $\hat{N}_i$ for various values of β. The results, for a few selected values of β are:

β	R^2	∇R^2	$\hat{N}_1$	$\hat{N}_2$	$\hat{N}_3$
0.00	0.83167	-0.40718	112.500	150.000	187.500
-0.40	0.95214	-0.15049	103.589	142.536	203.875
-0.80	0.97252	-0.00502	100.791	134.328	214.881
-1.00	0.97455	-0.01882	101.560	131.559	216.881
-1.40	0.98785	-0.03818	105.254	128.898	215.848
-1.80	0.99874	-0.01319	109.279	128.788	211.932
-2.00	1.00000	-0.00000	110.872	129.255	209.873
-2.40	0.99665	+0.01405	112.913	130.718	206.369
-2.80	0.99031	+0.01655	113.567	132.537	203.897
-3.00	0.98701	+0.01635	113.492	133.500	203.008
-4.00	0.97053	+0.01761	111.013	138.224	200.763

Thus we see that there is a perfect R^2 at a β of -2.00, which we would expect as our "observed" values of N_i were obtained by use of a -2.00 value for β. Further we see that the gradient, ∇R^2 is negative for β values greater than -2.00, and positive for β values less than -2.00, thus suggesting which way to change β from any point, in order to improve the R^2 value. A graph of R^2 and ∇R^2 versus β is given in Figure 1.

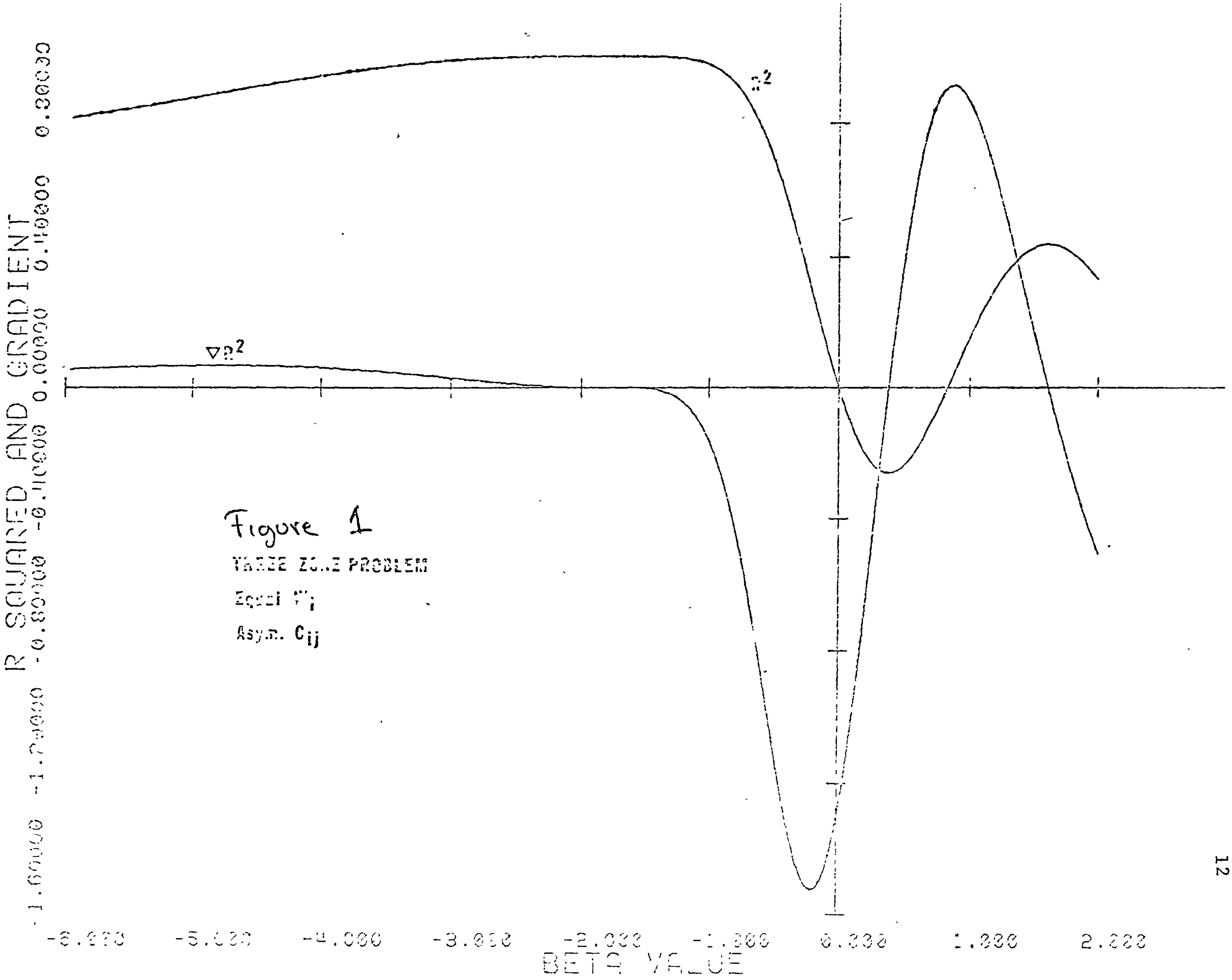


Figure 1

THREE ZONE PROBLEM

Equal V_i

Asym. C_{ij}

An immediately obvious problem appears from Figure 1. If the gradient search procedure had been initiated with a value of β greater than 0.4 (approximately), the value of ∇R^2 would have been positive and the search routine would have yielded a parameter estimate of 1.6 (approximately) for β. This was a disturbing result, ameleorated only partly by the suggestion that all search procedures use starting β values less than zero. The desire to know more about these functions led to the series of experiments which follows.

In the following experiments we explore the different shapes of criterion function and its gradient resulting from variations in:

a) the criterion itself, e.g. R^2 or Likelihood

b) the data set

c) the number of observations (zones) in the data set

d) the functional form of the model with respect to the travel function and the attractiveness measure

Our particular concerns are:

a) Is the criterion function unimodal or multimodal in the region of interest?

b) How steep or shallow is the function near its optimum?

c) How do the answers to the above affect calibration procedures and strategies for their use?

Three Zone Test Problem - R^2 Criterion: One Parameter Case

After getting the results obtained and shown in Figure 1 for the three zone problem described above, several variations were tried. The original cost matrix was asymmetric. New runs were made using a symmetric cost matrix. New runs were also made with equal values of attractiveness for all zones. The results of these tests are shown in Figures 2 and 3. Note that in all these tests the "observed" values of residents in each zone is the exact value which is yielded by the model equation with a value of -2.00 for β. By so doing we always know where the optimum of the criterion function should occur and thus are able to compare the various tests.

The results evident from these first tests, for three zones, are:

1) that changes in the data set have significant effects on the shape of the R^2 curve;
2) that the R^2 criterion function is sometimes bimodal, but never in the region of interest where β is less than zero; and
3) that the optimum is not very sharply defined, particularly on the left or "more negative" β value side of the curve.

Nine Zone Test Problem - R^2 Criterion

To examine how the above results might change with an increase in the number of zones, the following nine zone problem was created:

Nine Zone Test Problem

Employment:	Attractiveness:
$E_1 = 100$	$W_1 = 3$
$E_2 = 50$	$W_2 = 4$
$E_3 = 25$	$W_3 = 5$
$E_4 = 0$	$W_4 = 8$
$E_5 = 25$	$W_5 = 6$
$E_6 = 75$	$W_6 = 4$
$E_7 = 200$	$W_7 = 2$
$E_8 = 250$	$W_8 = 1$
$E_9 = 500$	$W_9 = 2$

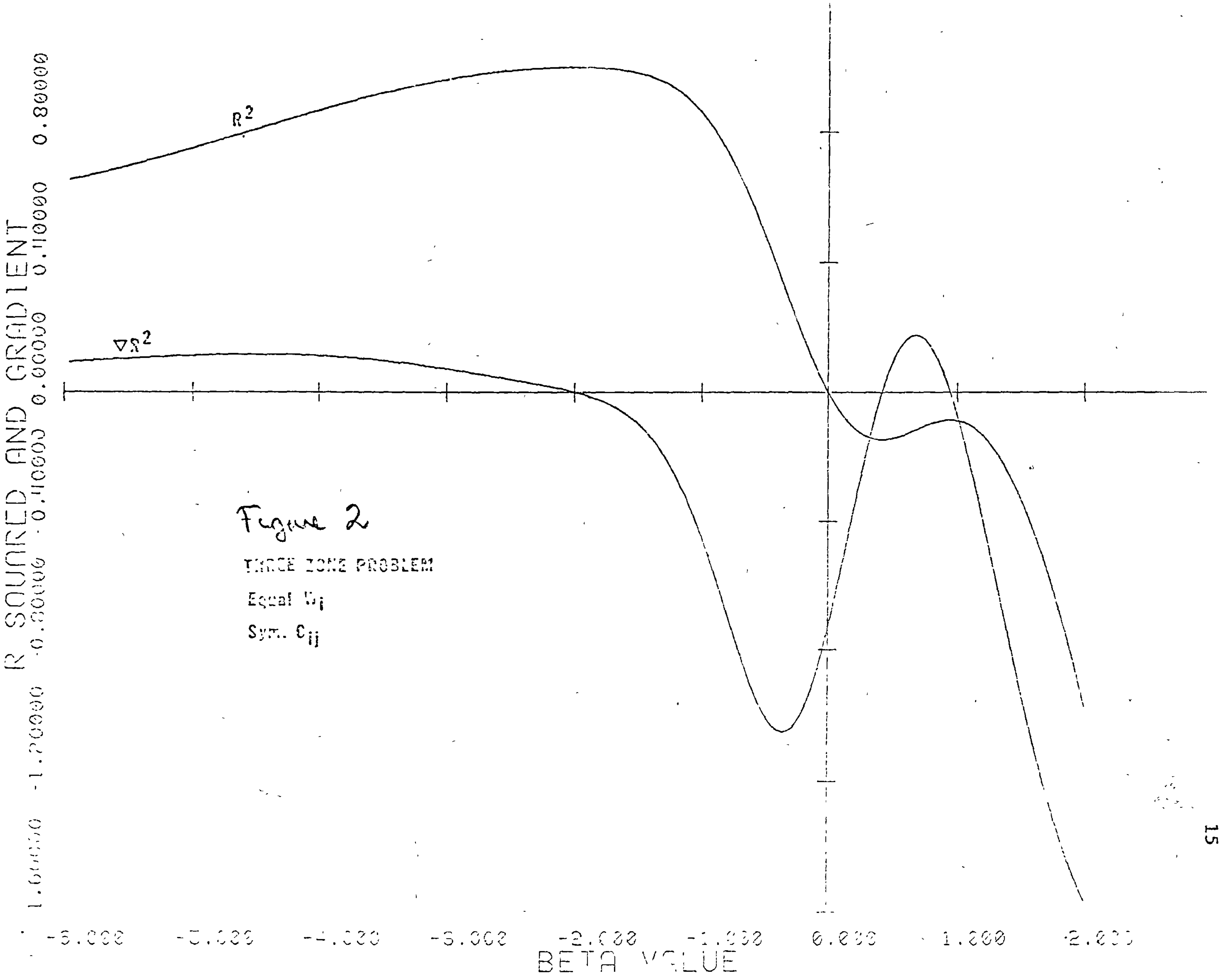


Figure 2

THREE ZONE PROBLEM

Equal ω_i

Sym. C_{ij}

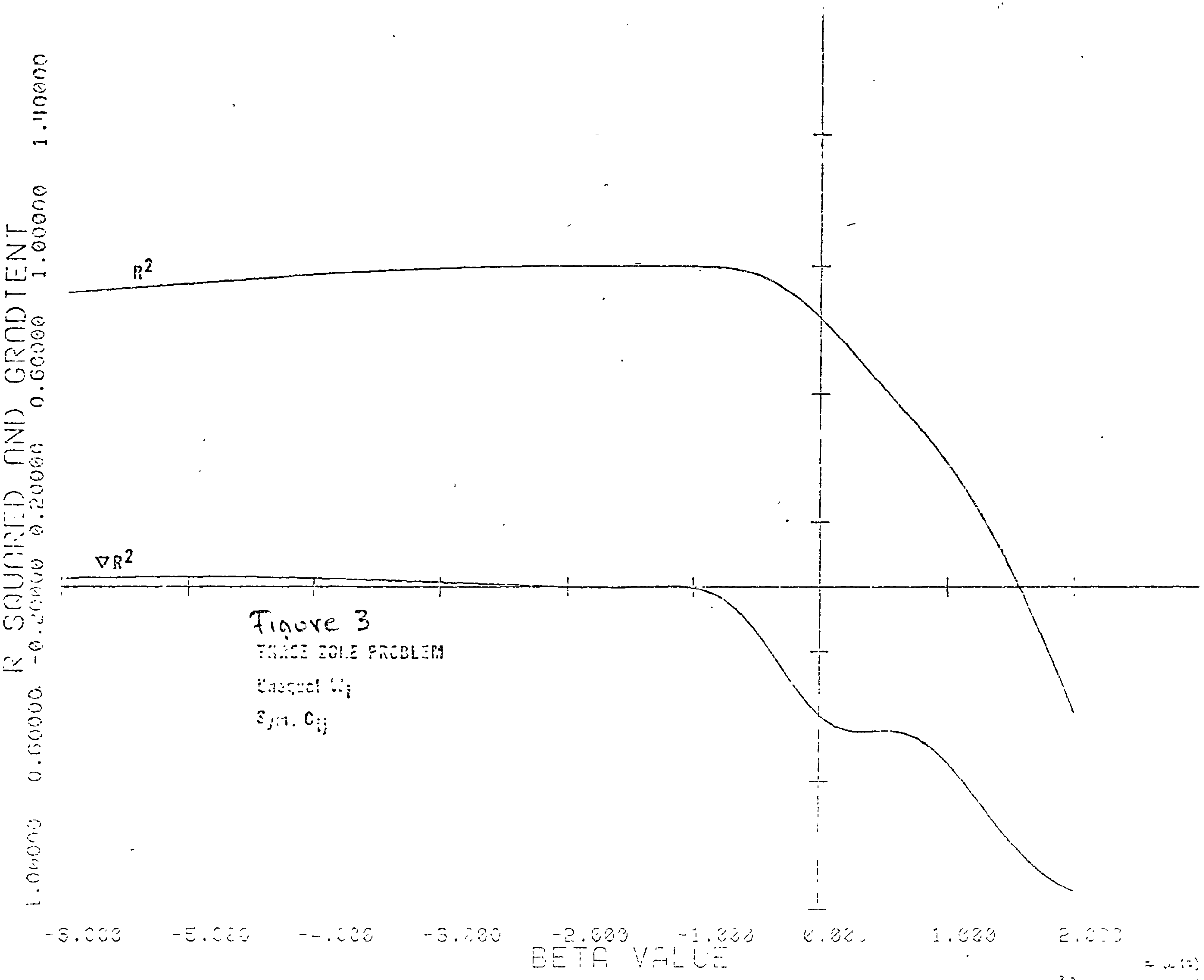


Figure 3
THREE ZONE PROBLEM
Unequal W_i
Sym. C_{ij}

Cost Matrix

	1	2	3	4	5	6	7	8	9
1	1.5	2.5	3.5	4.5	4.0	4.0	4.0	2.0	2.0
2	2.5	1.0	1.5	2.0	1.5	3.0	3.5	3.0	2.5
3	3.5	1.5	1.0	3.0	1.0	3.5	4.0	4.0	3.0
4	4.5	2.0	3.0	1.0	1.0	1.5	3.5	3.5	2.0
5	4.0	1.5	1.0	1.0	1.0	1.5	3.5	3.5	2.0
6	4.0	3.0	3.5	1.5	1.5	1.5	2.0	2.5	2.5
7	4.0	3.5	4.0	3.5	3.5	2.0	2.0	2.0	2.5
8	2.0	3.0	4.0	3.5	3.5	2.5	2.0	2.0	2.5
9	2.0	2.5	3.0	2.0	2.0	2.5	2.5	2.5	2.5

Again, the R^2 criterion and ∇R^2 were calculated and graphed. Note also that the nine zone data were used with a $\beta = -2.00$ to generate the "observed" data. Thus, a perfect "fit" will be found for $\beta = -2.00$, and we there expect $R^2 = 1.00$ and $\nabla R^2 = 0.00$ as per the previous example.

For this case, two runs were made, both using the symmetric cost matrix. In the second case, however, the attractiveness indices were set equal to 4.0 in all nine zones. The two graphs are shown in Figures 4 and 5.

A welcome result here is that the R^2 function now seems to be unimodal. It is also encouraging to note that the sharpness of the optimum is much increased. The peak is still sharper to the right of the optimum than to the left. Again, the different data sets do show different curves, though they are much closer in shape than were the various three zone tests.

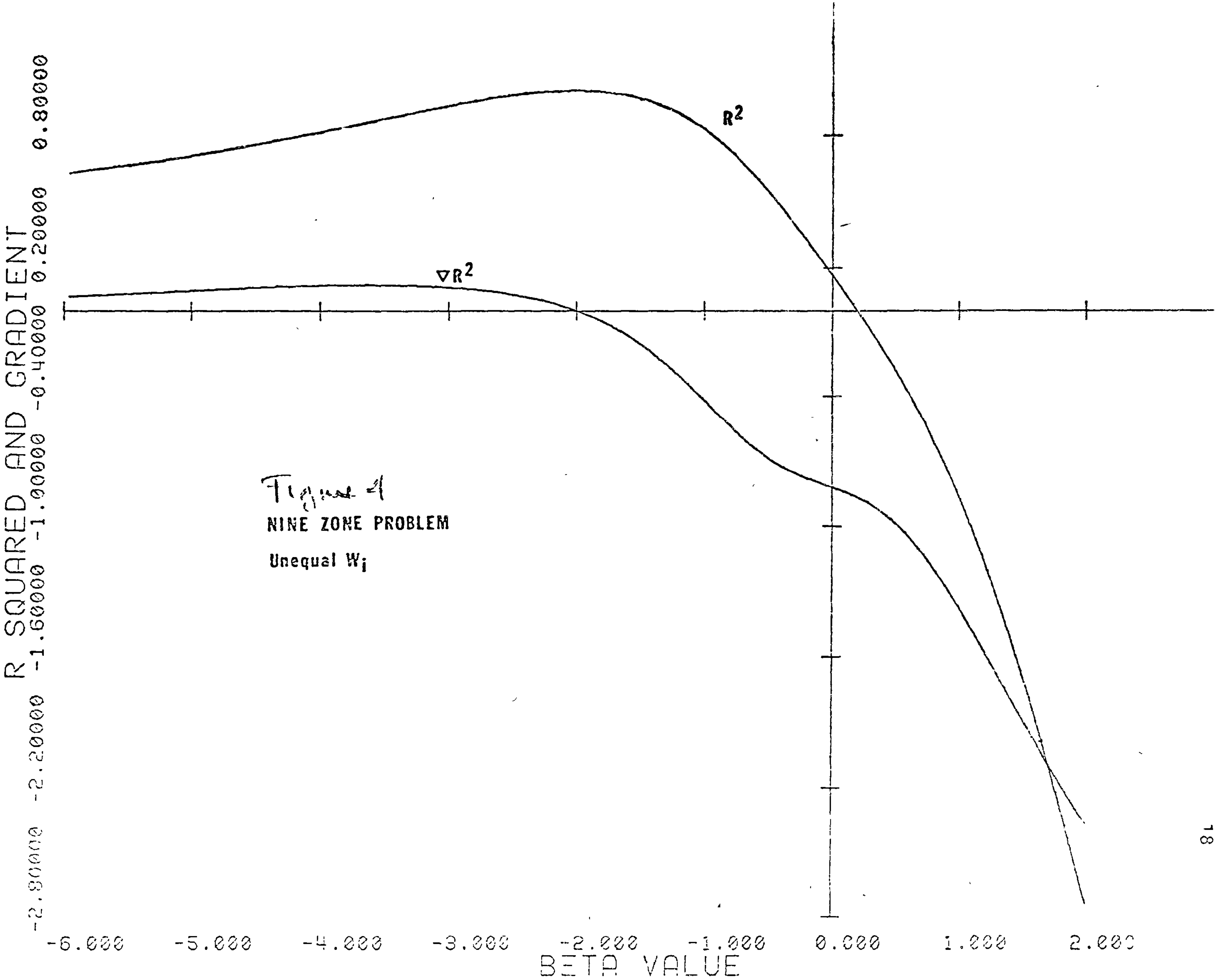


Figure 4

NINE ZONE PROBLEM

Unequal W_i

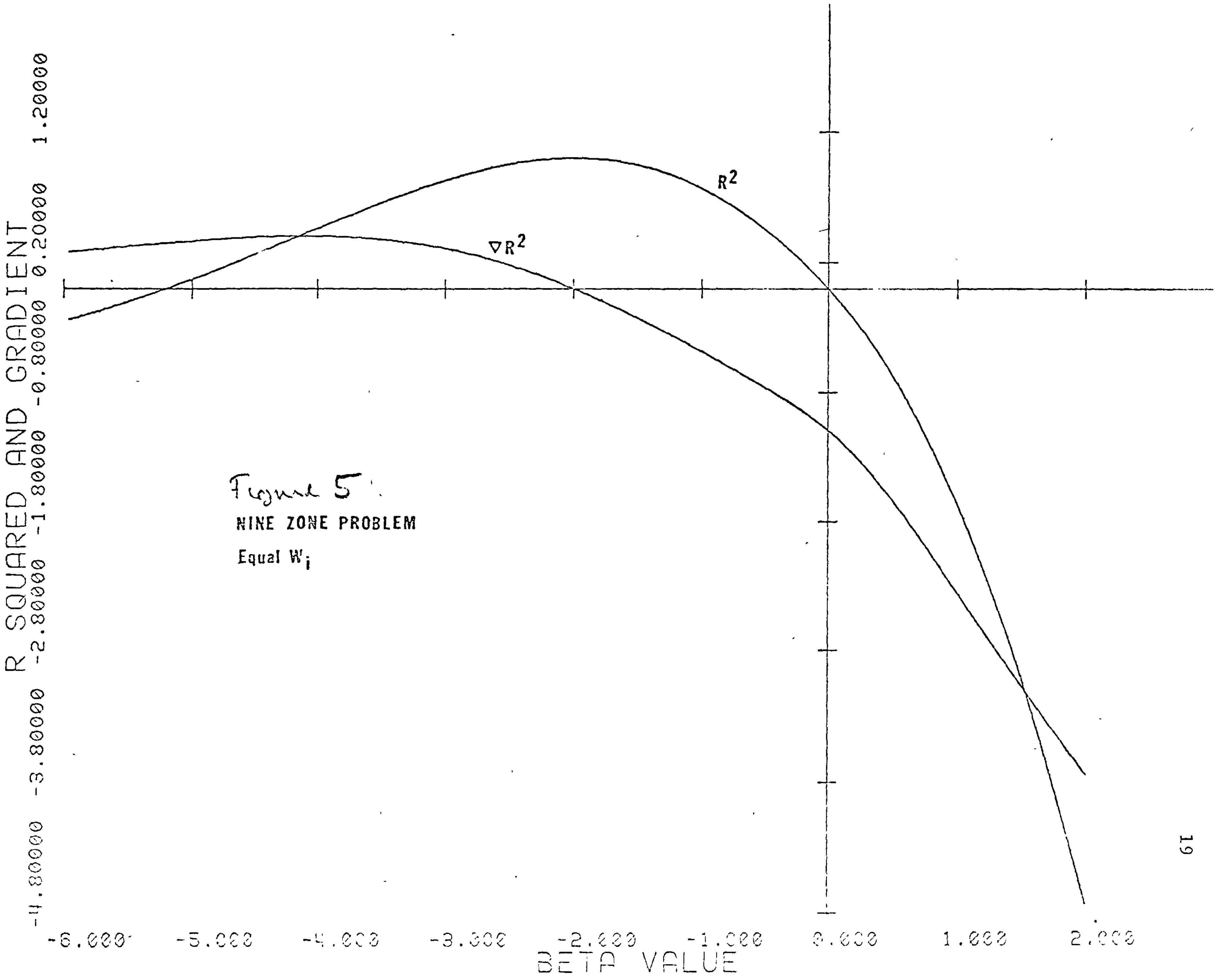


Figure 5

NINE ZONE PROBLEM

Equal W_i

We next move to a twenty-five zone problem. In this case, the data were as follows:

Twenty-Five Zone Test Problem

Employment:	Attractiveness:	Employment:	Attractiveness:
$E_1 = 25$	$W_1 = 5$	$E_{14} = 150$	$W_{14} = 1$
$E_2 = 25$	$W_2 = 5$	$E_{15} = 75$	$W_{15} = 3$
$E_3 = 50$	$W_3 = 5$	$E_{16} = 100$	$W_{16} = 3$
$E_4 = 25$	$W_4 = 5$	$E_{17} = 200$	$W_{17} = 2$
$E_5 = 25$	$W_5 = 5$	$E_{18} = 300$	$W_{18} = 1$
$E_6 = 75$	$W_6 = 4$	$E_{19} = 100$	$W_{19} = 4$
$E_7 = 50$	$W_7 = 4$	$E_{20} = 50$	$W_{20} = 4$
$E_8 = 100$	$W_8 = 4$	$E_{21} = 150$	$W_{21} = 3$
$E_9 = 75$	$W_9 = 3$	$E_{22} = 150$	$W_{22} = 3$
$E_{10} = 50$	$W_{10} = 4$	$E_{23} = 75$	$W_{23} = 4$
$E_{11} = 75$	$W_{11} = 2$	$E_{24} = 75$	$W_{24} = 4$
$E_{12} = 250$	$W_{12} = 2$	$E_{25} = 50$	$W_{25} = 4$
$E_{13} = 400$	$W_{13} = 1$		

The cost matrix is shown on the next page.

Again, the value of R^2 and ∇R^2 were plotted against the value of β. Also, as before, the optimum value of β is at -2.00, the value used to generate the "observed" data. Finally, there were again two runs, one for the W_i values shown above, and the second with a value of 4.0 for all the W_i.

The plots of these two runs are shown in Figures 6 and 7. We note here that the R^2 curve shows signs of levelling off, an unwelcome result that suggests that calibration of such a model might become tedious as a search for an optimum along such a smooth curve would proceed rather slowly.

A positive result here is that the differences between the R^2 curves for the equal W_i and unequal W_i curves are rather miminimal. This suggests that data set differences would not seriously affect the calibration procedure though they will, of course, give different optimal values of β.

At this point, however, it becomes desirable to investigate the possibility of using a different criterion function.

The Likelihood Criterion Function: One Parameter Case

An alternative criterion (actually in most cases a preferable one) is the likelihood function. This function has the form:

$$L = \prod_i (\hat{N}_i)^{N_i} \tag{35}$$

where, again

$\hat{N}_i$ = the estimated number of households locating in zone i

N_i = the observed number of households in zone i

Due to the difficulty of dealing with the product function $(\prod_i)$, it is

```
1.5 1.0 2.0 3.0 4.0 1.0 1.4 2.2 3.2 4.1 2.0 2.2 2.8 3.6 4.5 3.0 3.2 3.6 4.2 5.0 4.0 4.1 4.5 5.0 5.7
1.0 1.5 1.0 2.0 3.0 1.4 1.0 1.4 2.2 3.2 2.2 2.0 2.2 2.8 3.6 3.2 3.0 3.2 3.6 4.2 4.1 4.0 4.1 4.5 5.0
2.0 1.0 1.5 1.0 2.0 2.2 1.4 1.0 1.4 2.2 2.8 2.2 2.0 2.2 2.8 3.6 3.2 3.0 3.2 3.6 4.5 4.1 4.0 4.1 4.5
3.0 2.0 1.0 1.5 1.0 3.2 2.2 1.4 1.0 1.4 3.6 2.8 2.2 2.0 2.2 4.2 3.6 3.2 3.0 3.2 5.0 4.5 4.1 4.0 4.1
4.0 3.0 2.0 1.0 1.5 4.1 3.2 2.2 1.4 1.0 4.5 3.6 2.8 2.2 2.0 5.0 4.2 3.6 3.2 3.0 5.7 5.0 4.5 4.1 4.0
1.0 1.4 2.2 3.2 4.1 1.5 1.0 2.0 3.0 4.0 1.0 1.4 2.2 3.2 4.1 2.0 2.2 2.8 3.6 4.5 3.0 3.2 3.6 4.2 5.0
1.4 1.0 1.4 2.2 3.2 1.0 1.5 1.0 2.0 3.0 1.4 1.0 1.4 2.2 3.2 2.2 2.0 2.2 2.8 3.6 3.2 3.0 3.2 3.6 4.2
2.2 1.4 1.0 1.4 2.2 2.0 1.0 1.5 1.0 2.0 2.2 1.4 1.0 1.4 2.2 2.8 2.2 2.0 2.2 2.8 3.6 3.2 3.0 3.2 3.6
3.2 2.2 1.4 1.0 1.4 3.0 2.0 1.0 1.5 1.0 3.2 2.2 1.4 1.0 1.4 3.6 2.8 2.2 2.0 2.2 4.2 3.6 3.2 3.0 3.2
4.1 3.2 2.2 1.4 1.0 4.0 3.0 2.0 1.0 1.5 4.1 3.2 2.2 1.4 1.0 4.5 3.6 2.8 2.2 2.0 5.0 4.2 3.6 3.2 3.0
2.0 2.2 2.8 3.6 4.5 1.0 1.4 2.2 3.2 4.1 1.5 1.0 2.0 3.0 4.0 1.0 1.4 2.2 3.2 4.1 2.0 2.2 2.8 3.6 4.5


2.2 2.0 2.2 2.8 3.6 1.4 1.0 1.4 2.2 3.2 1.0 1.5 1.0 2.0 3.0 1.4 1.0 1.4 2.2 3.2 2.2 2.0 2.2 2.8 3.6
2.8 2.2 2.0 2.2 2.8 2.2 1.4 1.0 1.4 2.2 2.0 1.0 1.5 1.0 2.0 2.2 1.4 1.0 1.4 2.2 2.8 2.2 2.0 2.2 2.8
3.6 2.8 2.2 2.0 2.2 3.2 2.2 1.4 1.0 1.4 3.0 2.0 1.0 1.5 1.0 3.2 2.2 1.4 1.0 1.4 3.6 2.8 2.2 2.0 2.2
4.5 3.6 2.8 2.2 2.0 4.1 3.2 2.2 1.4 1.0 4.0 3.0 2.0 1.0 1.5 4.1 3.2 2.2 1.4 1.0 4.5 3.6 2.8 2.2 2.0
3.0 3.2 3.6 4.2 5.0 2.0 2.2 2.8 3.6 4.5 1.0 1.4 2.2 3.2 4.1 1.5 1.0 2.0 3.0 4.0 1.0 1.4 2.2 3.2 4.1
3.2 3.0 3.2 3.6 4.2 2.2 2.0 2.2 2.8 3.6 1.4 1.0 1.4 2.2 3.2 1.0 1.5 1.0 2.0 3.0 1.4 1.0 1.4 2.2 3.2
3.6 3.2 3.0 3.2 3.6 2.8 2.2 2.0 2.2 2.8 2.2 1.4 1.0 1.4 2.2 2.0 1.0 1.5 1.0 2.0 2.2 1.4 1.0 1.4 2.2
4.2 3.6 3.2 3.0 3.2 3.6 2.8 2.2 2.0 2.2 3.2 2.2 1.4 1.0 1.4 3.0 2.0 1.0 1.5 1.0 3.2 2.2 1.4 1.0 1.4
5.0 4.2 3.6 3.2 3.0 4.5 3.6 2.8 2.2 2.0 4.1 3.2 2.2 1.4 1.0 4.0 3.0 2.0 1.0 1.5 4.1 3.2 2.2 1.4 1.0
4.0 4.1 4.5 5.0 5.7 3.0 3.2 3.6 4.2 5.0 2.0 2.2 2.8 3.6 4.5 1.0 1.4 2.2 3.2 4.1 1.5 1.0 2.0 3.0 4.0
4.1 4.0 4.1 4.5 5.0 3.2 3.0 3.2 3.6 4.2 2.2 2.0 2.2 2.8 3.6 1.4 1.0 1.4 2.2 3.2 1.0 1.5 1.0 2.0 3.0
4.5 4.1 4.0 4.1 4.5 3.6 3.2 3.0 3.2 3.6 2.8 2.2 2.0 2.2 2.8 2.2 1.4 1.0 1.4 2.2 2.0 1.0 1.5 1.0 2.0
5.0 4.5 4.1 4.0 4.1 4.2 3.6 3.2 3.0 3.2 3.6 2.8 2.2 2.0 2.2 3.2 2.2 1.4 1.0 1.4 3.0 2.0 1.0 1.5 1.0
5.7 5.0 4.5 4.1 4.0 5.0 4.2 3.6 3.2 3.0 4.5 3.6 2.8 2.2 2.0 4.1 3.2 2.2 1.4 1.0 4.0 3.0 2.0 1.0 1.5
```

25 Zone Cost Matrix

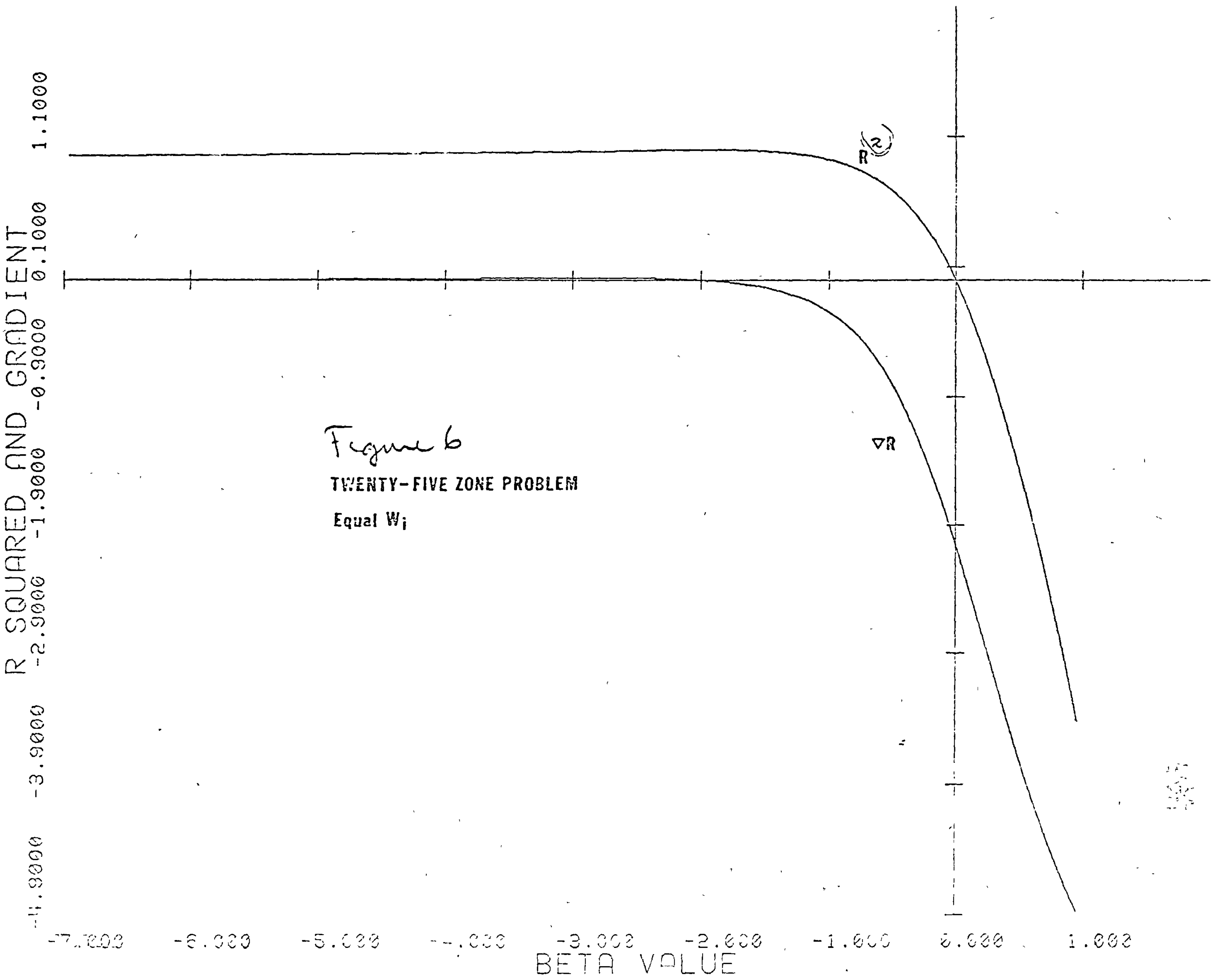


Figure 6

TWENTY-FIVE ZONE PROBLEM

Equal W_i

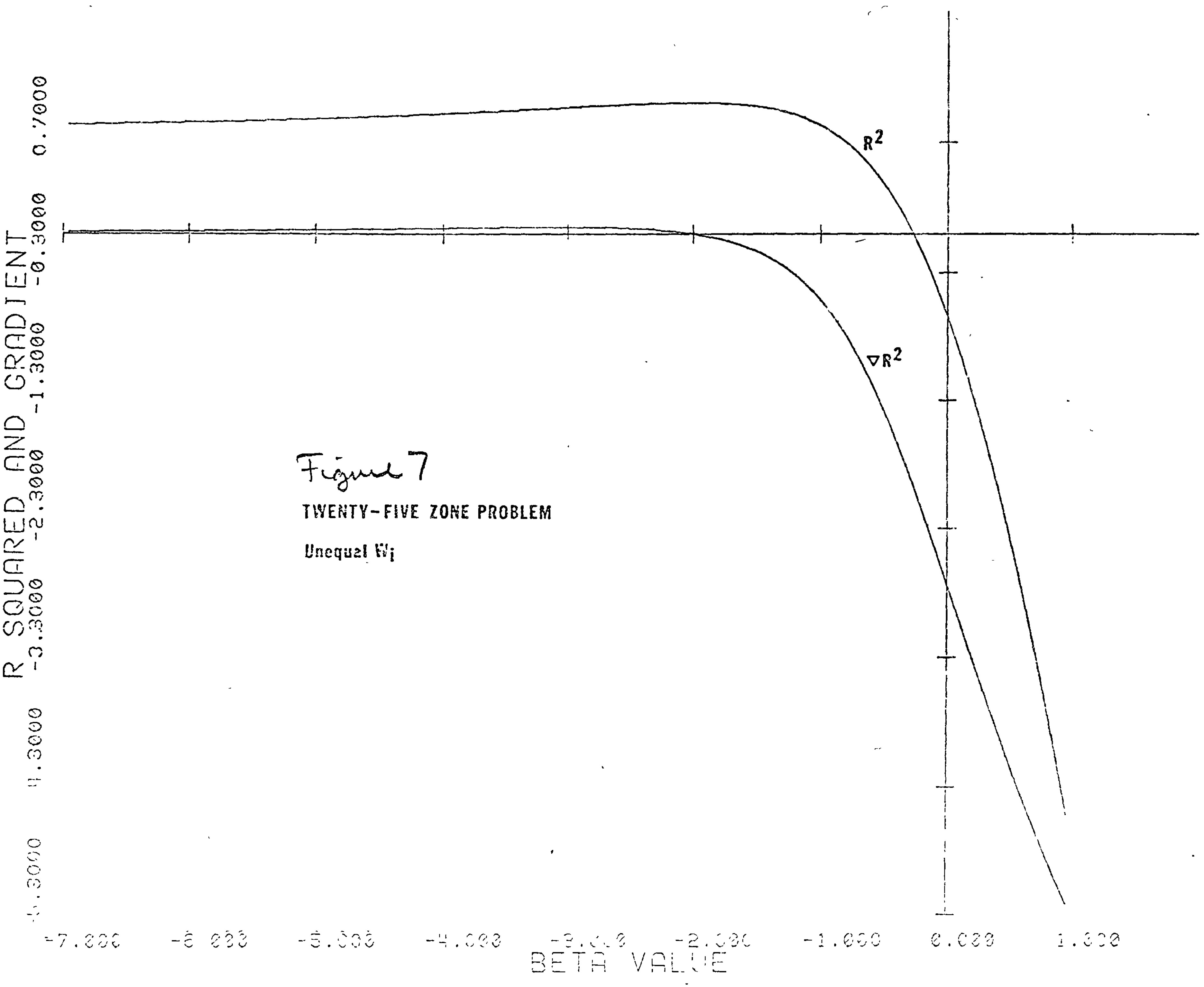


Figure 7

TWENTY-FIVE ZONE PROBLEM

Unequal W_i

convenient to proceed in terms of the logarithm of the likelihood function. Thus

$$\ln L = \sum_i N_i \ln \hat{N}_i \tag{36}$$

We may now construct a likelihood criterion function such that for a perfect fit between observed and estimated we get L = 0. We redefine L such that

$$L = \sum_i N_i \ln \hat{N}_i - \sum_i N_i \ln N_i \tag{37}$$

The problem then is to maximize the value of L (which will have its maximum at C = 0.0). As was the case for R^2, this involves taking the derivative of the criterion function with respect to β.

$$\nabla L = \frac{\partial L}{\partial \beta} = \frac{\partial L}{\partial \hat{N}_i}\left(\frac{\partial \hat{N}_i}{\partial \beta}\right) \tag{38}$$

$$= \sum_i N_i \left(\frac{1}{\hat{N}_i}\right)\left(\frac{\partial \hat{N}_i}{\partial \beta}\right) \tag{39}$$

Referring back to equation (32), we see that the partial derivative of N_i with respect to β is already worked out. Thus, we have for the gradient of the log likelihood criterion the following:

$$\nabla L = \sum_i \left(\frac{N_i}{\hat{N}_i}\right)\left[\sum_j \left(\frac{E_j L_{ij}}{M_j}\right)\left(C_{ij} - \frac{\sum_k C_{kj} L_{kj}}{M_j}\right)\right] \tag{40}$$

In this series of plots, we observe the decreasing effect of the values of the data set with increasing numbers of zones, as we observed using the R^2 criterion. In addition, we observe, if we notice the values on the Likelihood, Gradient axes of the plots, that the curve becomes much more steeply peaked for the 9 and 25 zone examples. While this is not so obvious in Figures 8 through 13, Figure 14 may help to display this phenomenon. In Figure 14,

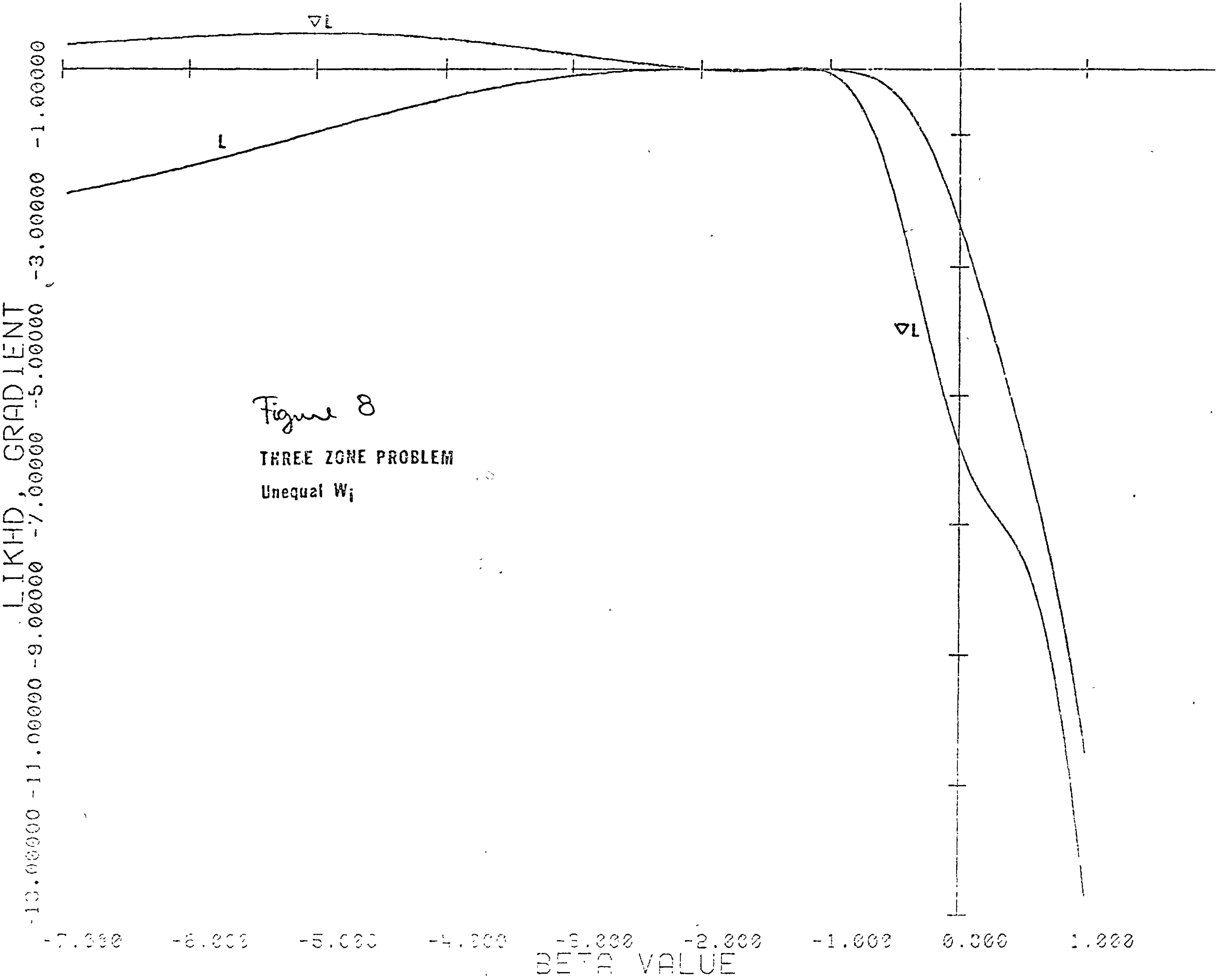


Figure 8

THREE ZONE PROBLEM

Unequal W_i

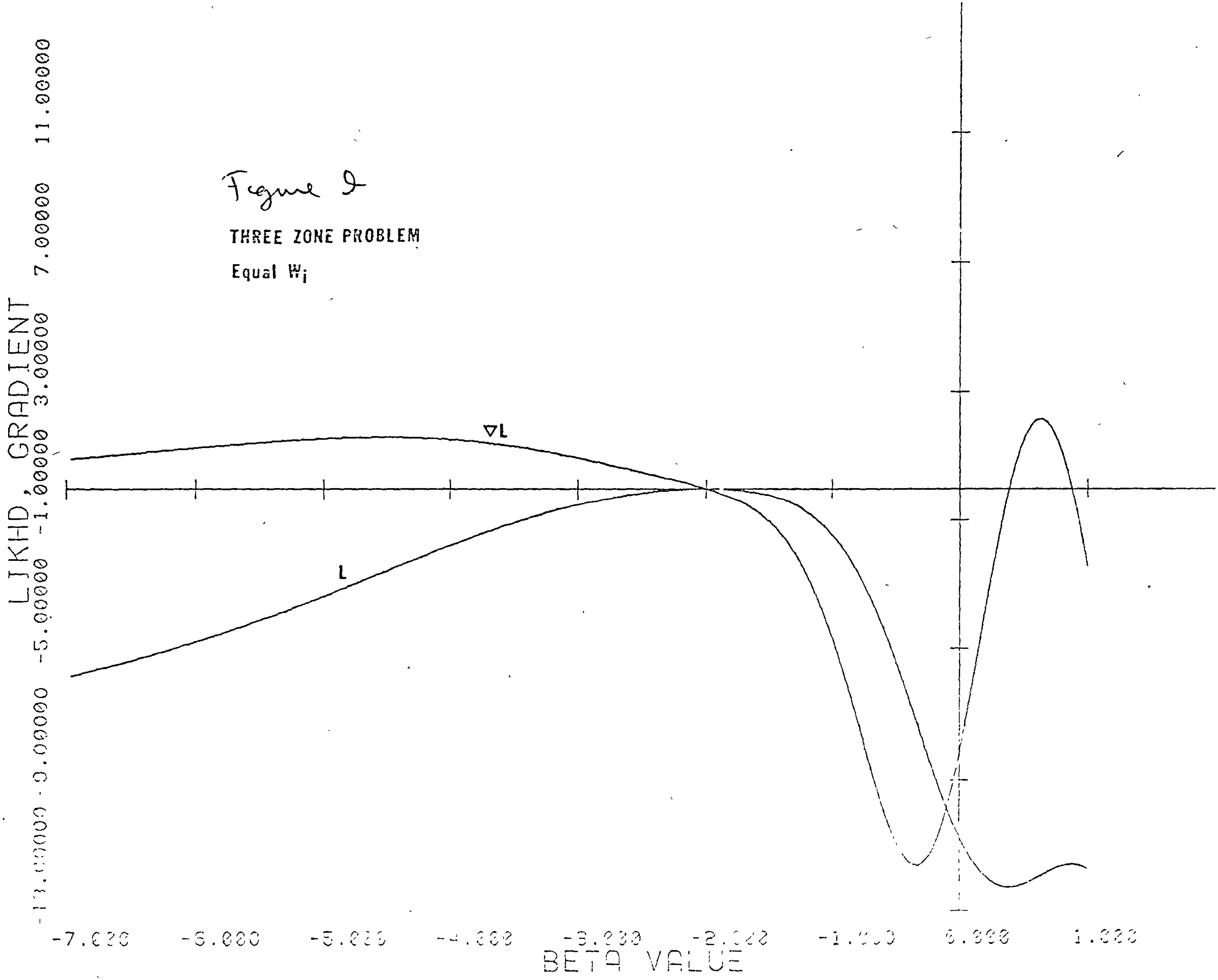


Figure 2

THREE ZONE PROBLEM

Equal W_i

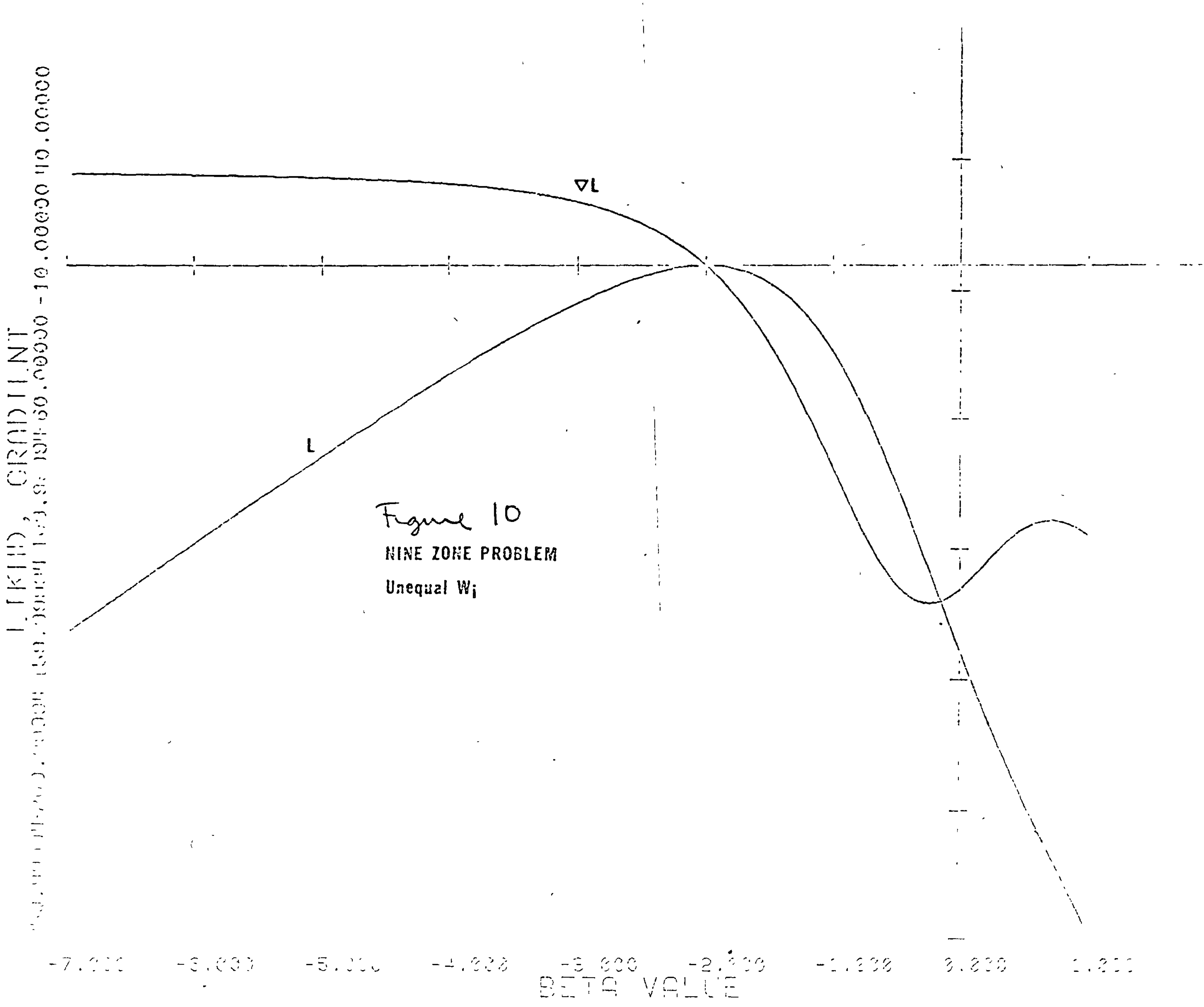


Figure 10

NINE ZONE PROBLEM

Unequal W_i

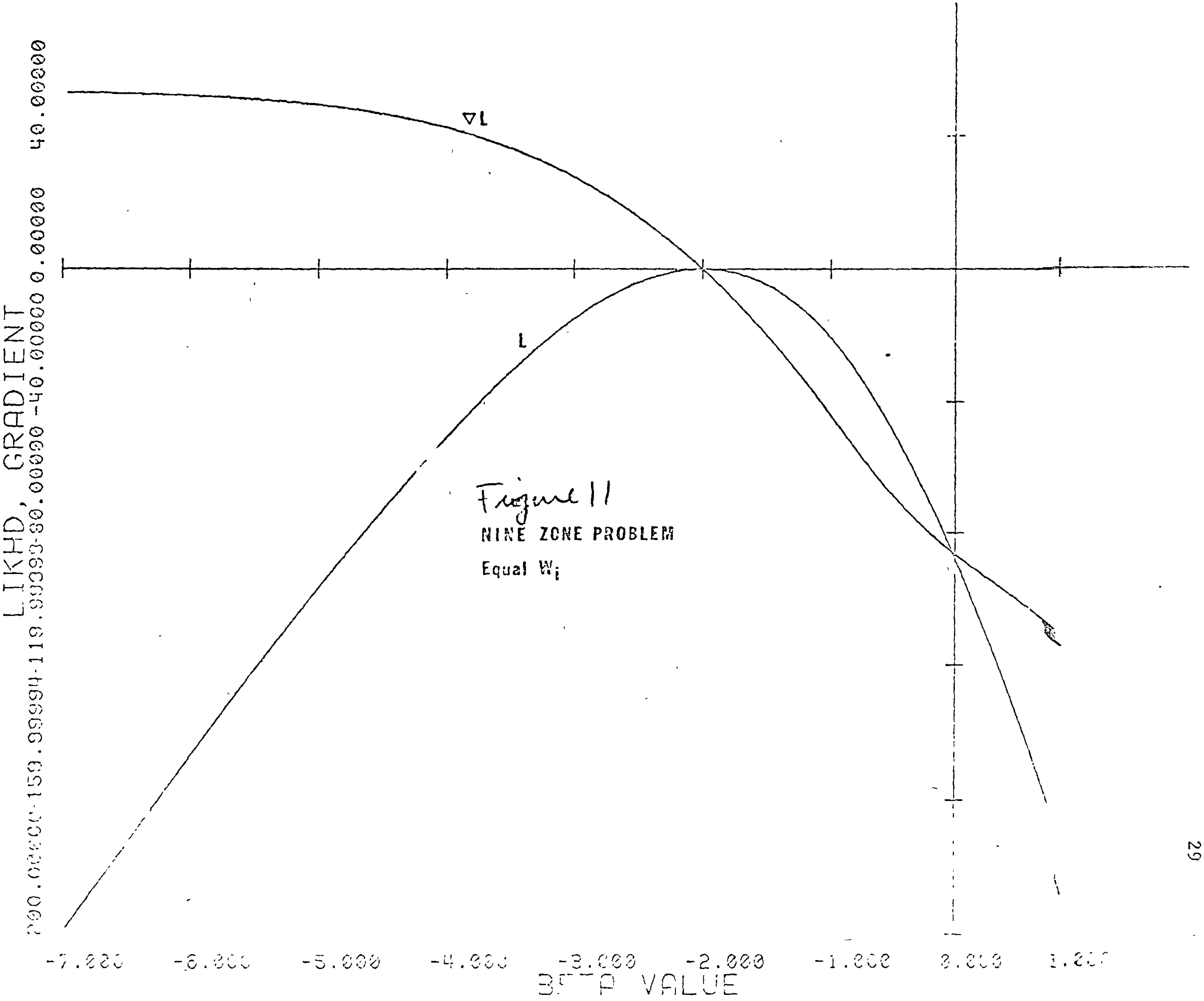


Figure 11

NINE ZONE PROBLEM

Equal W_i

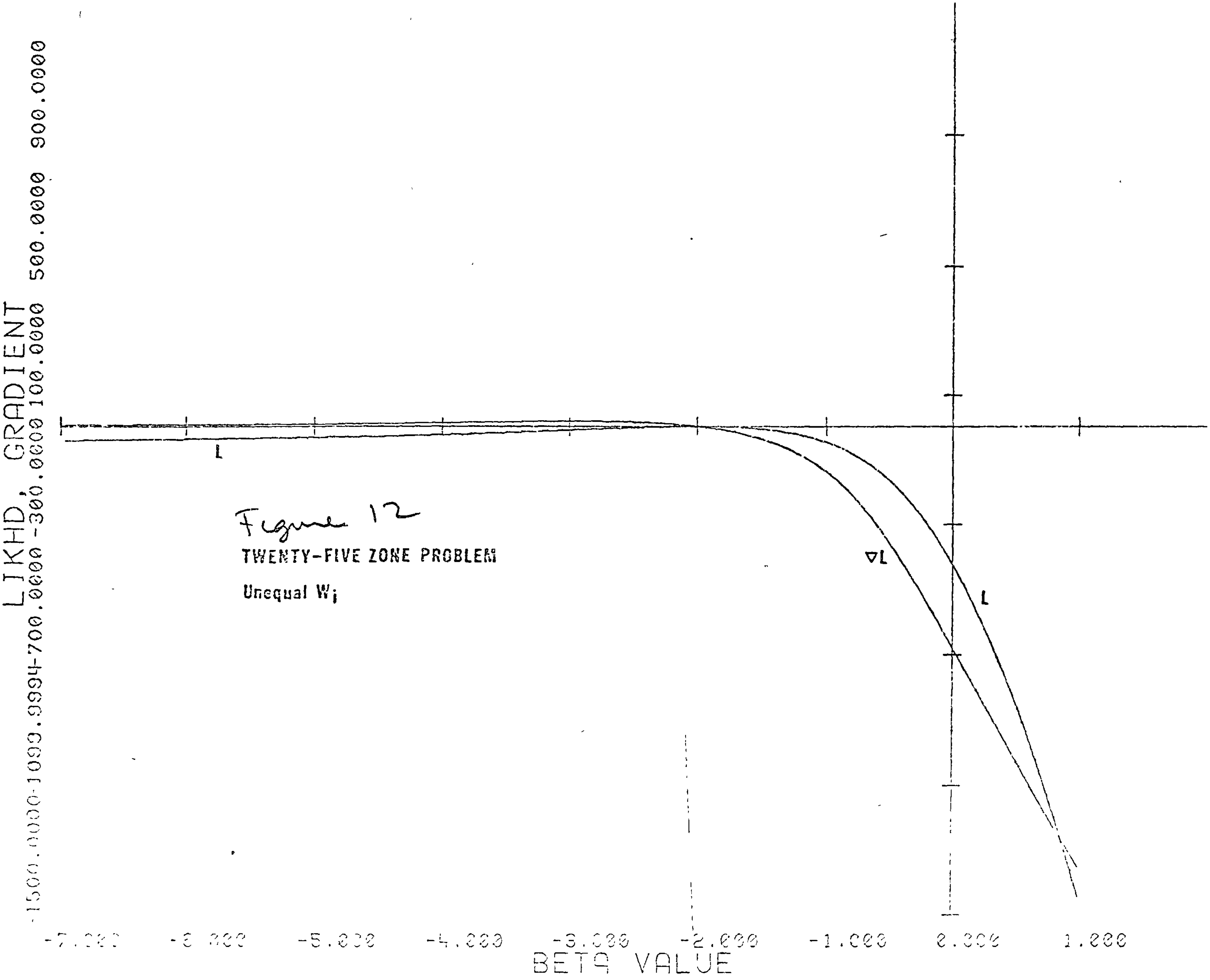


Figure 12

TWENTY-FIVE ZONE PROBLEM

Unequal W_i

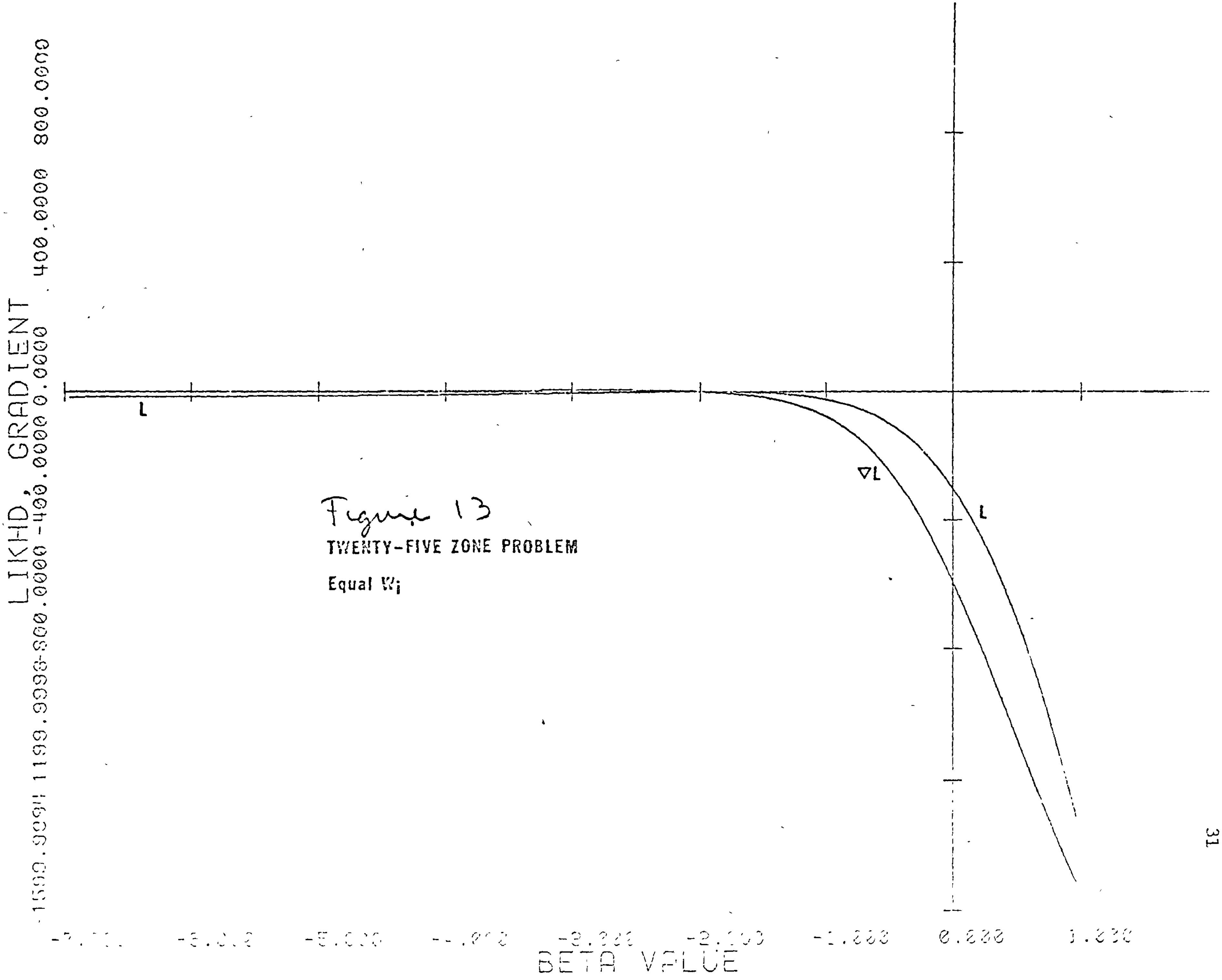
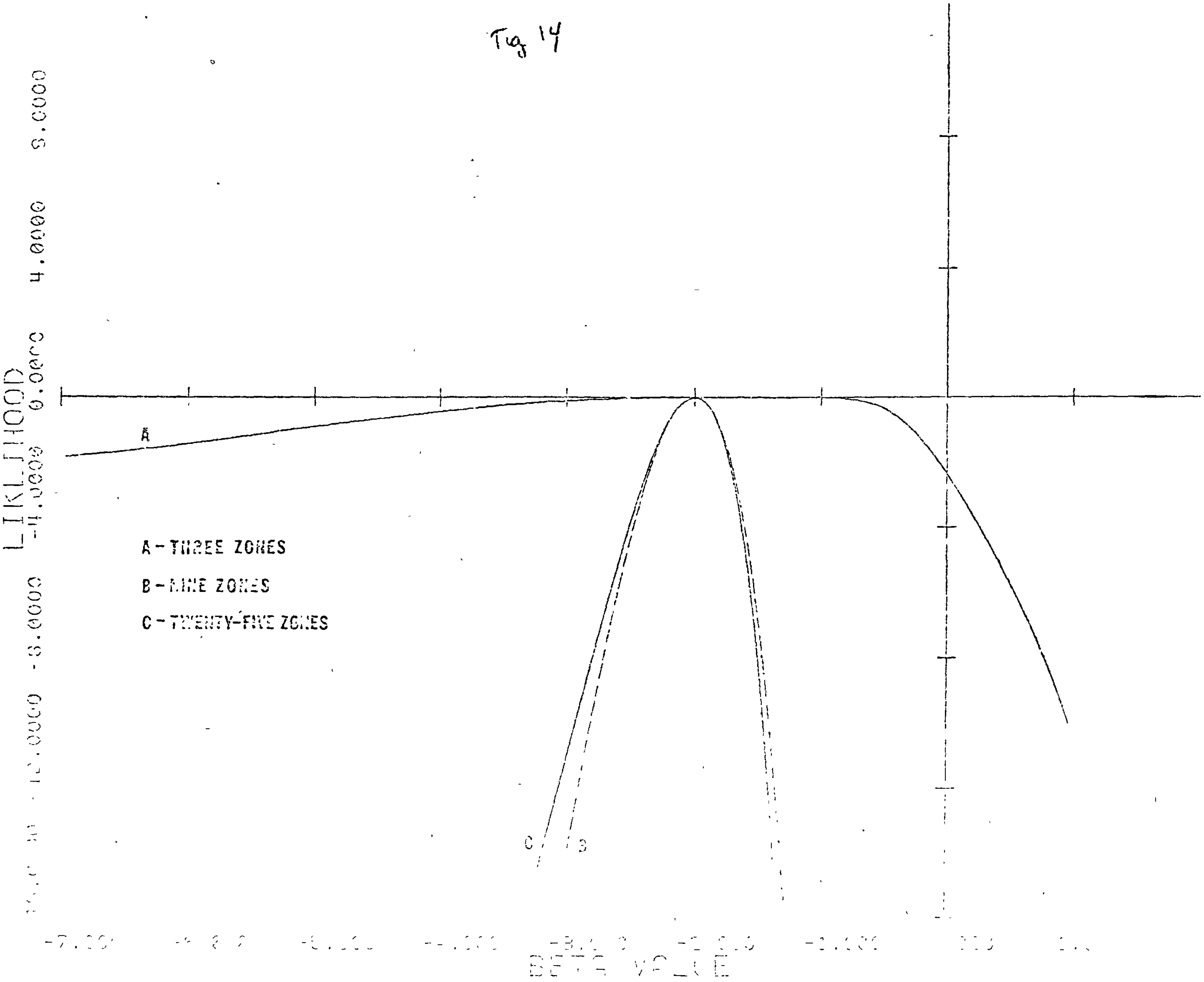


Figure 13

TWENTY-FIVE ZONE PROBLEM

Equal W_i

Fig 14


LIKLIHOOD
8.0000
4.0000
0.0000
-4.0000
-8.0000
-12.0000
BETA VALUE
A
B
C
A - THREE ZONES
B - NINE ZONES
C - TWENTY-FIVE ZONES

the L function is plotted for all three test problems (three zone, nine zone, and twenty-five zone, all with unequal W_i) with the same likelihood axis scaling. Here we see the rather spread-out peak of the 3 zone problem and the much sharper peaks of the 9 zone and 25 zone problems. Note that the latter two peaks overlap, rather than the 25 zone peak bein entirely within the 9 zone peak as expected. This may well be due to the data base differences in the two problems. The result does suggest that the peak of the likelihood function becomes very sharp even for 9 zones, but that it does not get much "sharper" as the number of zones increases beyond that. It would have been interesting to see what would result from a 100 zone problem but it was decided to move on to the somewhat more interesting two and three parameter problems rather than continuing with the one parameter example.

It is possible to draw the following conclusions at this point in the investigation:

1) Different data sets yield different shapes of criterion and gradient functions.
2) These differences become less significant as the number of zones increase and are probably of no consequence at the numbers of zones used for actual applications of the model.
3) The steepness of the function near its optimum is very much greater for the likelihood criterion than for the R^2 criterion, confirming the appropriateness of the use of likelihood for calibration applications.

The Likelihood Criterion Function: Two Parameter Case

The distance or travel function used throughout the examples given above was a simple declining exponential function

$$\mathcal{F}(C_{ij}) = \exp(\beta C_{ij}) \tag{41}$$

A function which has been found to yield better data fits for residential location models is the two parameter Tanner function

$$\mathcal{F}(C_{ij}) = C_{ij}^{\alpha}\exp(\beta C_{ij}) \tag{42}$$

Substituting this function into equation (28), we get

$$\hat{N}_i = \sum_j E_j \left[\frac{W_i C_{ij}^{\alpha}\exp(\beta C_{ij})}{\sum_i W_i C_{ij}^{\alpha}\exp(\beta C_{ij})}\right] \tag{43}$$

This, of course, means that equation (40) for ∇L will be changed as we now require partial derivatives of L with respect to α and with respect to β. We confine further discussion to the likelihood criterion, and derive the necessary derivatives.

To simplify what follows, let

$$B_i = \sum_k C_{ik}^{\alpha}\exp(\beta C_{ik}) \tag{44}$$

Then

$$\hat{N}_j = \sum_i T_i [C_{ij}^{\alpha}\exp(\beta C_{ij})\ B_i^{-1}] \tag{45}$$

Now, we begin taking derivatives, first recalling that

$$\frac{\partial (C_{ij}^{\alpha})}{\partial \alpha} = C_{ij}^{\alpha}\ \ln C_{ij} \tag{46}$$

$$\frac{\partial (\exp(\beta C_{ij}))}{\partial \beta} = C_{ij}\exp(\beta C_{ij}) \tag{47}$$

Now, by the chain rule for derivatives

$$\frac{\partial \hat{N}_j}{\partial \alpha} = \sum_i T_i \left[C_{ij}^{\alpha} \exp(\beta C_{ij}) \left(\frac{\partial B^{-1}}{\partial \alpha} \right) + B^{-1} \left(\frac{\partial \left(C_{ij}^{\alpha} \exp(\beta C_{ij}) \right)}{\partial \alpha} \right) \right] \qquad (48)$$

Continuing to expand and to take derivatives,

$$= \sum_i T_i \left[-B_i^{-2} \left[\sum_k \left(C_{ik}^{\alpha} \left\{ \exp(\beta C_{ik}) \right\} \ln C_{ik} \right) \right] C_{ij}^{\alpha} \exp(\beta C_{ij}) + \left(C_{ij}^{\alpha} \left\{ \exp(\beta C_{ij}) \right\} \ln C_{ij} \right) B_i^{-1} \right] \qquad (49)$$

$$= \sum_i T_i C_{ij}^{\alpha} \left\{ \exp(\beta C_{ij}) \right\} B_i^{-1} \left[\ln C_{ij} - \sum_k \left(C_{ik}^{\alpha} \left\{ \exp(\beta C_{ik}) \right\} \ln C_{ik} \right) B_i^{-1} \right] \qquad (50)$$

$$= \sum_i T_{ij} \left[\ln C_{ij} - \sum_k \left(C_{ik}^{\alpha} \left\{ \exp(\beta C_{ik}) \right\} \ln C_{ik} \right) B_i^{-1} \right] \qquad (51)$$

We next calculate the derivative with respect to β.

$$\frac{\partial \hat{N}_j}{\partial \beta} = \sum_i T_{ij} \left[C_{ij} - \sum_k C_{ik}^{\alpha} \left\{ \exp(\beta C_{ik}) \right\} C_{ik} B_i^{-1} \right] \qquad (52)$$

$$= \sum_i T_{ij} \left[C_{ij} - \sum_k C_{ik}^{\alpha+1} \left\{ \exp(\beta C_{ik}) \right\} B_i^{-1} \right] \qquad (53)$$

Finally, we may write the derivatives of the criterion with respect to each parameter.

Referring to equation (11), we have

$$\frac{\partial L}{\partial \alpha} = \frac{\partial L}{\partial \hat{N}_i}\left(\frac{\partial \hat{N}_i}{\partial \alpha}\right) \tag{54}$$

and

$$\frac{\partial L}{\partial \beta} = \frac{\partial L}{\partial \hat{N}_i}\left(\frac{\partial \hat{N}_i}{\partial \beta}\right) \tag{55}$$

so, by substitution from equations (39), (51), and (54)

$$\frac{\partial L}{\partial \alpha} = \sum_i \left(\frac{N_i}{\hat{N}_i}\right)\left(\frac{\partial \hat{N}_i}{\partial \alpha}\right) \tag{56}$$

$$= \sum_i \left(\frac{N_i}{\hat{N}_i}\right)\left[\sum_i T_{ij}\left[\ln C_{ij} - \sum_k \left(C_{ik}^{\alpha}\left\{\exp(\beta C_{ik})\right\} \ln C_{ik}\right) B_i^{-1}\right]\right] \tag{57}$$

and, again by substitution from equations (39), (53), and (55)

$$\frac{\partial L}{\partial \beta} = \sum_i \left(\frac{N_i}{\hat{N}_i}\right)\left(\frac{\partial \hat{N}_i}{\partial \beta}\right) \tag{58}$$

$$= \sum_i \left(\frac{N_i}{\hat{N}_i}\right)\left[\sum_i T_{ij}\left[C_{ij} - \sum_k C_{ik}^{(\alpha+1)}\left\{\exp(\beta C_{ik})\right\} B_i^{-1}\right]\right] \tag{59}$$

Holding these calculations in abeyance, we may do some explorations of the shape of the likelihood function surface using the Tanner function in place of the simple declining exponential function. First, using the same nine zone problem from before, we assign the values $\alpha = 2.0$, $\beta = -1.5$ and calculate a new set of N_i which we know will fit the model perfectly at the given parameter values. Next we plot the values of the log likelihood function against values of the parameters to get an impression of the shape of the surface.

In Figure 15 is a plot of L vs. β for specified values of α. Here we see a series of relatively sharp peaks whose optima are not strikingly different. These optima in the L vs. β function are clearly shifted about by changes in the value of α. That each curve is rather like the next suggests that we are dealing with a narrow ridge as the shape of the L surface.

In Figure 16 is a plot of L vs. α for specified values of β. This plot is somewhat less clear than the previous one. What we have is, again, a series of peaks, this time much less sharp and greatly shifted about by changes in β. Thus we see only the left hand portion of the peaks where β = -3.0, -2.5, -2.0, and -1.5, while we see only the right hand portions where β = -1.0, -0.5, and zero. While this confuses the picture a bit, it is quite clear that the values of α and β are interrelated. Additionally it seems that within a moderate range of the perfect fit parameters, while variation either in α or β will result in a variation in the other in determining the function's optimum, the actual value of the optimum (in terms of L) is not very different.

All this may be made more understandable by reference to Figure 17 where a contour plot of the L surface in terms of α and β is given. Here the curves shown in Figure 15 are cross-sectional slices through the surface at fixed values of β. Thus it is readily seen that we are dealing, in the case of the Tanner function, with a rather long and narrow ridge which slopes very gently along its long axis, and rather sharply across its short axis. As a consequence of knowing the shape of this surface we clearly demonstrate the considerable relationship between the two parameters of the Tanner function. From one given value of α to another, the likelihood maximizing value of β changes substantially. While the change is less in α from one given value of β to another, it is still

Fig. 15

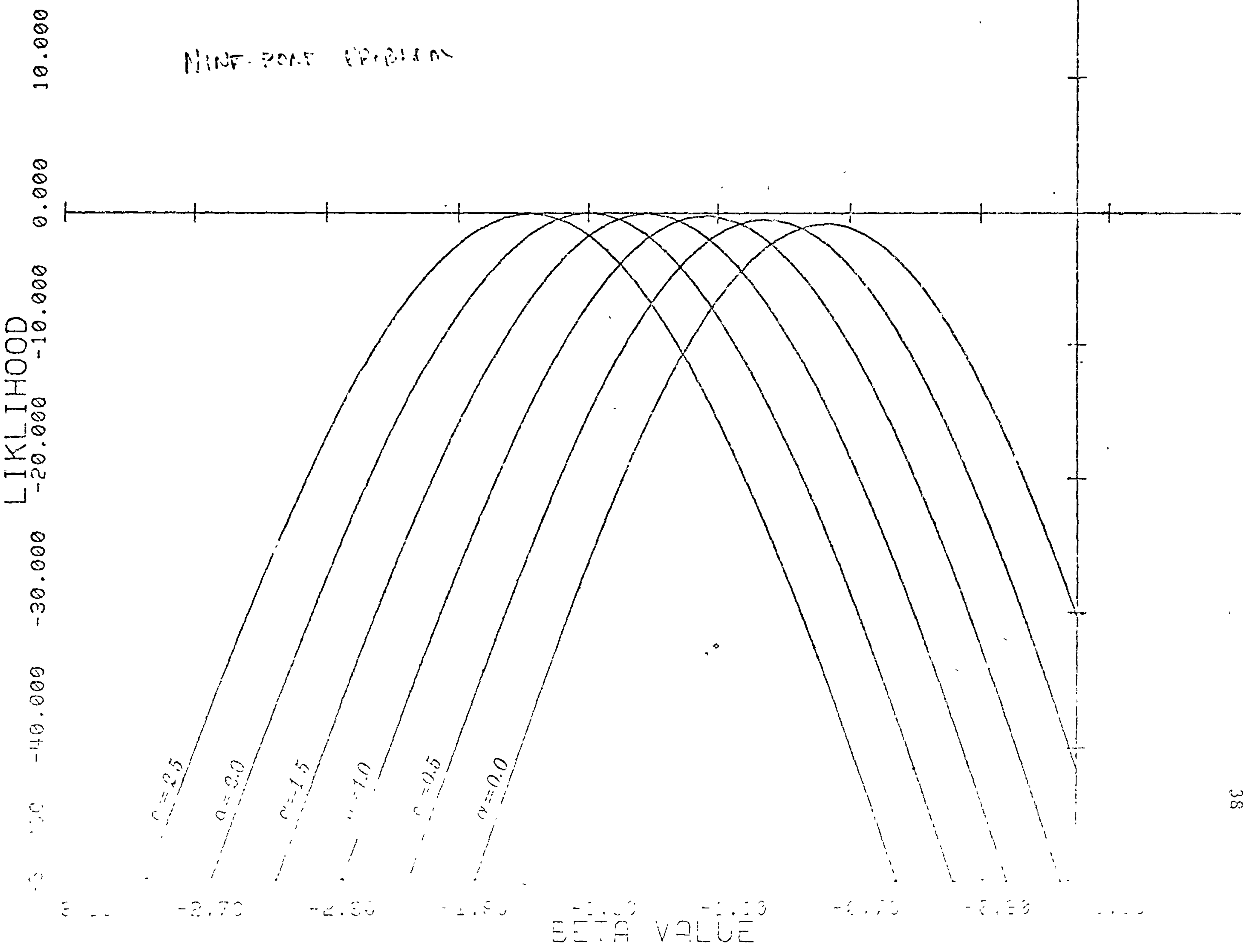
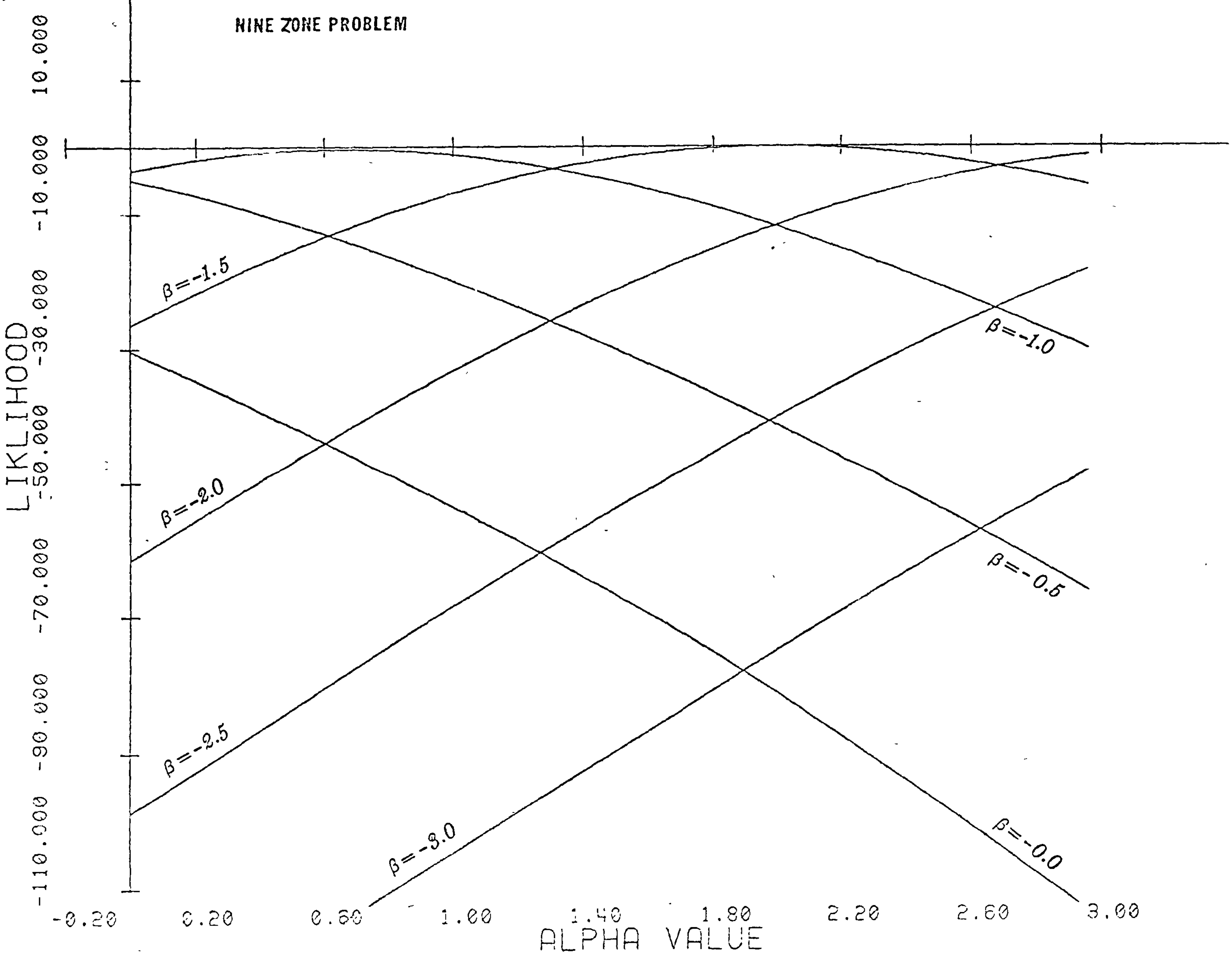

NINE-POINT PROBLEM
LIKLIHOOD
10.000
0.000
-10.000
-20.000
-30.000
-40.000
-50.000
α=2.5
α=2.0
α=1.5
α=1.0
α=0.5
α=0.0
-2.70
-2.30
-1.90
-1.50
-1.10
-0.70
-0.30
BETA VALUE

Fig 16


NINE ZONE PROBLEM
LIKLIHOOD
10.000
-10.000
-30.000
-50.000
-70.000
-90.000
-110.000
ALPHA VALUE
-0.20
0.20
0.60
1.00
1.40
1.80
2.20
2.60
3.00
β=-1.5
β=-2.0
β=-2.5
β=-3.0
β=-1.0
β=-0.5
β=-0.0

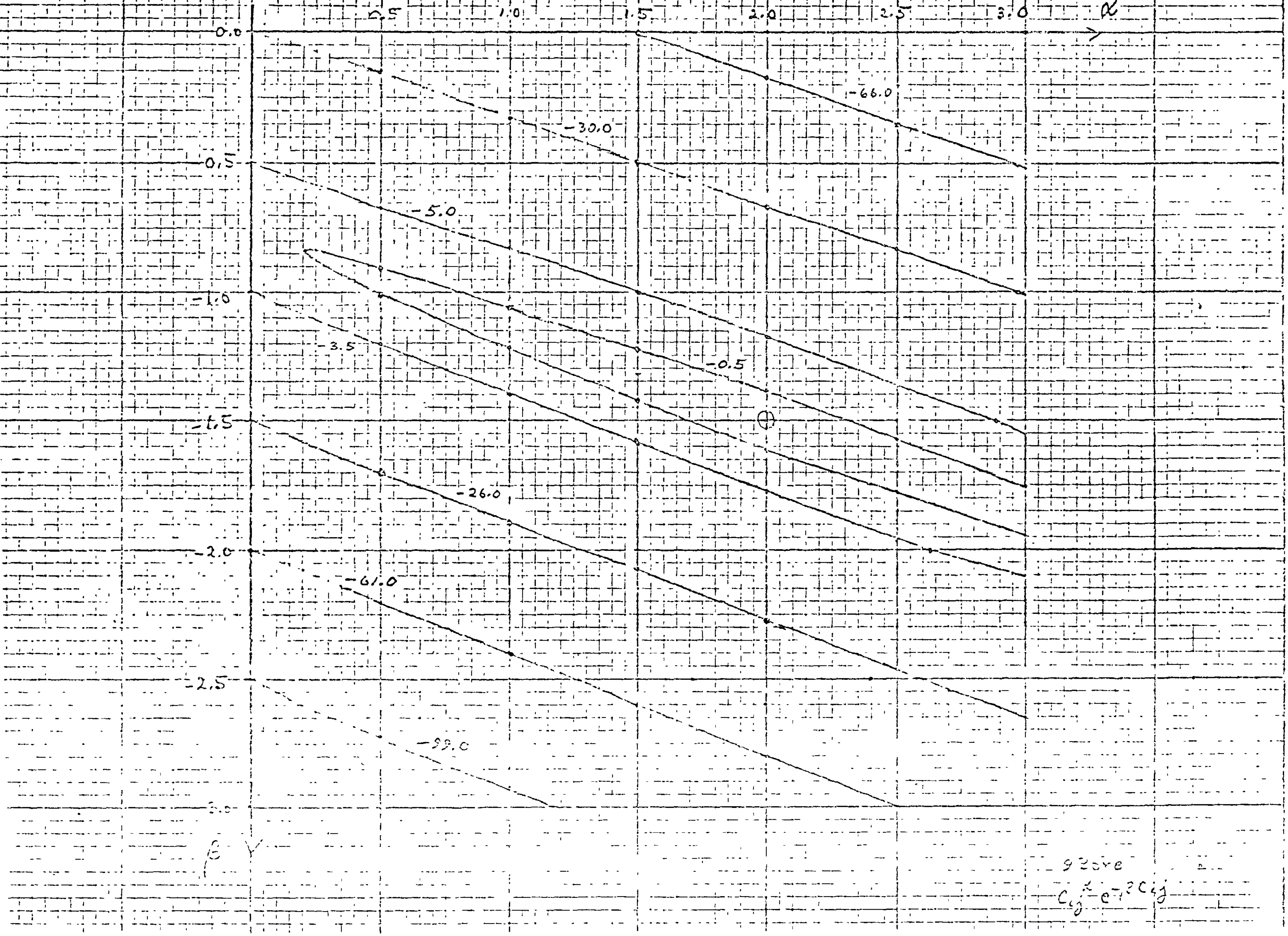

0.5
1.0
1.5
2.0
2.5
3.0
α
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
-66.0
-30.0
-5.0
-0.5
-3.5
θ
-26.0
-61.0
-99.0
β
9 zone
$C_{ij}^2 - e^{-2C_{ij}}$

Fig 17

significant. From this we draw two conclusions before proceeding. First, any calibration of a model using the Tanner function must involve simultaneous estimation of α and β. Second, a simultaneous search technique such as gradient search will be more efficient than a sequential search procedure in finding the optimal values of α and β.

The Likelihood Criterion Function: Three Parameter Case.

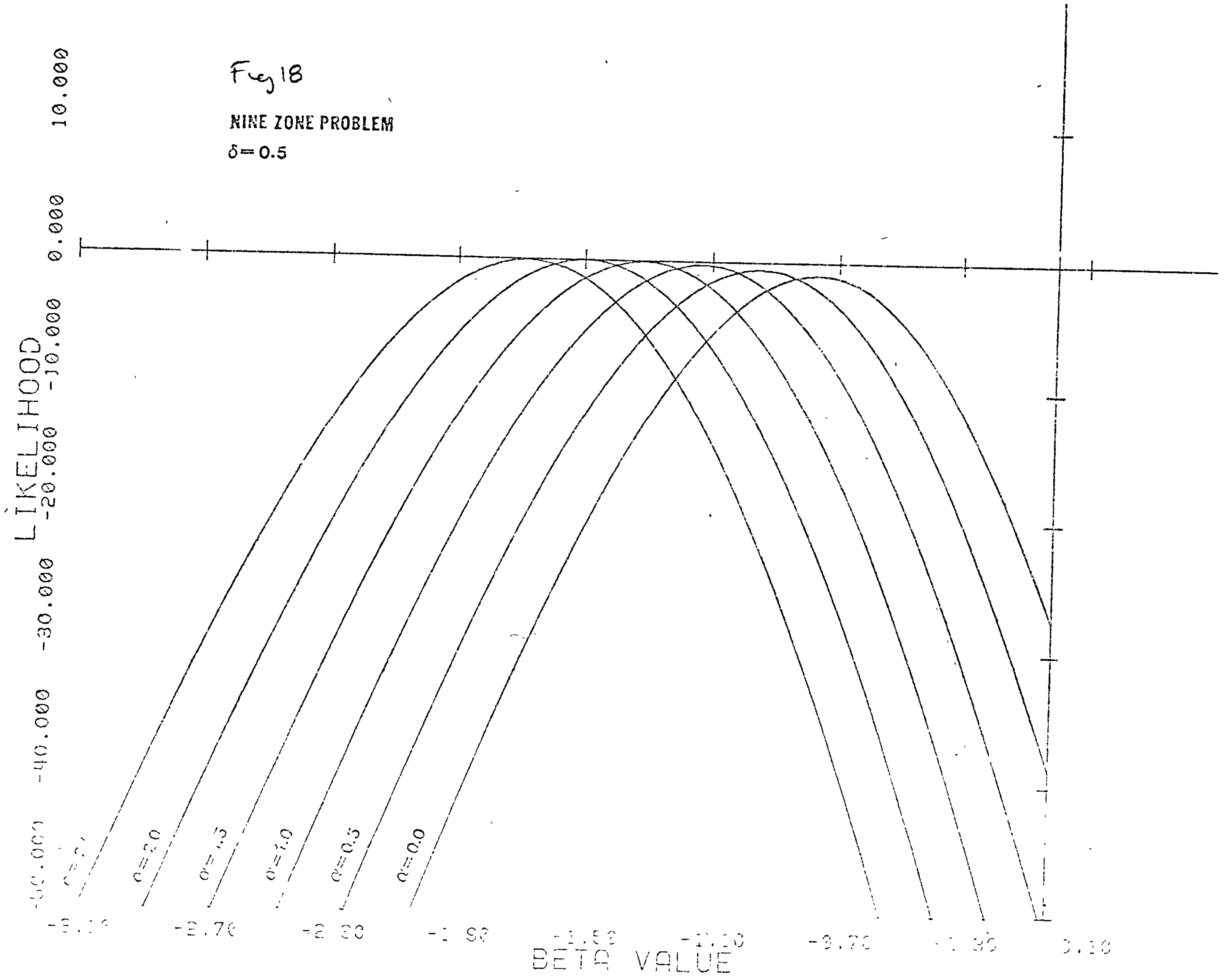
In the prior discussion, the attractiveness measure W_i has been used with 1.0 as an implicit exponent. Recent work with residential location models (Putman, 1977) suggests that it may be useful to consider the exponent of W_i as a parameter to be estimated empirically. Further, W_i may itself be a multivariate function with several parameters to be estimated. The question to be asked here is to what extent will the introduction of additional parameters will alter the shape of the criterion surface?

To investigate this question the equation for estimating zonal residents was altered by the addition of a parametric exponent to the attractiveness measure. Thus equation (43) becomes

$$\hat{N}_i = \sum_j E_j \left[\frac{W_i^{\delta} \, C_{ij}^{\alpha} \, \exp(\beta \, C_{ij})}{\sum_i W_i^{\delta} \, C_{ij}^{\alpha} \, \exp(\beta \, C_{ij})} \right] \tag{60}$$

The "observed" values of residents in zone i were then calculated using equation (60) with $\delta = 0.5$, 3.0, and -3.0 and with $\alpha = 2.0$ and $\beta = -1.5$ in all cases. Then a series of plots were drawn, analogous to Figures 15 and 16. In Figure 18 is shown the L vs. β curves for $\delta = 0.5$ and specified values of α. This figure does not differ markedly from Figure 15.

Fig 18


NINE ZONE PROBLEM
δ=0.5
LIKELIHOOD
10.000
0.000
-10.000
-20.000
-30.000
-40.000
-50.000
α=2.
α=2.0
α=1.5
α=1.0
α=0.5
α=0.0
-3.10
-2.70
-2.30
-1.90
-1.50
-1.10
-0.70
-0.30
0.10
BETA VALUE

similarly, Figure 19 shows L vs. α curves for $\delta = 0.5$ and specified values of β. This figure is similar to Figure 16. Then for δ values of 3.0 and -3.0 the same plots are repeated in Figures 20 to 23. From all these it may be seen that adding the δ parameter to the model does shift the optimal values of α and β, but does not alter the general shapes of the response surfaces between ln L and α and between ln L and β.

The next question is what shape response surface exists between ln L and δ for various values of α and β. In Figure 24 we plot ln L vs. δ for a fixed value of $\beta = -1.5$ and specified values of α. We have what appears to be another long sharp ridge on the surface. In Figure 25 we plot ln L vs. δ for $\alpha = 2.0$ and specified values of β. In Figure 26 we plot the ln L surface in terms of β vs. δ, and in Figure 27 the ln L surface in terms of α vs. δ. In both cases the abovementioned ridges may be clearly seen.

Having done these explorations of the shape of the ln L surface, we may return to the question of gradient search determination of best fit parameters.

In adding the additional parameter δ to the model we must add an additional derivative calculation to get the partial derivative of ln L with respect to δ. Without going through all the steps, it is clear that this derivative will be of the same form as the one for α (equation 51). So we may write

$$\frac{\partial \hat{N}_j}{\partial \delta} = \sum_i T_{ij} \left[\ln W_i - \sum_k \left(C_{ik}^{\alpha} \left\{ \exp(\beta C_{ij}) \right\} \ln W_i \right) B_i^{-1} \right] \qquad (61)$$

Fig 19

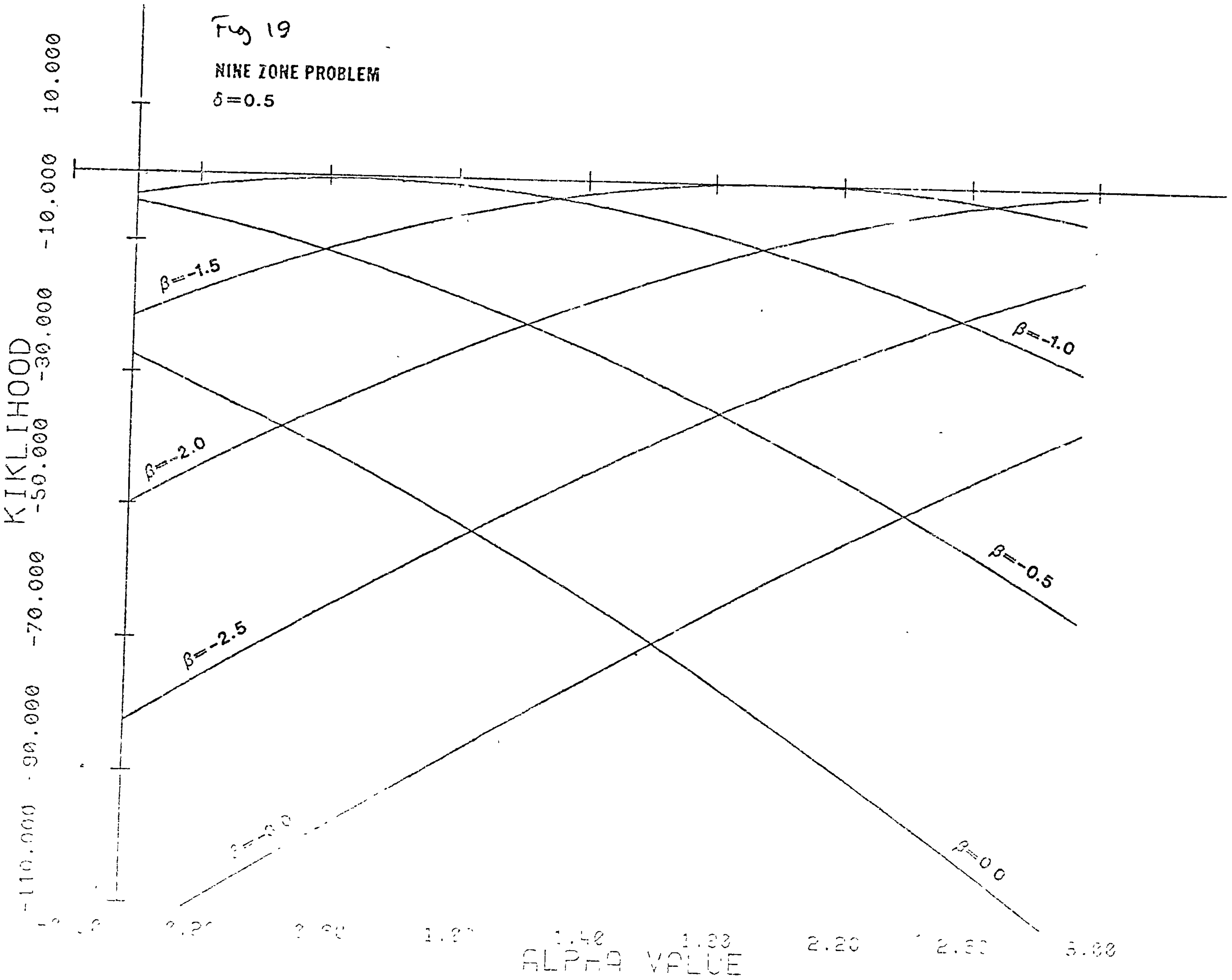

NINE ZONE PROBLEM
δ=0.5
10.000
-10.000
-30.000
-50.000
-70.000
-90.000
-110.000
KIKLIHOOD
β=-1.5
β=-2.0
β=-2.5
β=-3.0
β=-1.0
β=-0.5
β=0.0
-0.20
0.20
0.60
1.00
1.40
1.80
2.20
2.60
3.00
ALPHA VALUE

Fig 20

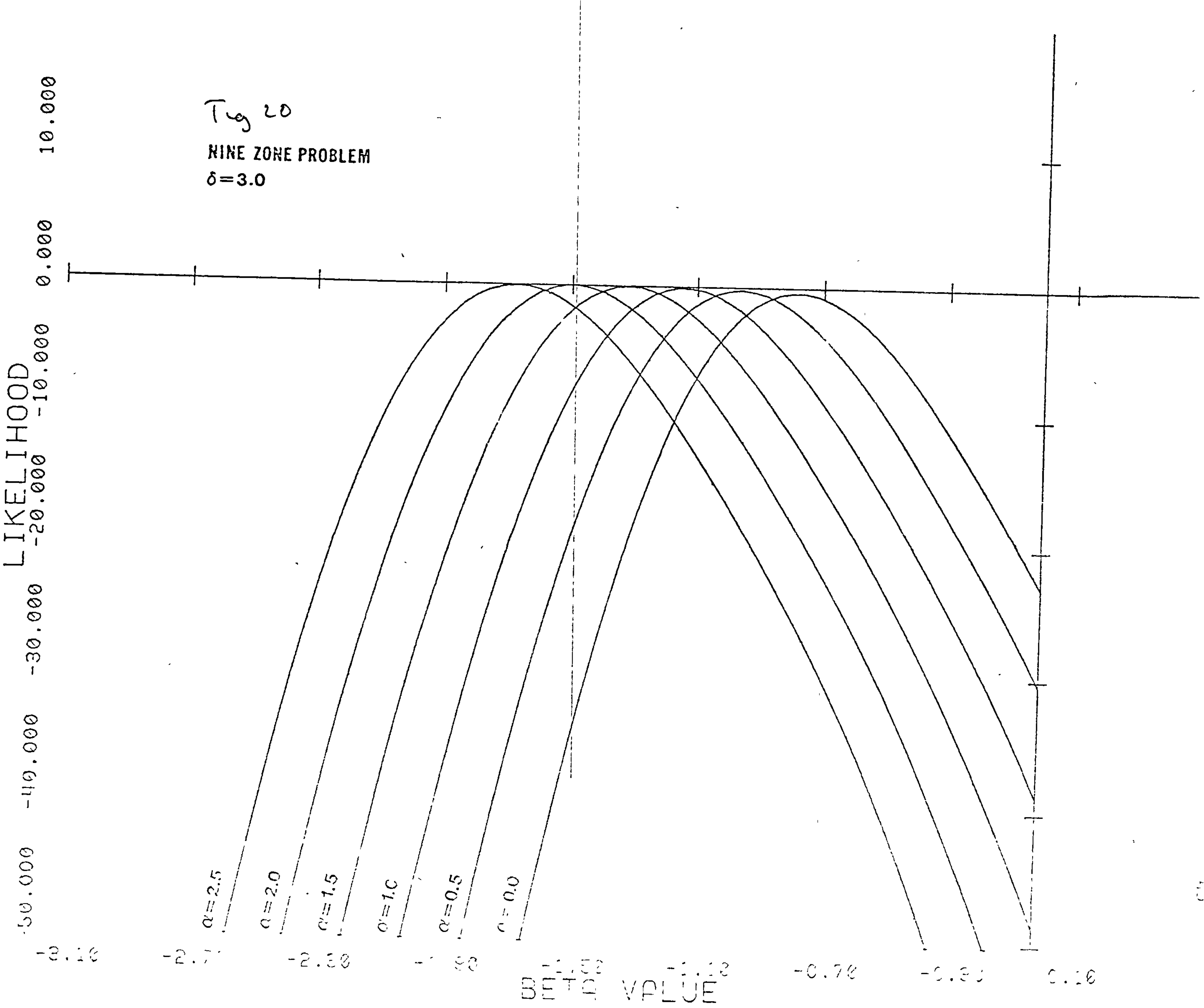

NINE ZONE PROBLEM
$\delta=3.0$
LIKELIHOOD
10.000
0.000
-10.000
-20.000
-30.000
-40.000
-50.000
α=2.5
α=2.0
α=1.5
α=1.0
α=0.5
α=0.0
-3.10
-2.70
-2.30
-1.90
-1.50
-1.10
-0.70
-0.30
0.10
BETA VALUE

Fig 21

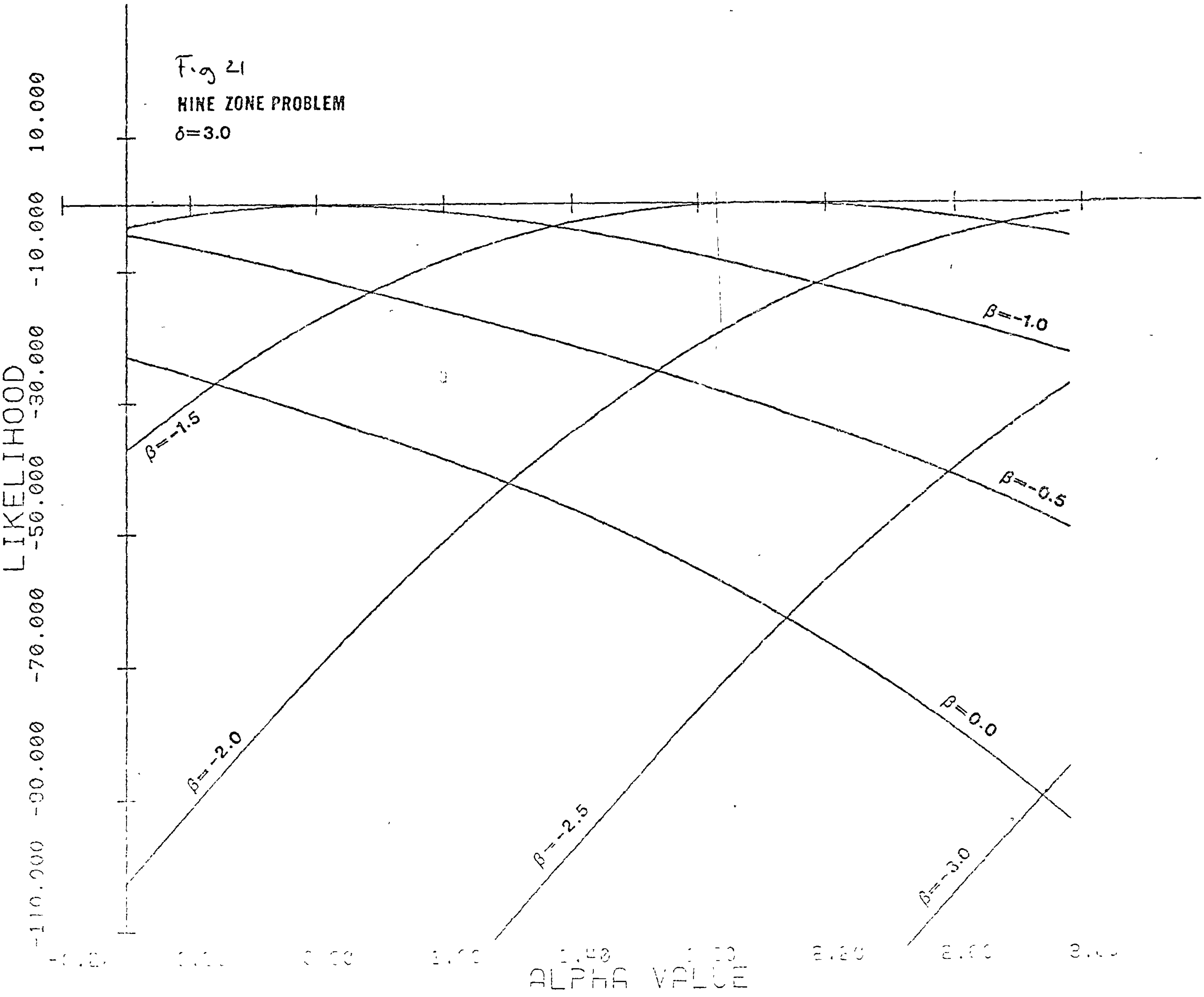
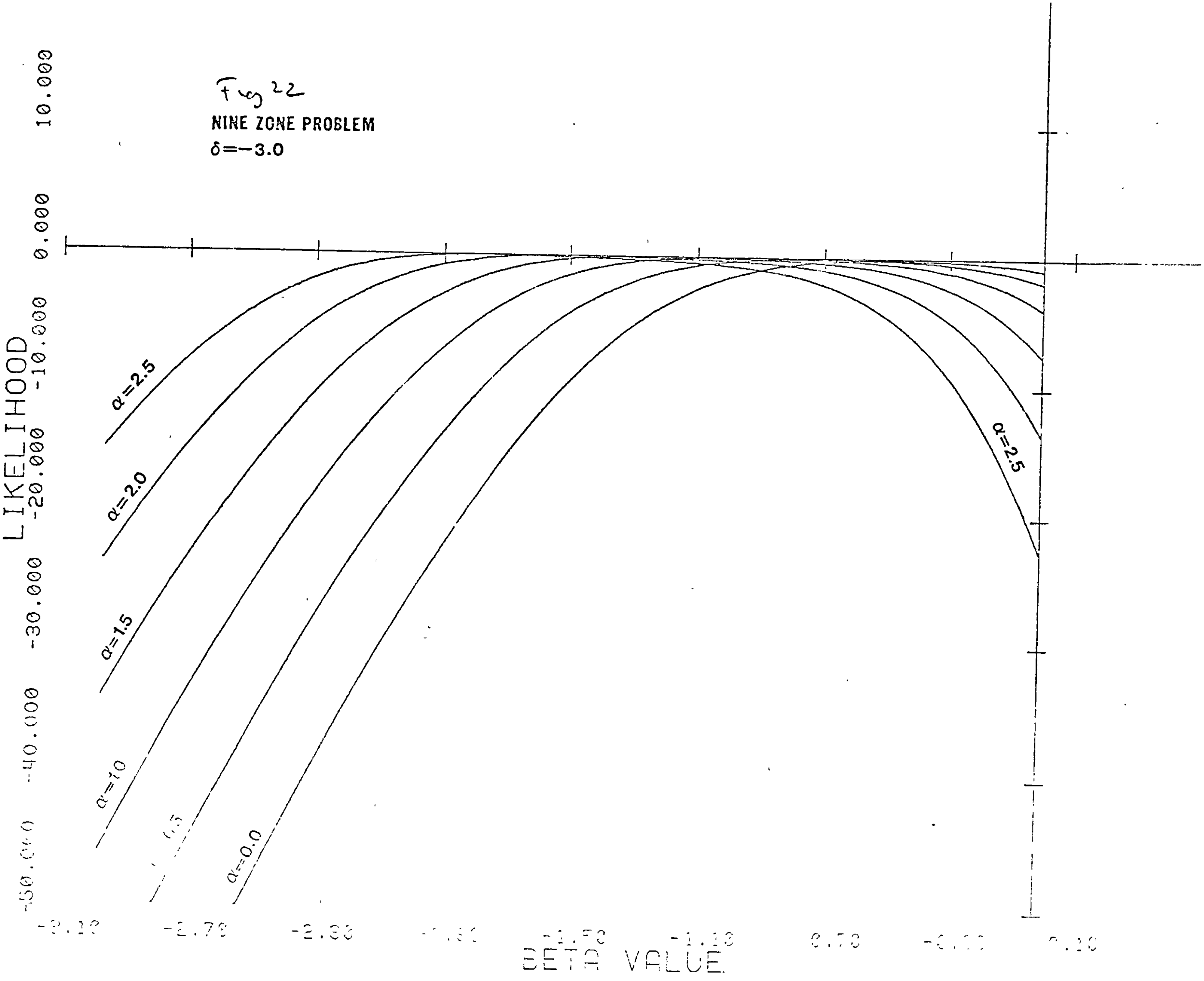

NINE ZONE PROBLEM
δ=3.0
LIKELIHOOD
10.000
-10.000
-30.000
-50.000
-70.000
-90.000
-110.000
-0.20
0.20
0.60
1.00
1.40
1.80
2.20
2.60
3.00
ALPHA VALUE
β=-1.0
β=-1.5
β=-0.5
β=0.0
β=-2.0
β=-2.5
β=-3.0

Fig 22


NINE ZONE PROBLEM
δ=−3.0
LIKELIHOOD
10.000
0.000
-10.000
-20.000
-30.000
-40.000
-50.000
BETA VALUE
-3.10
-2.70
-2.30
-1.90
-1.50
-1.10
-0.70
-0.30
0.10
α=2.5
α=2.0
α=1.5
α=1.0
0.5
α=0.0
α=2.5

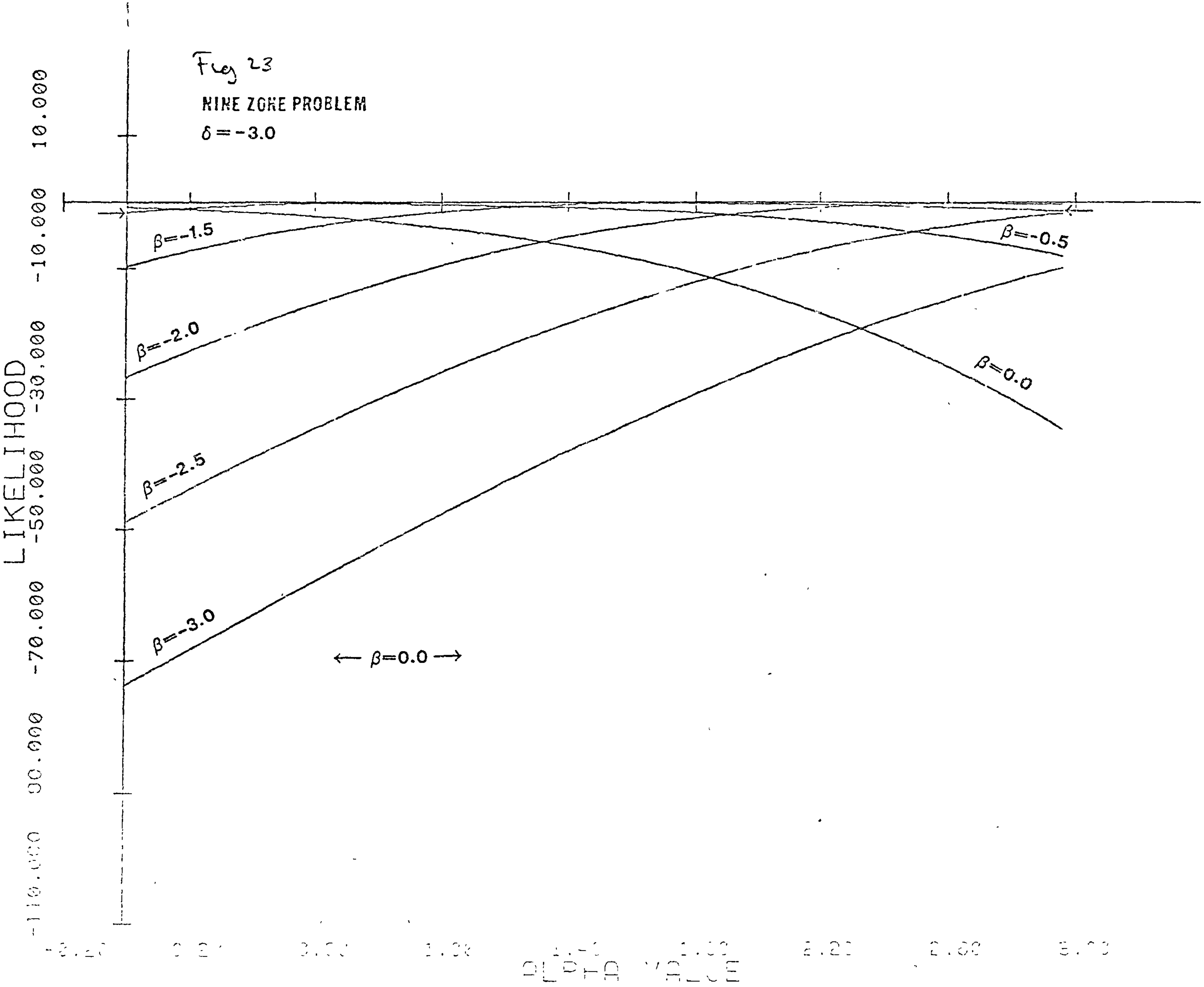


Fig 23

NINE ZONE PROBLEM
$\delta = -3.0$
$\beta=-1.5$
$\beta=-0.5$
$\beta=-2.0$
$\beta=0.0$
$\beta=-2.5$
$\beta=-3.0$
← $\beta=0.0$ →
LIKELIHOOD
10.000
-10.000
-30.000
-50.000
-70.000
-90.000
-110.000
-0.20
0.20
0.60
1.00
1.40
1.80
2.20
2.60
3.00
ALPHA VALUE

Fig 24

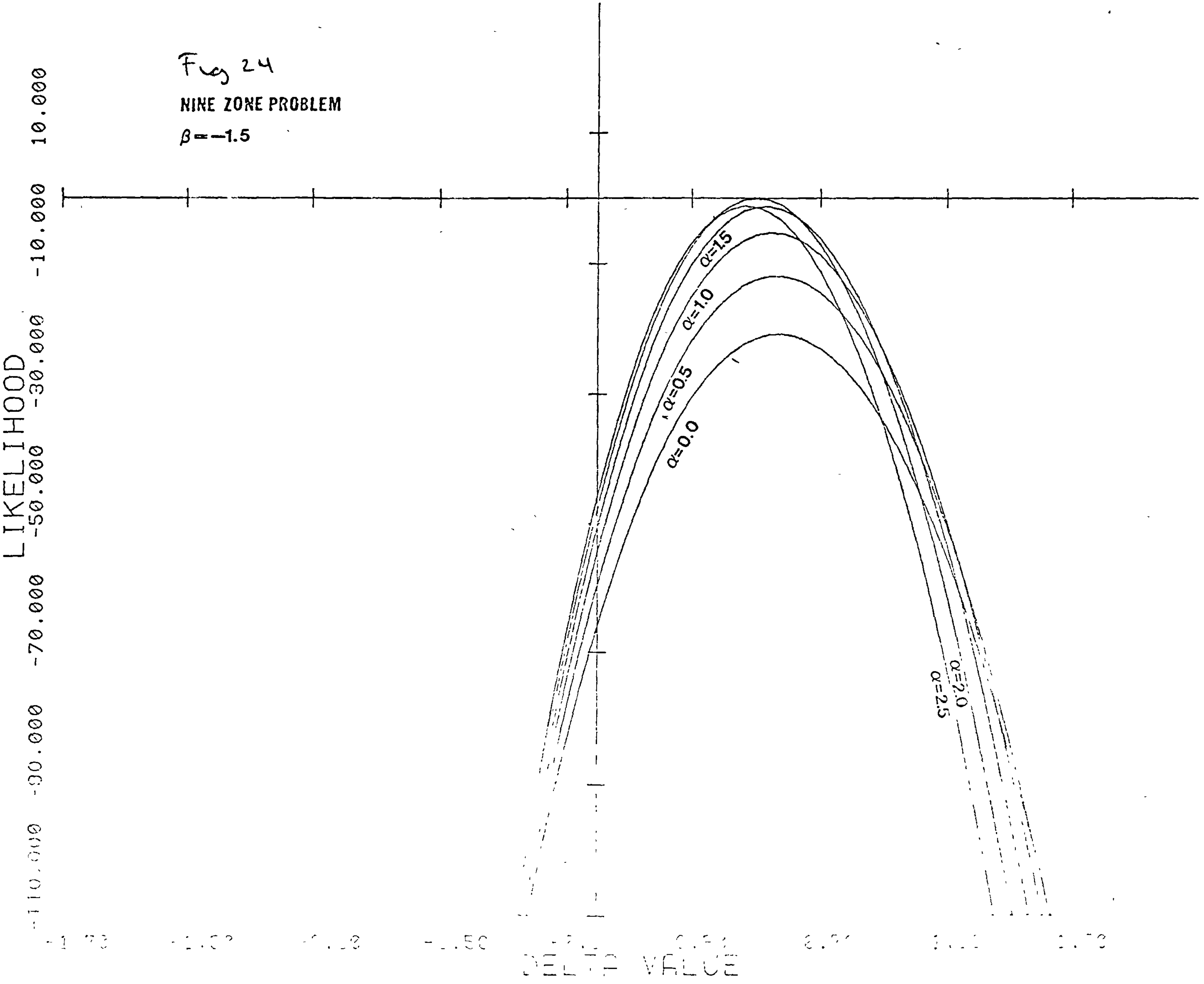
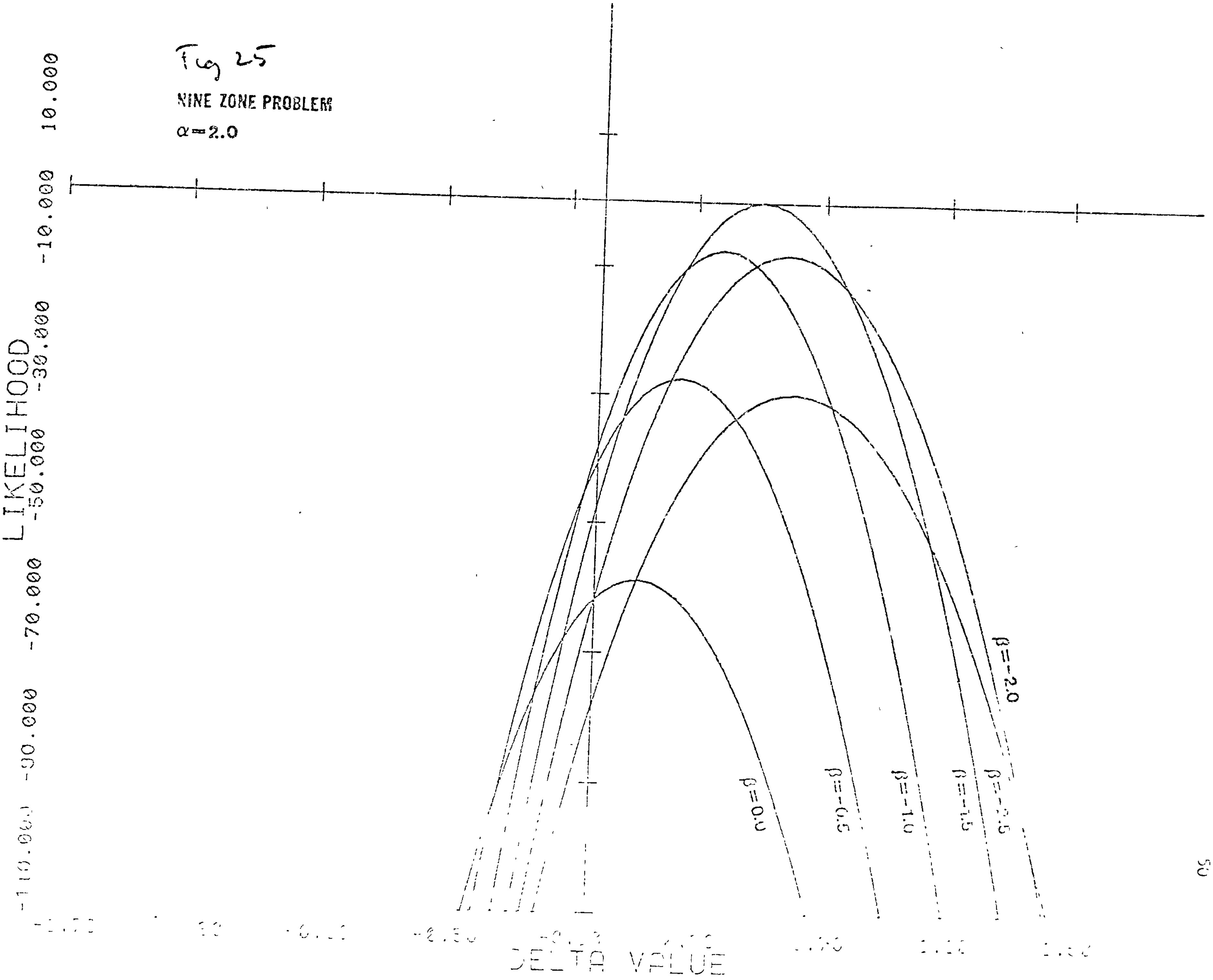
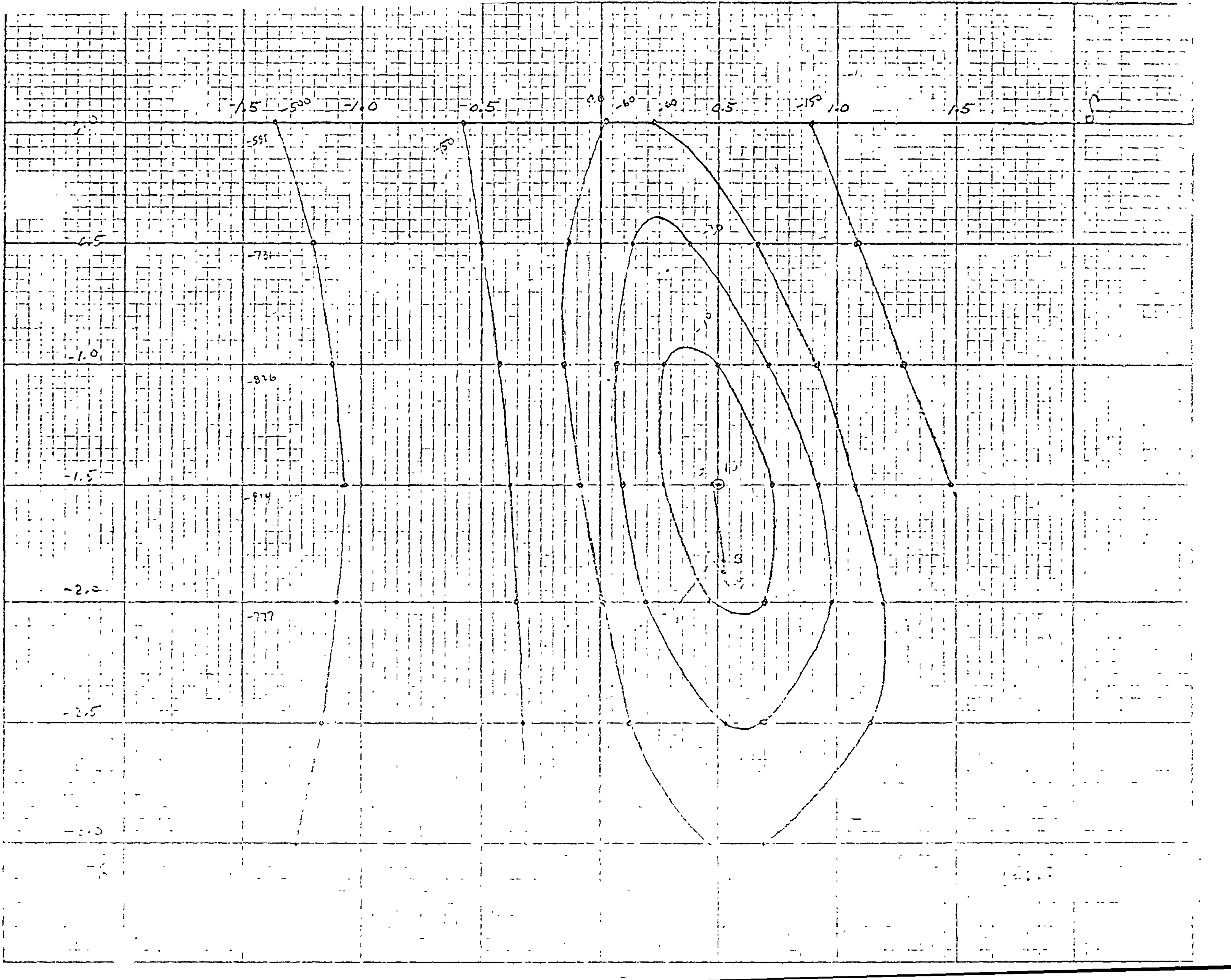

NINE ZONE PROBLEM
β=−1.5
LIKELIHOOD
10.000
-10.000
-30.000
-50.000
-70.000
-90.000
-110.000
α=1.5
α=1.0
α=0.5
α=0.0
α=2.0
α=2.5
-1.70
-1.50
1.70
DELTA VALUE

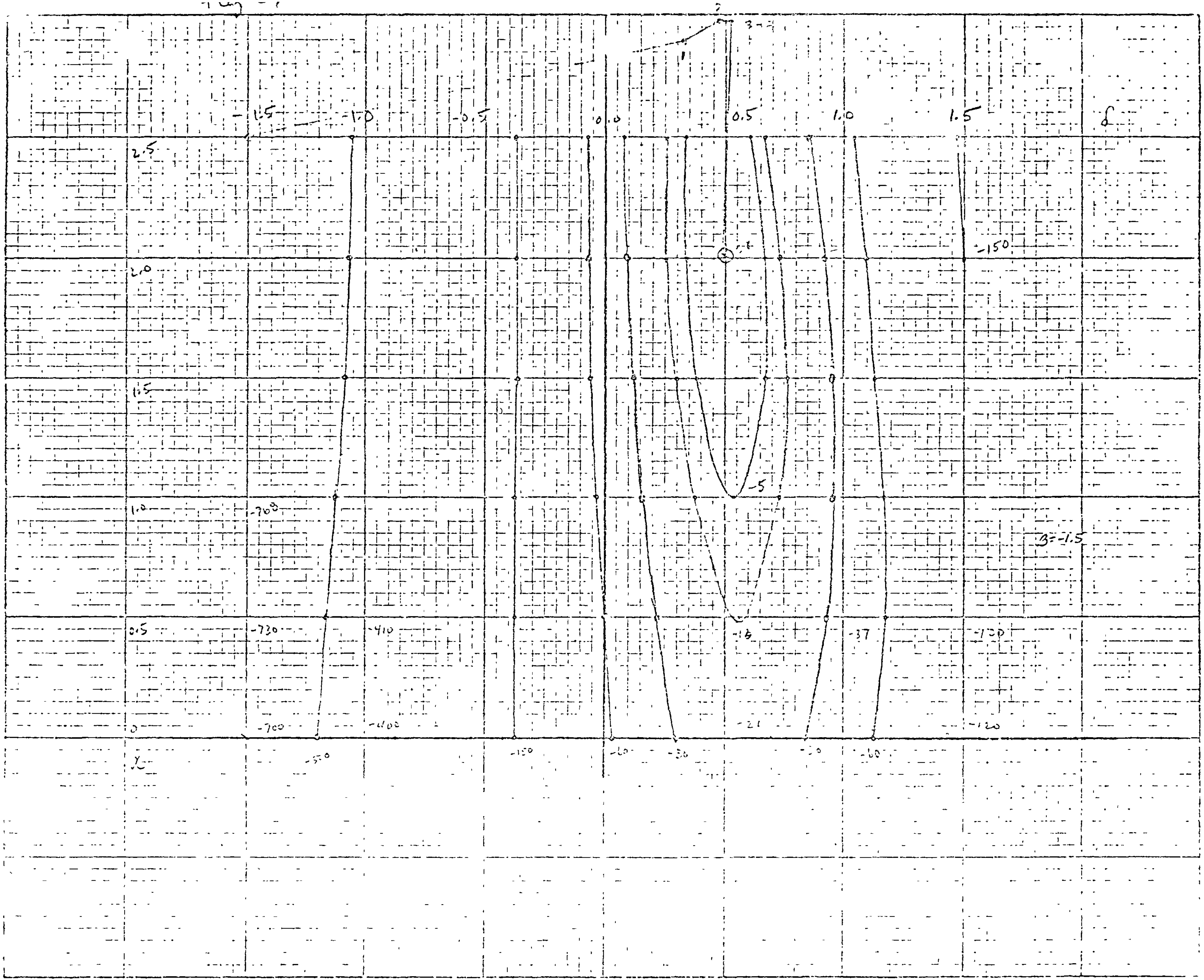
Fig 25

NINE ZONE PROBLEM

$\alpha = 2.0$




-1.5
-500
-1.0
-0.5
0.0
-60
-60
0.5
-150
1.0
1.5
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
-551
-731
-826
-814
-777
-150
-70
-10
-10
-3



2
3
1
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
δ
2.5
2.0
1.5
1.0
0.5
0
γ
−150
−5
β=−1.5
−700
−730
−410
−18
−37
−120
−700
−400
−21
−120
−550
−150
−60
−30
−20
−60

Again, referring to equation (11) we get

$$\frac{\partial L}{\partial \delta} = \frac{\partial L}{\partial \hat{N}_i}\left(\frac{\partial \hat{N}_i}{\partial \delta}\right) \tag{62}$$

and, substituting from equations (39), (61), and (62)

$$\frac{\partial L}{\partial \delta} = \sum_i \left(\frac{N_i}{\hat{N}_i}\right)\left[\sum_i T_{ij}\left[\ln W_i - \sum_k \left(C_{ik}^{\alpha}\left\{\exp(\beta C_{ik})\right\} \ln W_i\right) B_i^{-1}\right]\right] \tag{63}$$

The partial derivative equations of ln L with respect to α and to β remain the same as for the two parameter case. Further, it should be clear that if the model were extended to include a multivariate attractiveness measure with more than one parameter, further partial derivative calculations would be necessary.

It is a rather straightforward matter to do a gradient search on the nine zone test problem data. The known optimum is at α = 2.00, β = 1.50, and δ = 0.50. Several searches were run each from a different starting point. Searches initiated at different starting points took different numbers of steps to reach the optimum. One sample search is shown below for a starting point of α = 2.5, β = -3.0, and δ = -1.5.

Iteration	Criterion	α	β	δ
1	-671.4975	2.5000	-3.0000	-1.5000
2	-10.4296	2.8776	-2.0269	0.3264
3	-0.4113	2.9639	-1.8154	0.4836
4	-0.1367	2.9351	-1.8688	0.5242
5	-0.1330	2.9220	-1.8604	0.5248
6	-0.1308	2.9116	-1.8595	0.5225
7	-0.1289	2.9094	-1.8583	0.5236
8	-0.1269	2.8975	-1.8542	0.5213
9	-0.1042	2.8159	-1.8210	0.5217
10	-0.0009	2.0125	-1.5025	0.4993

The parameters could have been pushed closer to their known optimal values by further iterations of the search procedure, but the above ten iterations suffice for illustration.

CONCLUSIONS

Various conclusions have been stated throughout the paper, and have in some places guided the direction of the research. Taken all-in-all, we have the following:

1. In the absence of spatial interaction data it is difficult (if not impossible) to calibrate singly constrained spatial interaction models by the procedures found in the current literature.

2. A procedure is suggested for use in such cases and is then examined in various ways. The procedure involves selecting a criterion function and then using the equations for the partial derivative(s) of the criterion with respect to the parameter(s) to guide a search for the best fit parameter values.

3. The investigation was begun using R^2 as the criterion and a three zone, one parameter sample problem. It was found that differences in the data set produced substantial differences in the shape of the criterion and gradient functions.

4. The investigation shifted to a nine zone, one parameter, sample problem. It was found that the differences in shapes of criterion and gradient functions due to differences in data sets, were reduced. The R^2 criterion was found to be very flat in the region of its optimum thus suggesting the desirability of finding a different criterion.

5. The procedure was repeated for a twenty-five zone, one parameter problem, yielding similar results.

6. Using a likelihood criterion, the three, nine, and twenty-five zone, one parameter problems were rerun. At the nine and twenty-five zone levels the likelihood criterion showed a much sharper peak near the optimum than did the R^2 criterion. The likelihood criterion was, therefore, adopted for futher investigation.

7. As both the nine and twenty-five zone problems had yielded similar results, for economy's sake the remainder of the investigation was performed at the nine zone level.

8. A nine zone, two parameter, sample problem was developed. A strong correlation between the two parameters was found. The surface was drawn in several ways and found to be unimodal.

9. A nine zone, three parameter sample problem was developed. Correlations were found to exist between all three parameters and the surface was drawn in several ways. It appeared again that the likelihood surface is unimodal, though search over this surface will be complicated by the existence of ridges between parameters.

10. A sample trace of a gradient search over the nine-zone, three parameter likelihood surface was presented. It was found that the number of search iterations could vary considerably depending upon the choice of starting points. Further, given the interrelatedness of the parameters it became quite obvious that it would always be necessary to estimate their values simultaneously rather than individually (as was often the case in past land use models practice).

This concludes Part 1 of this 2 part paper on calibration of urban residential models. In Part 2 we present the empirical results of applying these methods to data sets from more than twenty North American cities.

References

Batty, M (1976) Urban Modelling Cambridge University Press, Cambridge.

Batty, M. and S. Mackie (1972) "The Calibration of Gravity, Entropy, and Related Models of Spatial Interaction", Environment and Planning, Vol. 4 pp. 206-233.

Baxter, R. and I. Williams (1975) "An Automatically Calibrated Urban Model", Environment and Planning A, Vol. 7 pp. 3-20.

Openshaw (1976) "An Empirical Study of Some Spatial Interaction Models", Environment and Planning A, Vol. 8, pp. 23-42.

Putman, S. (1977) "Calibrating a Disaggregated Residential Allocation Model - DRAM", London Papers in Regional Science 7 Pion, London pp. 108-124.

Williams (1976). "A Comparison of Some Calibration Techniques for Doubly Constrained Models with an Exponential Cost Function" Transportation Research, Vol. 10, pp. 91-104.

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

MODELOS DE SIMULACION PARA LA PLANEACION INTREGRAL DEL USO DEL SUELO Y EL TRANSPORTE

TEMA: IMGRID.

SISTEMA PARA EL MANEJO DE SISTEMAS RETICULARES DE INFORMACION EN LA PLANEACION DEL USO DEL SUELO.

AUTORES: C.STEINITZ, D.SINTON, R. THOT.

PREPARACION: A. VILLANUEVA, B.HORNEDO.
M. GARCIA.

ENERO, 1978.

Palacio de Minería Calle de Tacuba 5, primer piso. México 1, D. F.

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*.

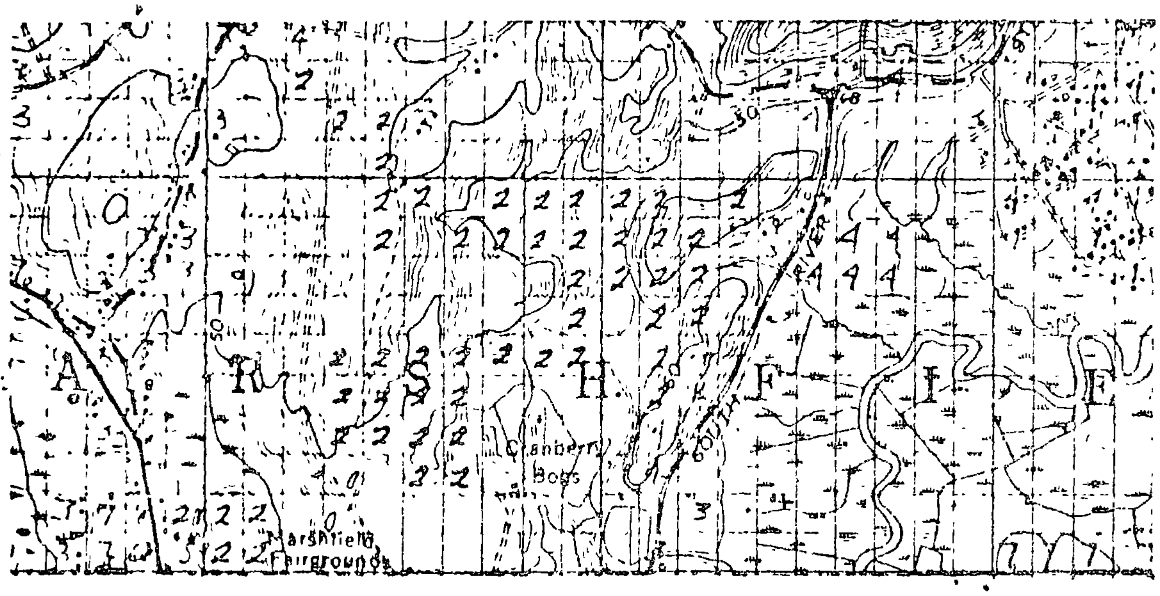
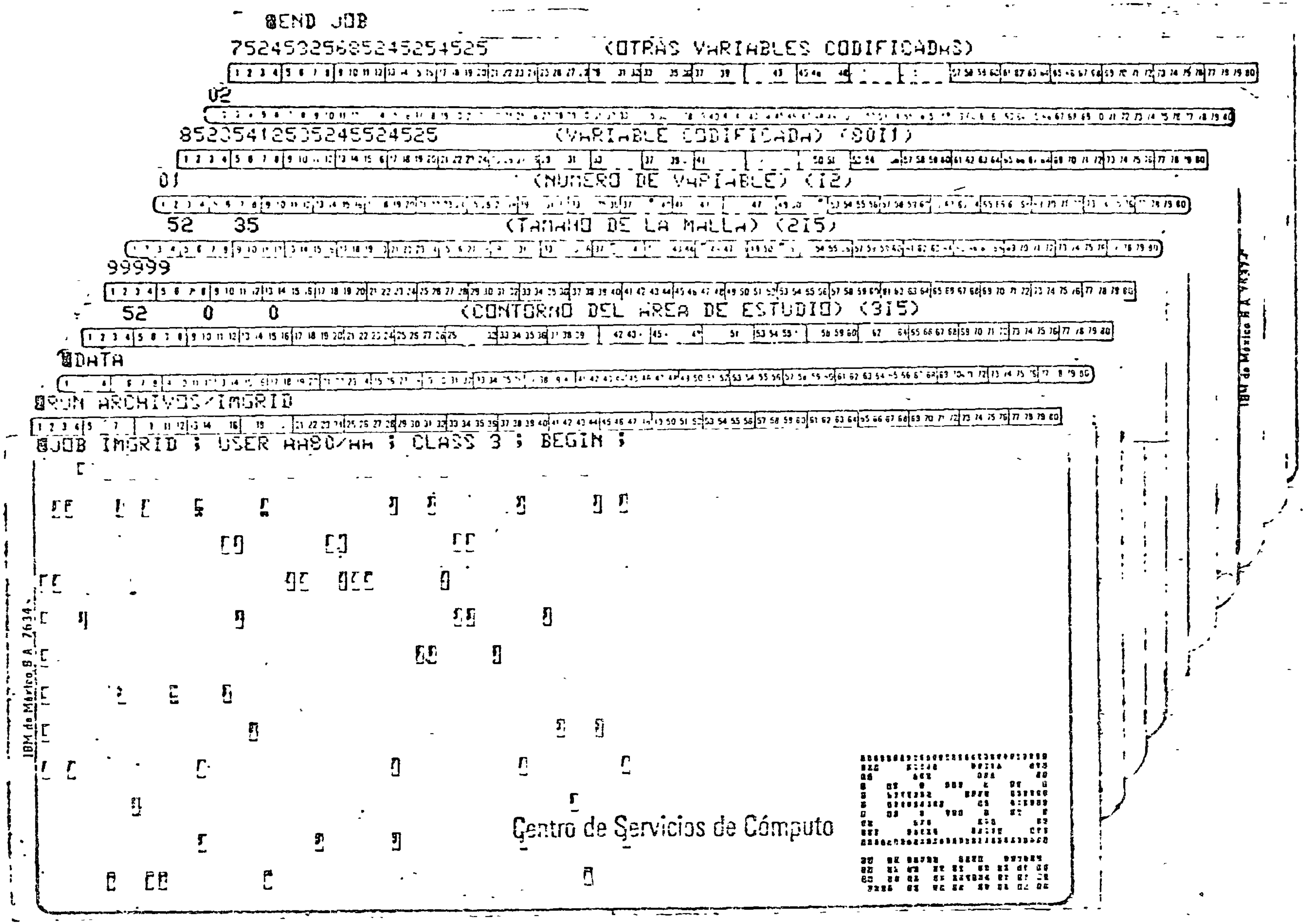


Figura 1. Ejemplo de retícula sobre mapa base.



```
?END JOB
752453256852452545254525      (OTRAS VARIABLES CODIFICADAS)
02
8520541253524552452 5         (VARIABLE CODIFICADA) (80I1)
01                             (NUMERO DE VARIABLE) (I2)
52   35                        (TAMAÑO DE LA MALLA) (2I5)
99999
52     0     0                 (CONTORNO DEL AREA DE ESTUDIO) (3I5)
?DATA
?RUN ARCHIVOS/IMGRID
?JOB IMGRID ; USER AA80/AA ; CLASS 3 ; BEGIN ;
```

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 [illegible] 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):

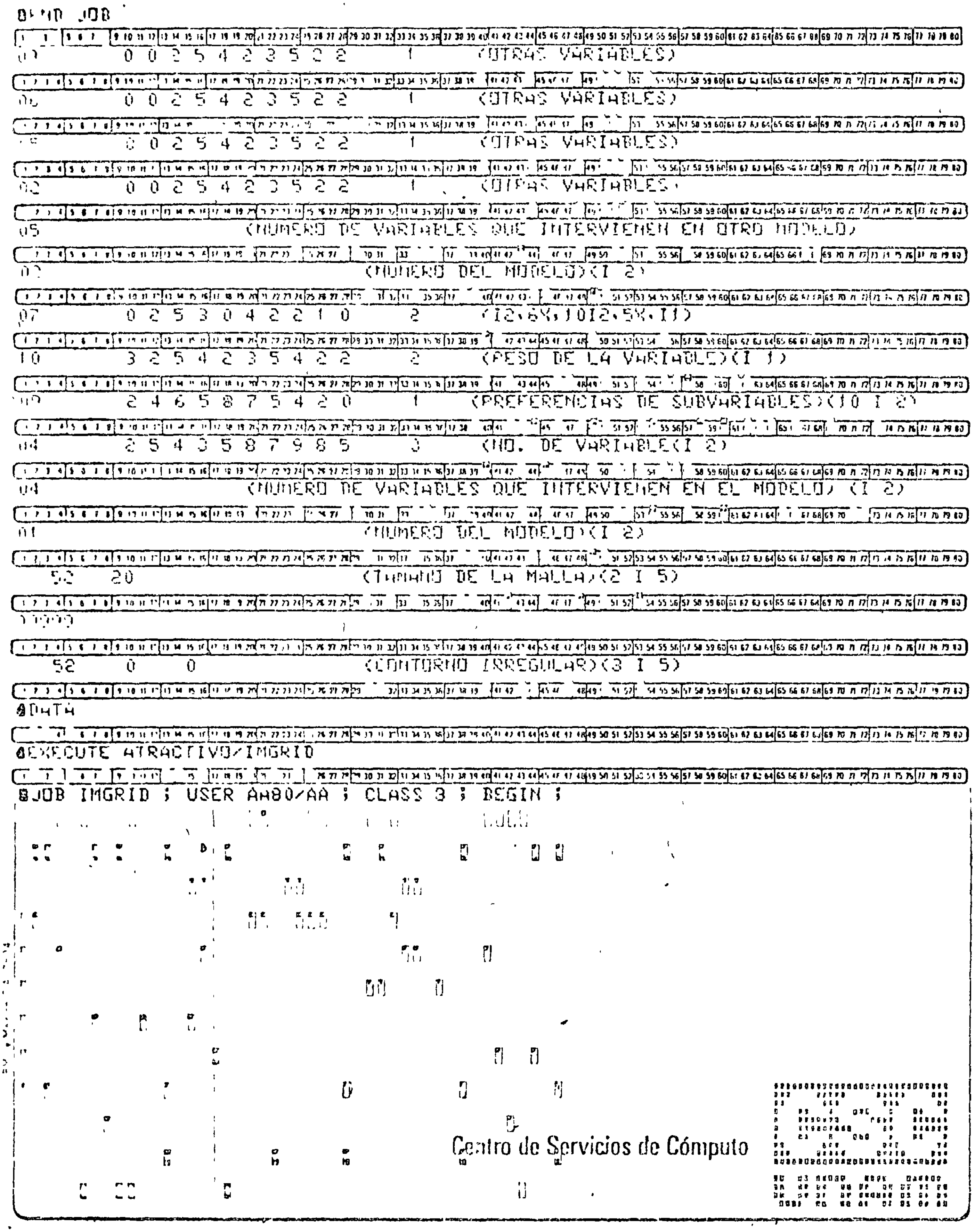
Nombre de la actividad (usos del suelo) para la que se construyó el modelo:

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)										Cual es la portancia de variable (
	Col-10	Col-12	Col-14	Col-16	Col-18	Col-20	Col-22	Col-24	Col-26	Col-28	Col-3

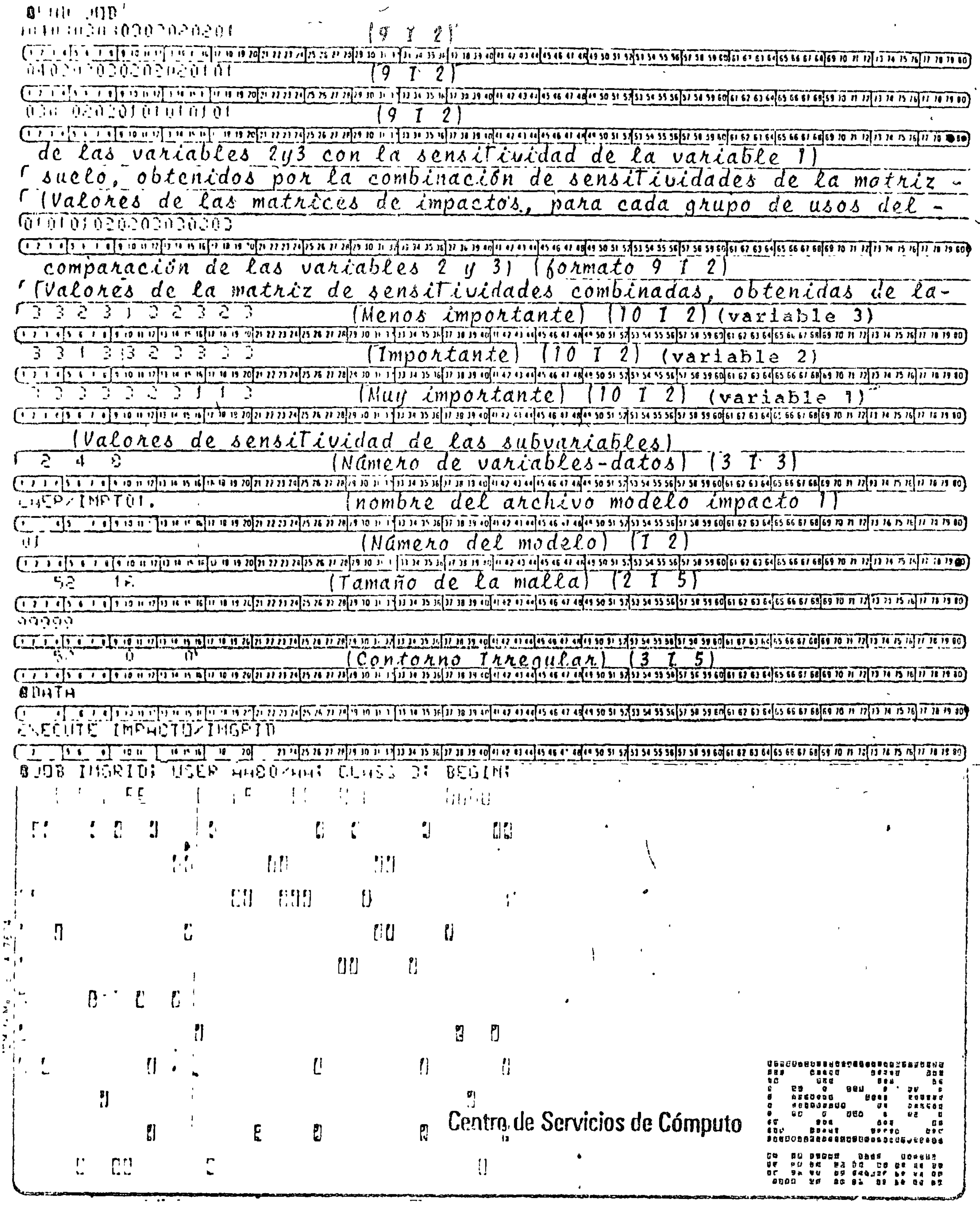
Figura 4.- Hoja de codificación para los MODELOS DE ATRACTIVIDAD.



```
?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)
08        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)
03                                 (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.



```
@END JOB
010103030303020201                    (9 I 2)
040202020202020101                    (9 I 2)
030302020101010101                    (9 I 2)
 de las variables 2y3 con la sensitividad de la variable 1)
 suelo, obtenidos por la combinación de sensitividades de la matriz -
 (Valores de las matrices de impactos, para cada grupo de usos del -
010101020202030303
 comparación de las variables 2 y 3) (formato 9 I 2)
 (Valores de la matriz de sensitividades combinadas, obtenidas de la-
3 3 2 3 1 3 2 3 2 3                   (Menos importante) (10 I 2) (variable 3)
3 3 1 3 3 2 3 3 3 3                   (Importante) (10 I 2) (variable 2)
3 3 3 3 3 2 3 1 1 3                   (Muy importante) (10 I 2) (variable 1)
 (Valores de sensitividad de las subvariables)
2 4 8                                 (Número de variables-datos) (3 I 3)
LHEP/IMPT01.                          (nombre del archivo modelo impacto 1)
01                                    (Número del modelo) (I 2)
52  16                                (Tamaño de la malla) (2 I 5)
99999
5  0  0                               (Contorno Irregular) (3 I 5)
@DATA
?EXECUTE IMPACTO/INGRID
@JOB INGRID; USER AABO/AAA; CLASS 3; BEGIN
```

Centro de Servicios de Cómputo

Figura 8.- Forma en la que debe ordenar su paquete de tarjetas para utilizar el programa IMPACTO/INGRID, en el sistema B-6700 del CSC, y obtener la evaluación de hasta 30 modelos de impacto por corrida.

```
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
```

```
@@      @@   @@@@@@@@   @@@@@@@@      @@@@@@@@     @@@@@@
@@@    @@@  @@@@@@@@@@  @@@@@@@@@    @@@@@@@@@@   @@@@@@@@
@@@@  @@@@  @@      @@  @@     @@@   @@      @@  @@@    @@@
@@ @@@@ @@  @@      @@  @@      @@   @@      @@  @@@
@@  @@  @@  @@@@@@@@@@  @@     @@@   @@@@@@@@@@    @@@
@@      @@  @@@@@@@@@@  @@@@@@@@@    @@@@@@@@@@      @@@
@@      @@  @@      @@  @@@@@@@@     @@      @@        @@
@@      @@  @@      @@  @@           @@      @@  @@@    @@@
@@      @@  @@      @@  @@           @@      @@   @@@@@@@@
@@      @@  @@      @@  @@           @@      @@    @@@@@@


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 TAMAÑO 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 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




 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 = HGFG2
                          2 = ST
                          3 = LHBC
                          4 = CHE
                          5 = KHGH
                          6 = GZF
                          7 = GRGH
                          8 = GZL
                          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 LA EA
                          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

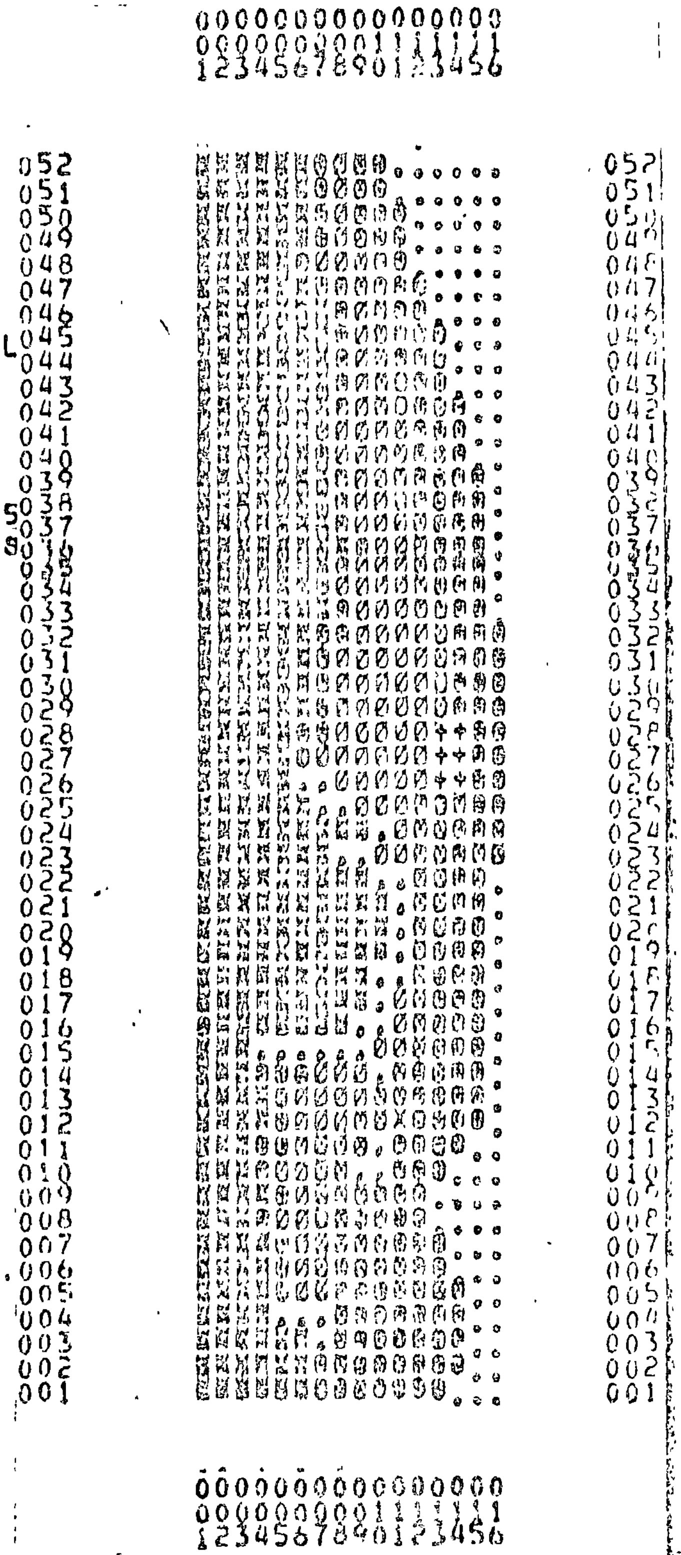
EA DE ESTUDIO VALLE PACIFICO LABORATORIO DE PLANEACION URBANA

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	3	4	5	6	7	8	9
FRECUENCIA	98	3	28	6	1	2	98	76	203	317



MAPA DE LA VARIABLE 4 DEL INVENTARIO DE DATOS VEGETACION POR ZONA

AREA DE ESTUDIO VALLE DEL PACIFICO LABORATORIO DE PLANEACION URBANA

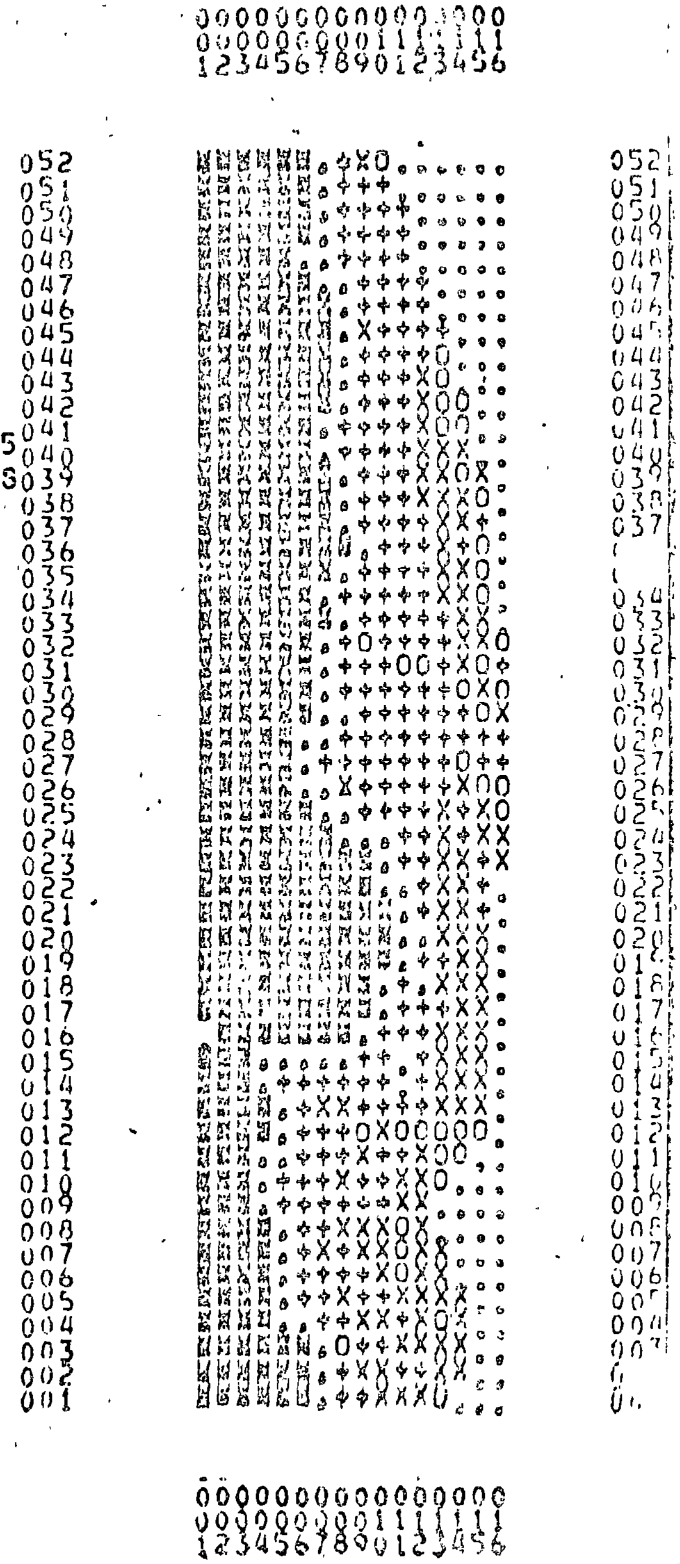
SUBVARIABLES DATO

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

SIMBOLOS

NIVELES	0	1	2	3	4	5	6	7	8	9
FRECUENCIA	101	0	69	200	106	39	0	0	0	517



MAPA DE LA VARIABLE 2 DEL INVENTARIO DE DATOS PORCIENTO DE PENDIENTE

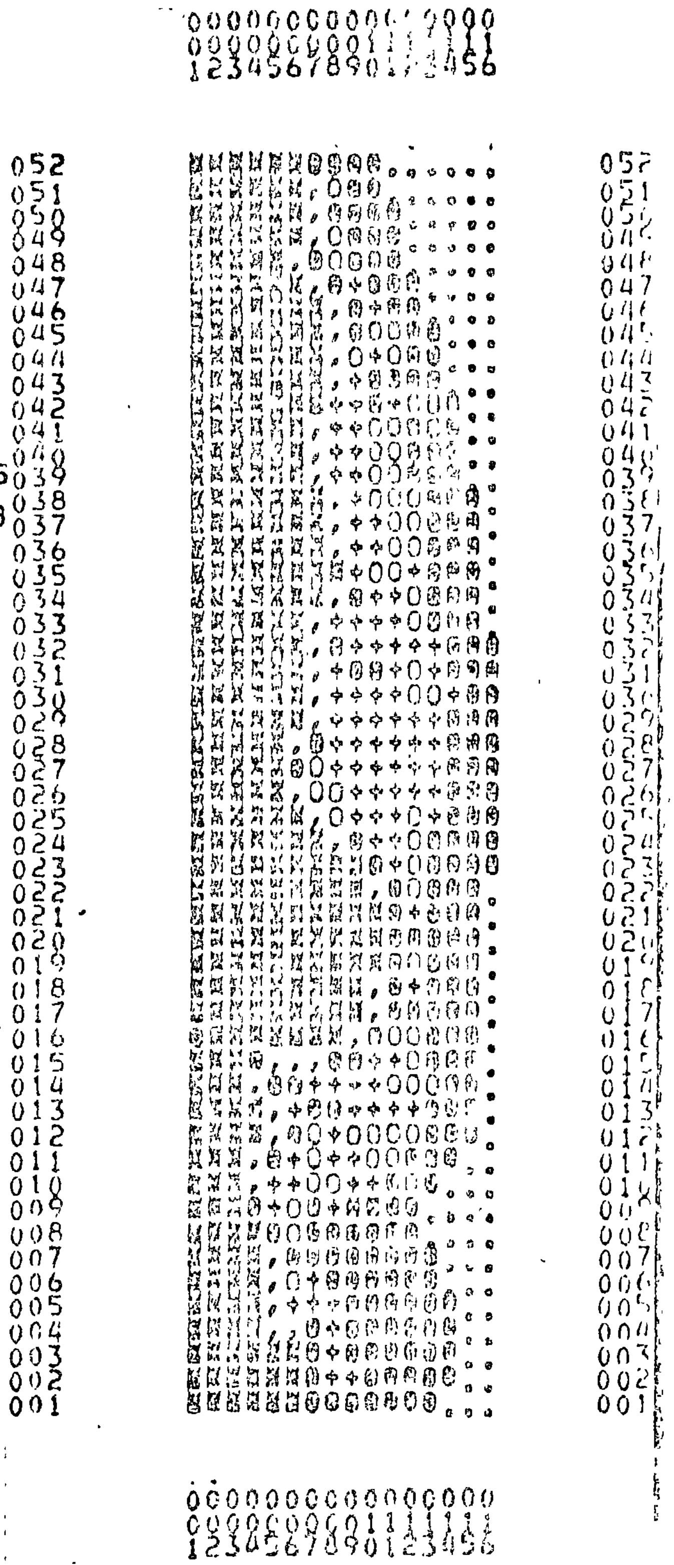
AREA DE ESTUDIO VALLE DEL PACIFICO LABORATORIO DE PLANEACION URBANA

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
          ==================================================
             .........          ,,,,,,,,,          ooooooooo
             .........          ,,,,,,,,,          ooooooooo
SIMBOLOS     ....0....          ,,,,1,,,,          oooo2oooo
             .........          ,,,,,,,,,          ooooooooo
             .........          ,,,,,,,,,          ooooooooo
          ==================================================
FRECUENCIA       99                  42                   0

          3              4              5              6
==========================================================
+++++++++      XXXXXXXXX      OOOOOOOOO      ØØØØØØØØØ
+++++++++      XXXXXXXXX      OOOOOOOOO      ØØØØØØØØØ
++++3++++      XXXX4XXXX      OOOO5OOOO      ØØØØ6ØØØØ
+++++++++      XXXXXXXXX      OOOOOOOOO      ØØØØØØØØØ
+++++++++      XXXXXXXXX      OOOOOOOOO      ØØØØØØØØØ
==========================================================
   100              0             54              0

          7              8              9
============================================
[illegible]    [illegible]    [illegible]
============================================
    7             214            316
```



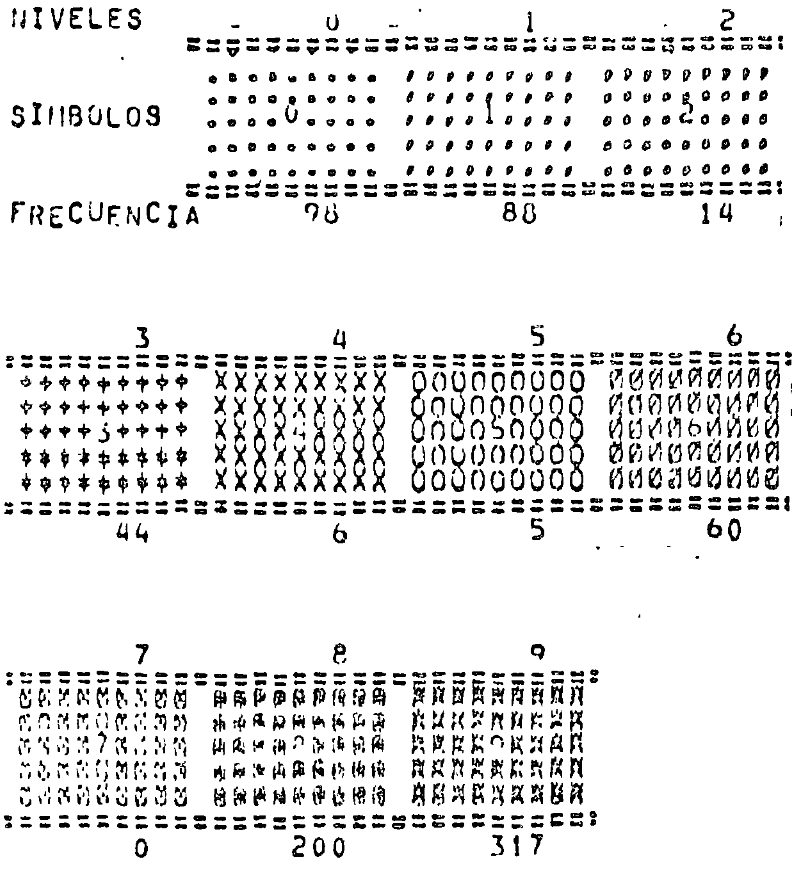
MAPA DE LA VARIABLE 3 DEL INVENTARIO DE DATOS ORIENTACION PENDIENTE

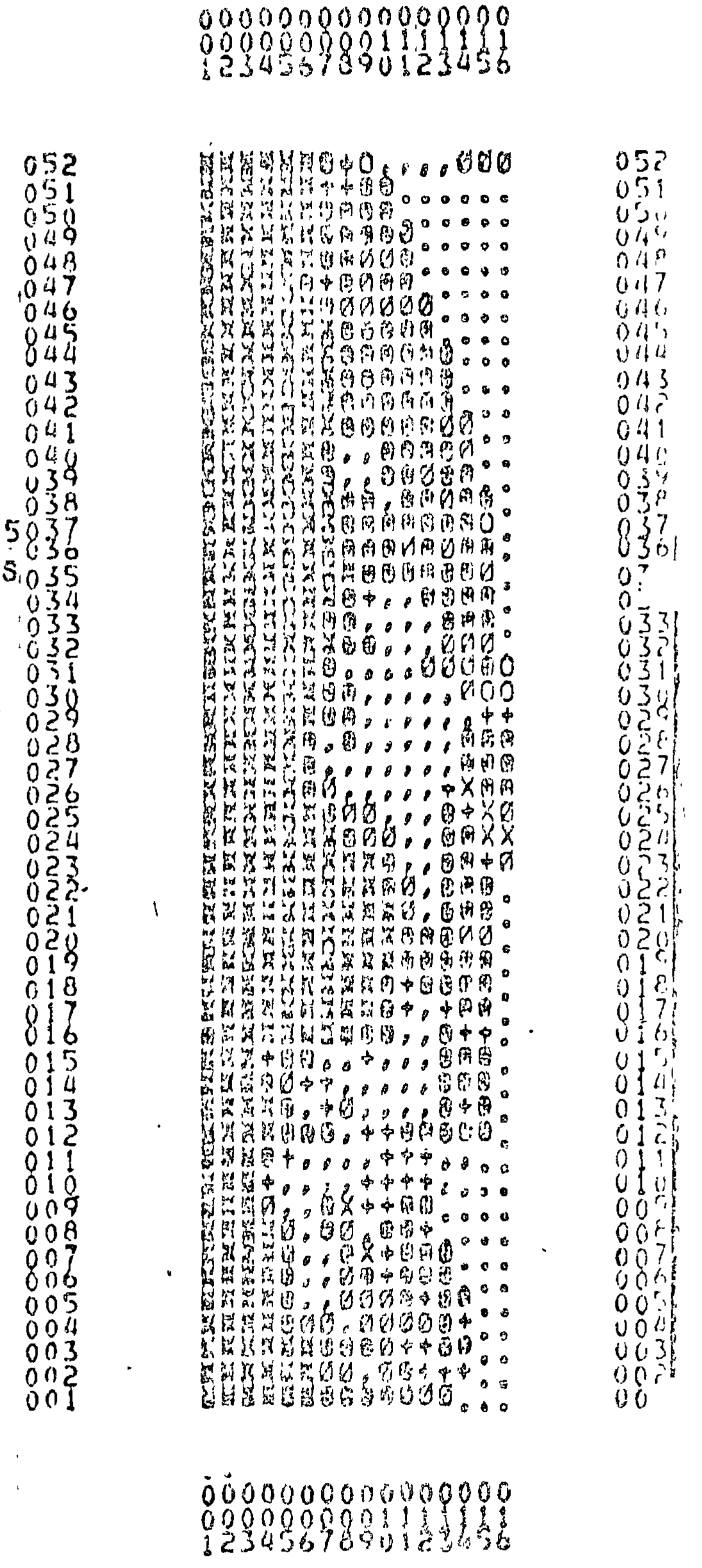
AREA DE ESTUDIO VALLE DEL PACIFICO LABORATORIO DE PLANEACION URBANA

SUBVARIABLES DATO

0 = NO HAY DATOS
1 = LLANURA
2 = NORTE
3 = 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





MAPA DE LA VARIABLE 5 DEL INVENTARIO DE DATOS DENSIDAD DE ARBOLES

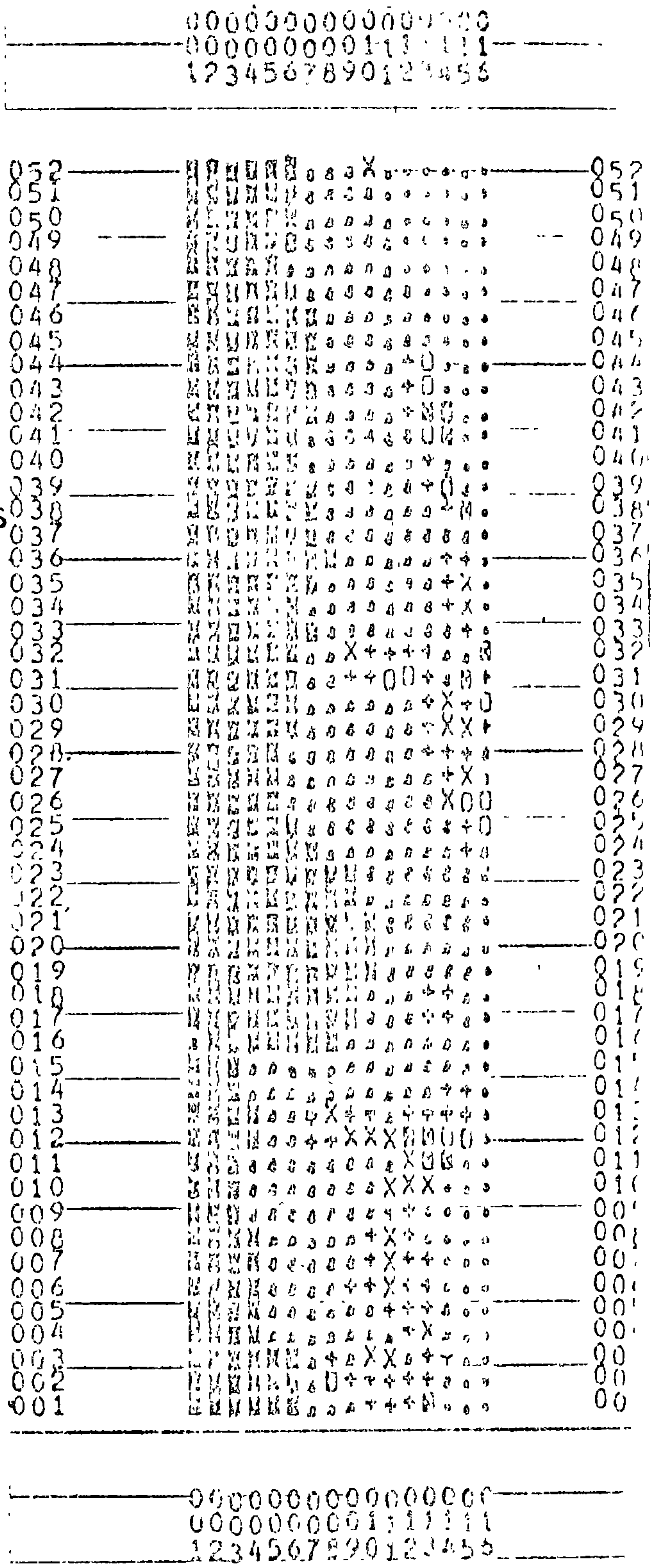
AREA DE ESTUDIO VALLE DEL PACIFICO LABORATORIO DE PLANEACION URBANA

SUBVARIABLES DATO

0 = NO HAY DATOS
2 = NADA
3 = 1-25%
4 = 26-50%
5 = 51-75%
7 = 76-100%
9 = OCEANO

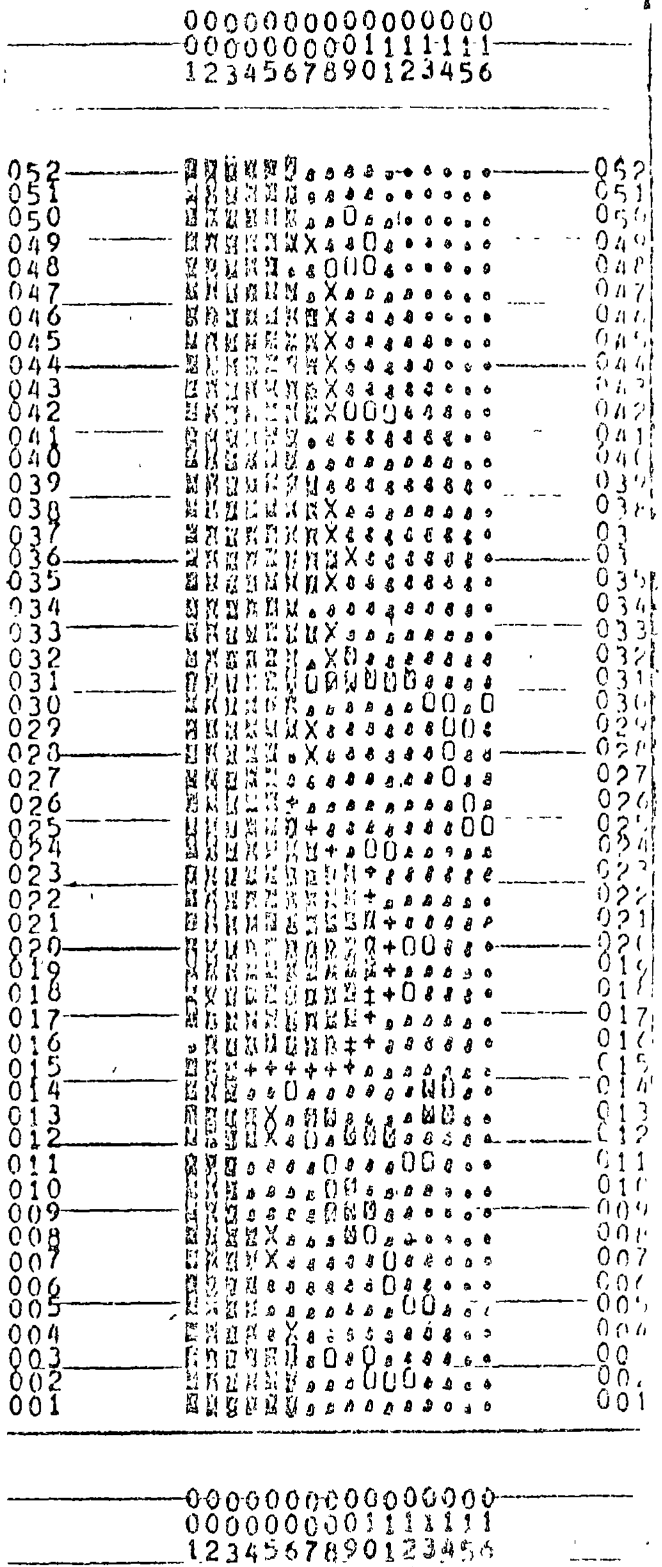
EL TAMANO DE LA CELULA DE LA MALLA ES DE 2. ACRES

```
NIVELES        0            1            2
SIMBOLOS       0            1            2
FRECUENCIA     98           0            300

     3          4          5          6
     3          4          5          6
     70         23         11         0

     7          8          9
     7          8          9
     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
SIMBOLOS	0	1	2
FRECUENCIA	97	1	318

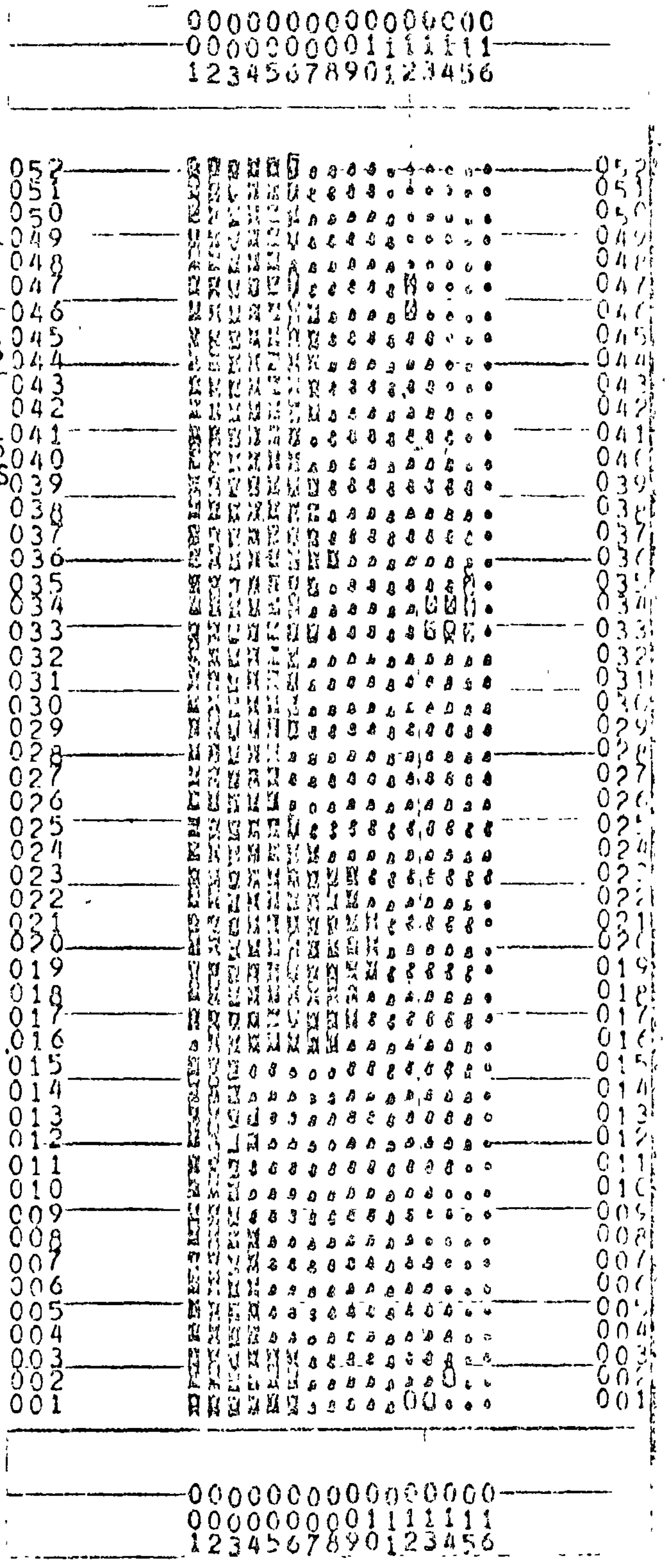
	3	4	5	6
	++++3++++	XXXX4XXXX	00005000	ØØØØ6ØØØØ
	19	20	38	0

	7	8	9
	7	8	9
	22	0	317

MAPA DE LA VARIABLE 7 DEL INVENTARIO DE DATOS ACCIDENTES GEOLOGICOS

AREA DE ESTUDIO VALLE DEL PACIFICO LABORATORIO DE PLANEACION URBANA

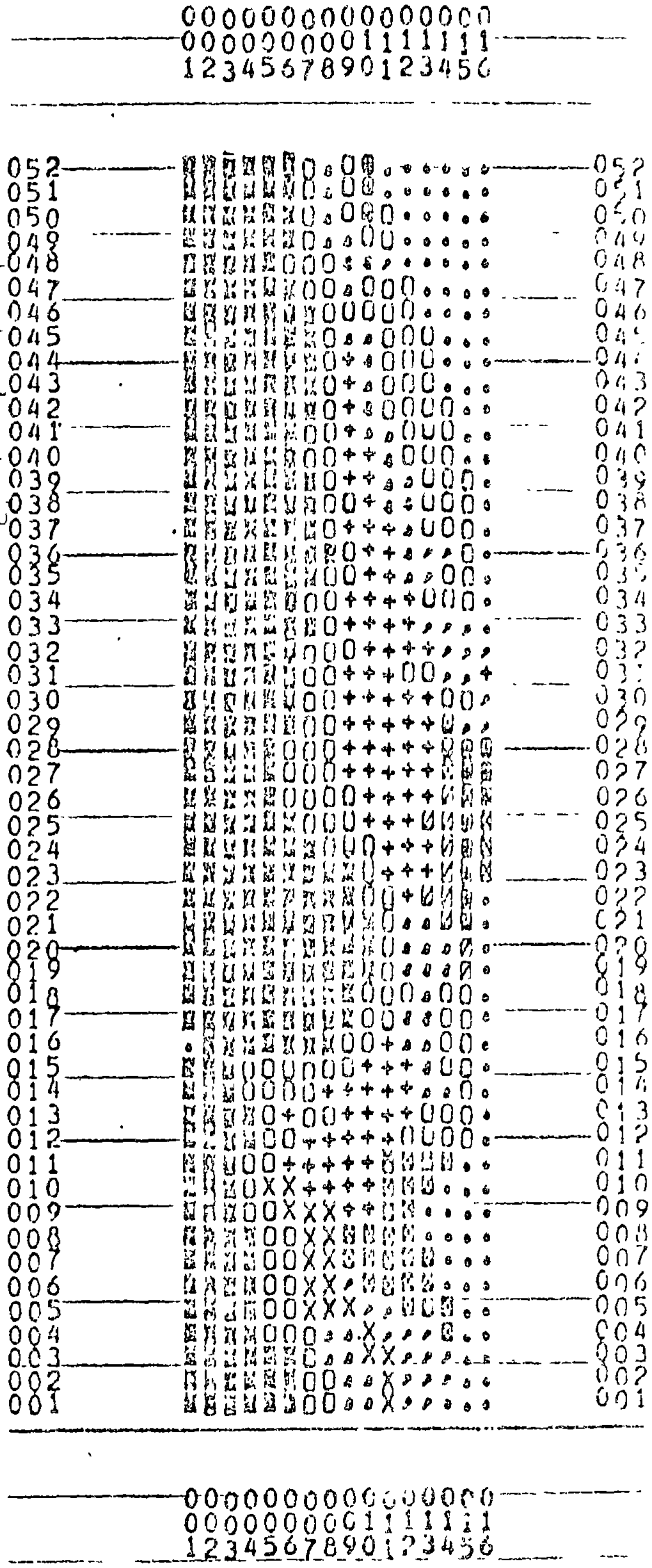
```
SUBVARIABLES DATO

      0 = NO HAY DATOS
      2 = AREA ESTABLE
      5 = DESLIZAMIENTOS ACTIVOS NO NATURALES
      7 = DESLIZAMIENTOS ACTIVOS DE TIERRA
      9 = OCEANO

EL TAMANO DE LA CELULA DE LA MALLA ES DE 2.5
                                        ACRES

NIVELES            0             1             2
          ===================================================
SIMBOLOS  ....0....   ////1////   8888288888
FRECUENCIA    98            0            405

          3             4             5             6
====================================================
++++3++++   XXXXXXXXX   000050000   0000600000
     0             0             3             0

          7             8             9
=========================================
[illegible]7[illegible]   [illegible]   [illegible]9[illegible]
     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 = MGFG2
2 = ST
3 = LMBC
4 = CHE
5 = RHGH
6 = GZF
7 = GPGH
8 = GZF
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	24	19	169	10	32	14	317

MAPA DE LA VARIABLE 9 DEL INVENTARIO DE DATOS PROXIMIDAD AL AGUA

AREA DE ESTUDIO VALLE DEL PACIFICO

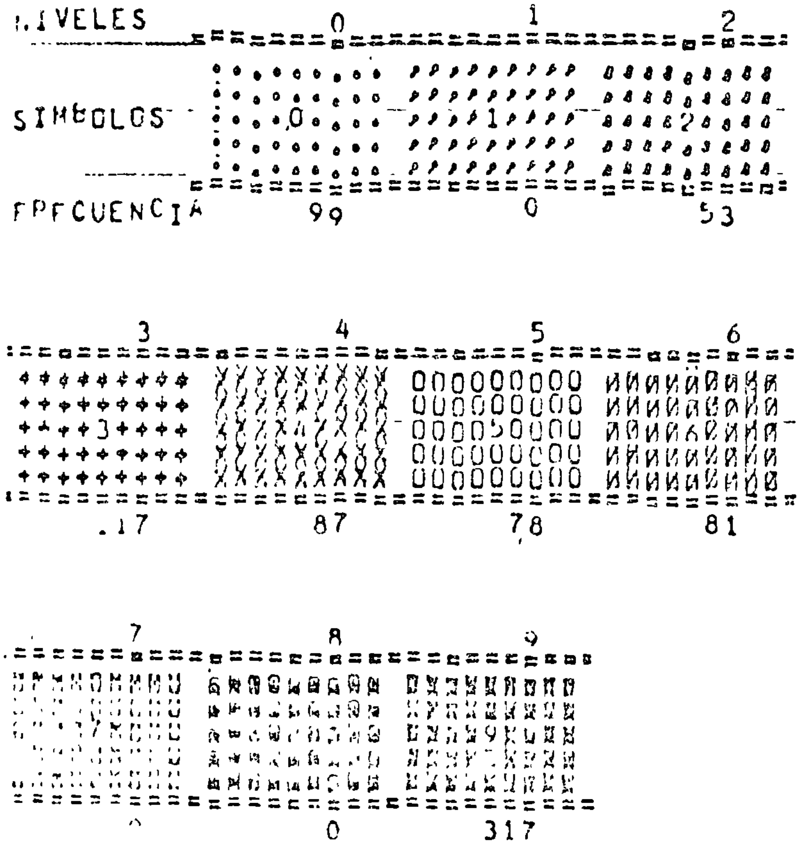
LABORATORIO DE PLANEACION URBANA

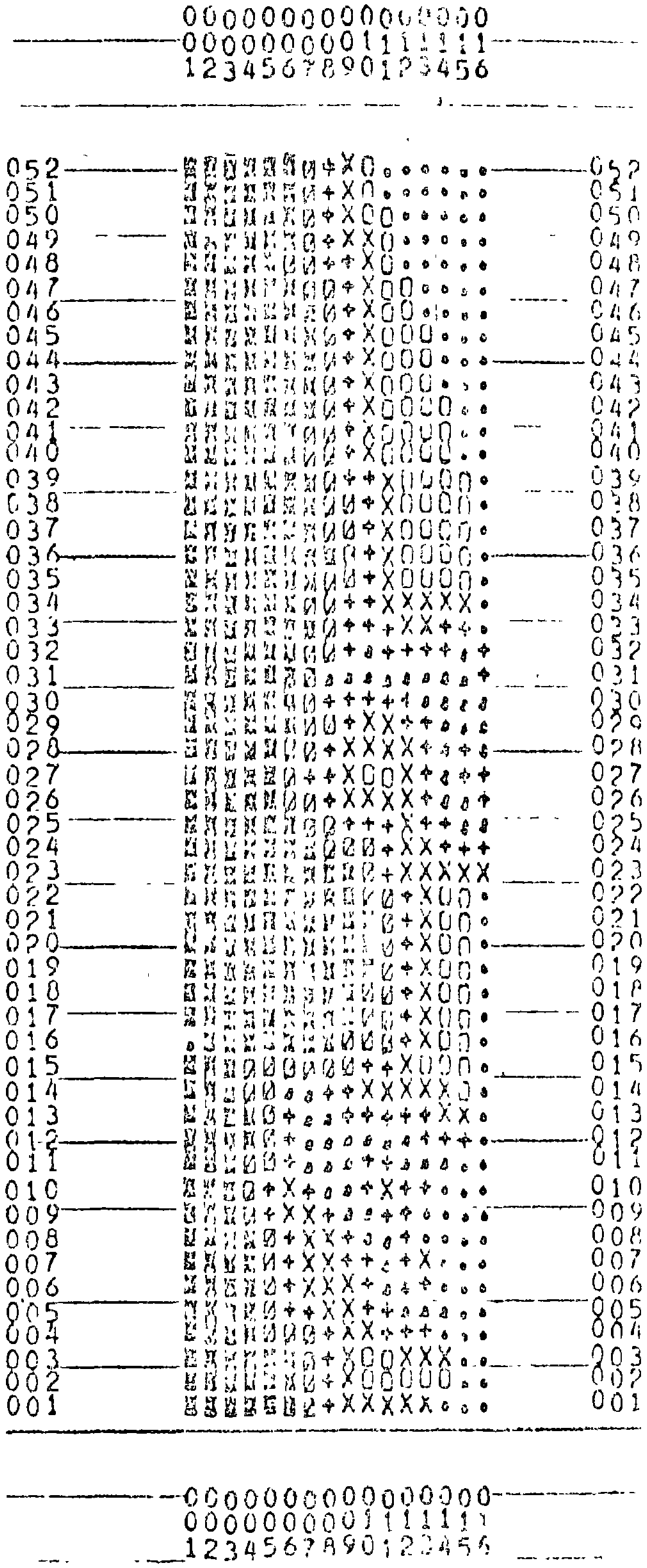
SUBVARIABLES DATO

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

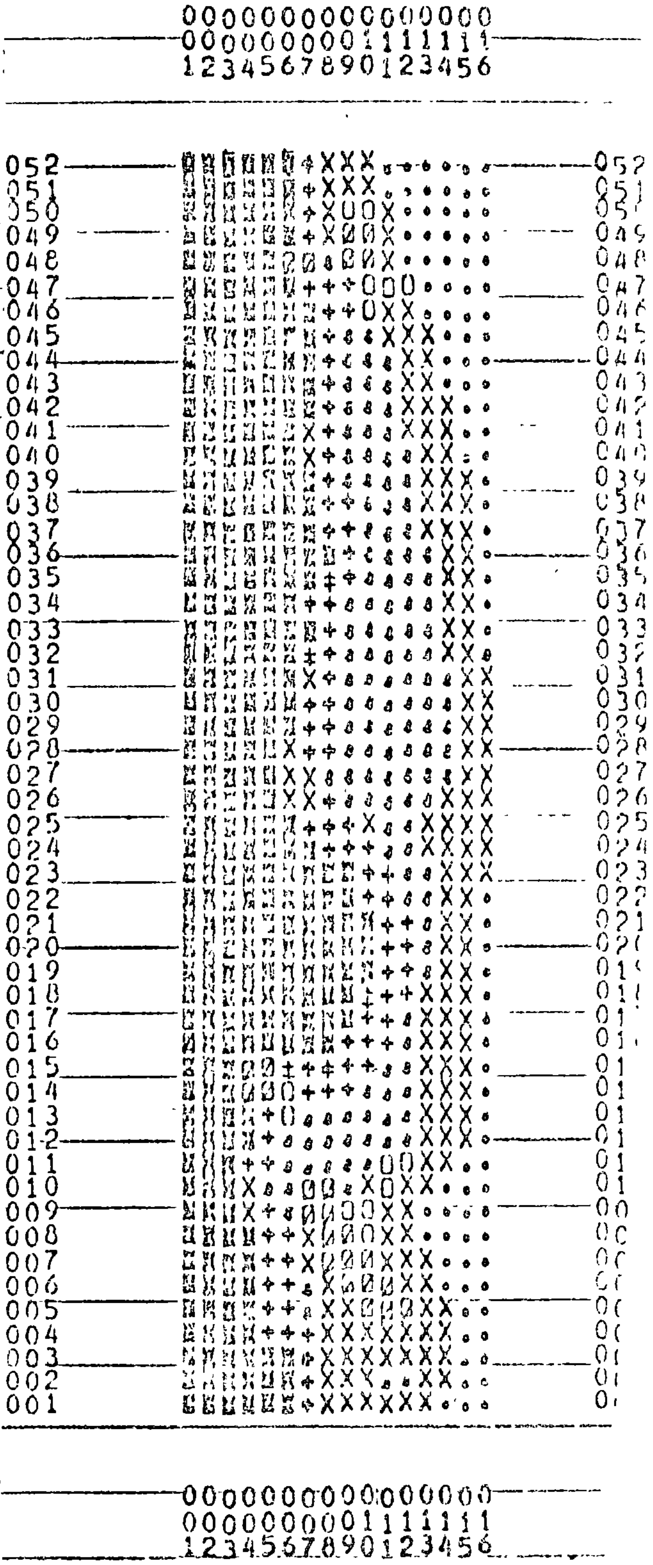




MAPA DE LA VARIABLE 10 DEL INVENTARIO DE DATOS   VISTAS

AREA DE ESTUDIO   VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA



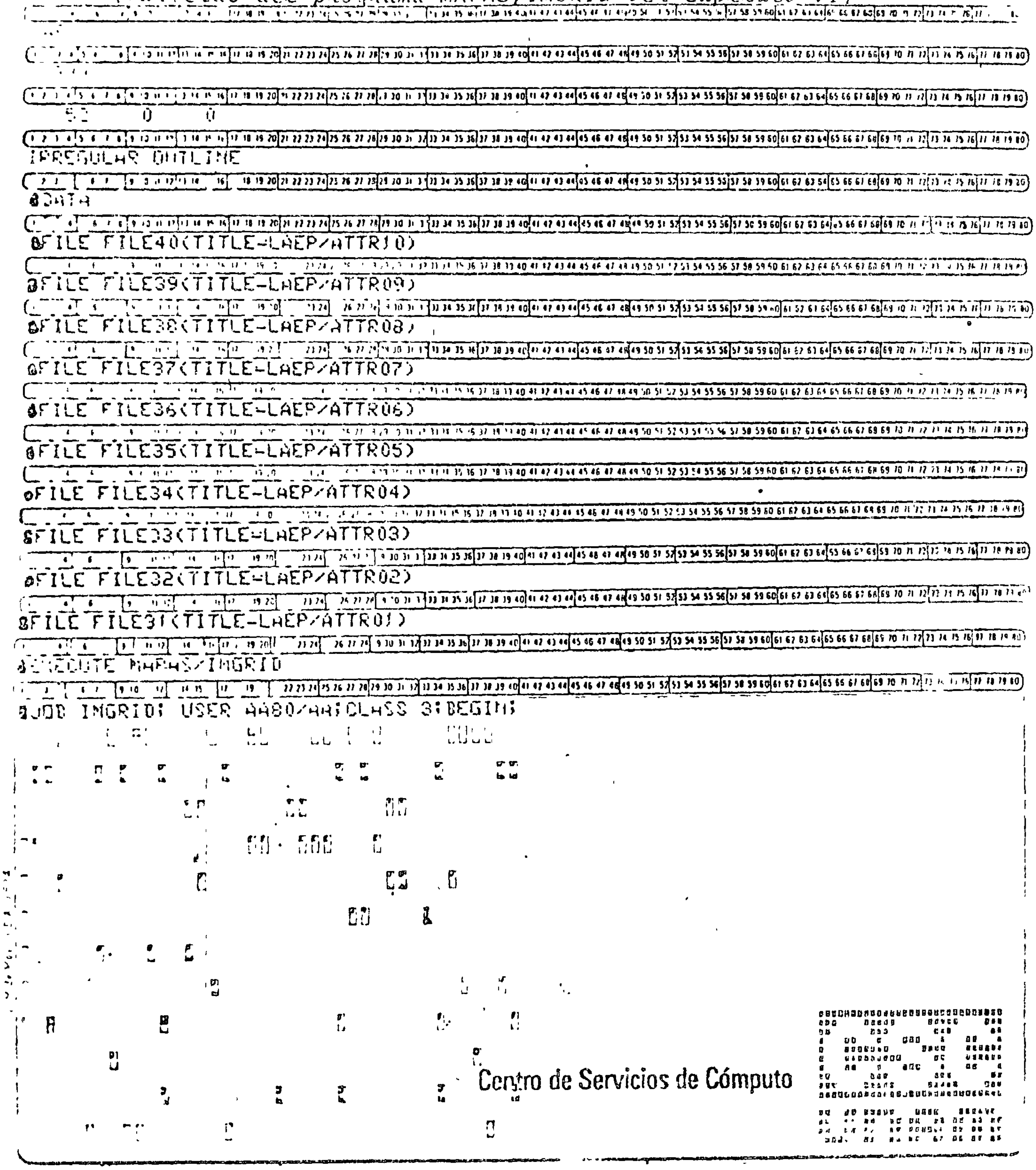
SUBVARIABLES DATO

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	•	ʼ	ˏ	+	X	O	▨	▨	▨	▨
FRECUENCIA	98	0	133	87	157	14	26	0	0	317

(Tarjetas del programa MAPAS/IMGRID ver capítulo VI)



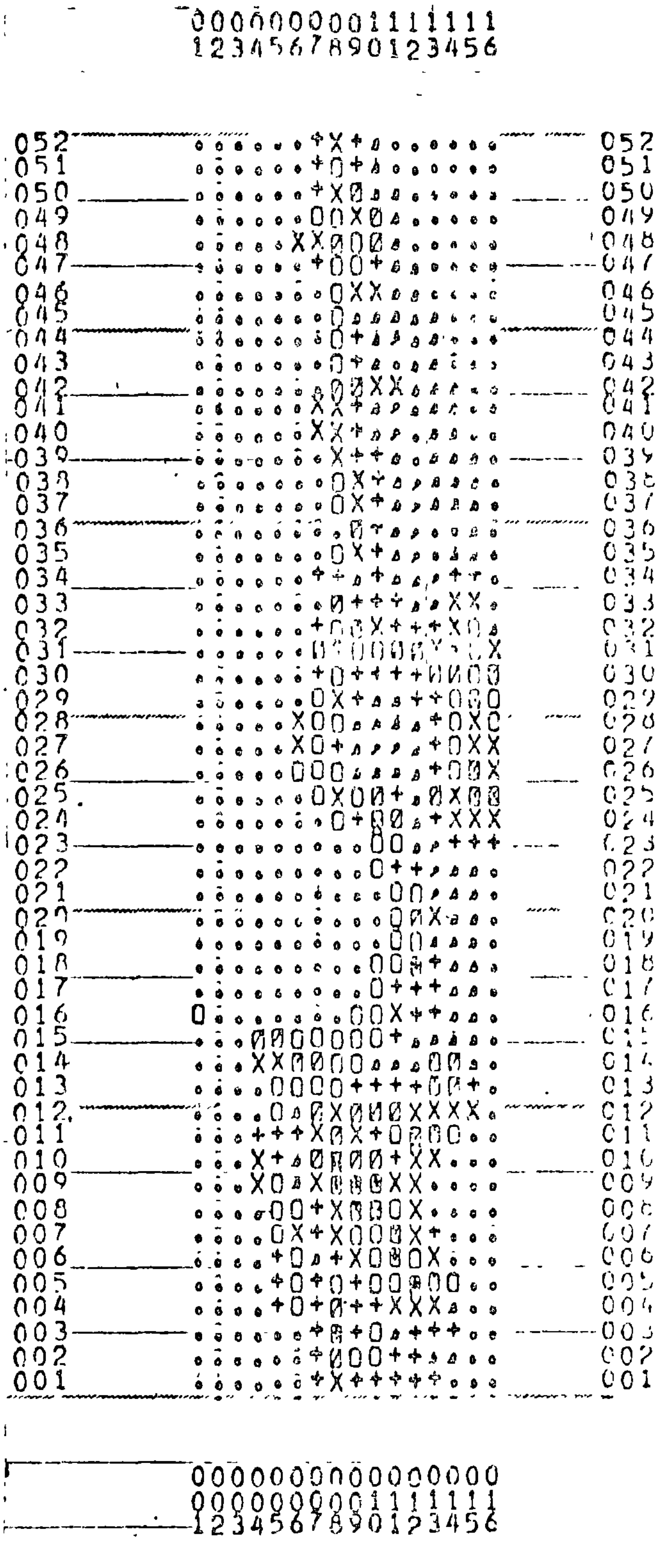
```
52    0    0
IRREGULAR OUTLINE
@DATA
@FILE FILE40(TITLE=LAEP/ATTR10)
@FILE FILE39(TITLE=LAEP/ATTR09)
@FILE FILE38(TITLE=LAEP/ATTR08)
@FILE FILE37(TITLE=LAEP/ATTR07)
@FILE FILE36(TITLE=LAEP/ATTR06)
@FILE FILE35(TITLE=LAEP/ATTR05)
@FILE FILE34(TITLE=LAEP/ATTR04)
@FILE FILE33(TITLE=LAEP/ATTR03)
@FILE 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 [illegible]FICACION 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
  ------------------------------

        EL TAMAÑO DE LA MALLA ES  52 RENGLONES Y  16 COLUMNAS
          TAMAÑO DE LA CELULA ES   1 CARACTERES EN SENTIDO VERTICAL
            CARACTERES EN SENTIDO HORIZONTAL



   7     LOS SIMBOLOS SON
          .,,+XOCOOCC123456/89
          .+XOOOOA
            /X*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

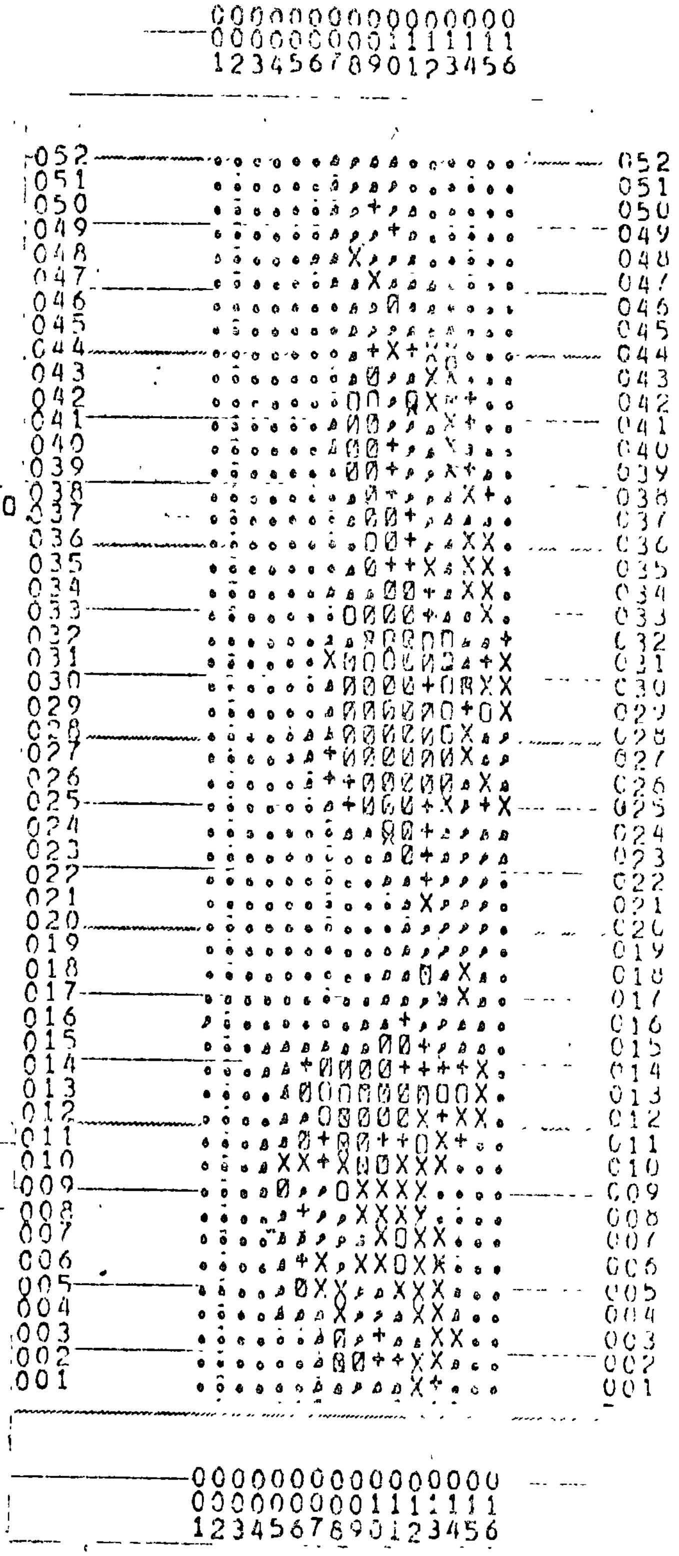
NO	NOMBRE	PESO
10	VISTAS	2
6	ZONAS LLANAS	2
9	VISTAS	2
2	PORCIENTO DE PENDIENTE	1

NIVELES	0	1	2	3	4	5	6	7	8	9
SIMBOLOS	0	1	2	3	4	5				
FRECUENCIA	420	13	101	94	64	82	33	14	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        0 0 0 9 0 2 0 0 0 0     2
       5        0 0 0 7 9 6 0 5 0 0     1
       6        0 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
```

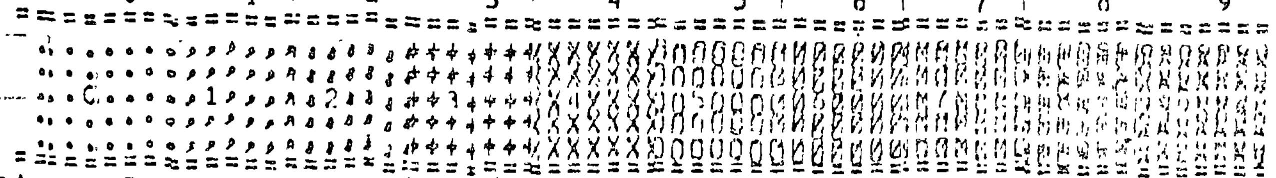
```
MODELO DE ATRACTIVIDAD # 2

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
 2      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        7       16       66       20        2        1
```

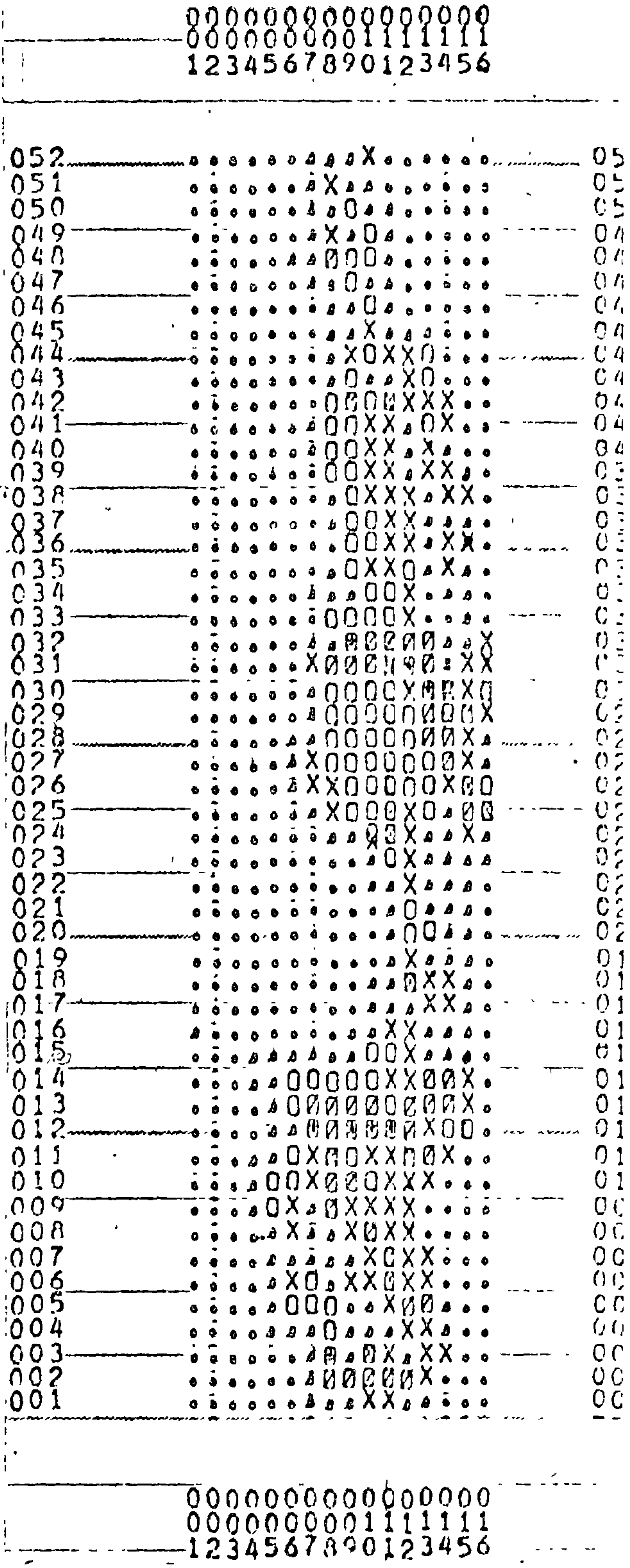
```
        ***MODELO DE ATRACTIVIDAD NUMERO   3***
[illegible]
ESTAS SON LAS TARJETAS DE LA HOJA DE CODIFICACION USADA EN EL MODELO
        2         0 0 0 9 0 7 0 0 0 0      1
        5         0 0 0 6 8 9 0 7 0 0      1
        6         0 0 0 0 0 9 0 7 0 0      1
        7         0 0 9 0 0 0 0 0 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

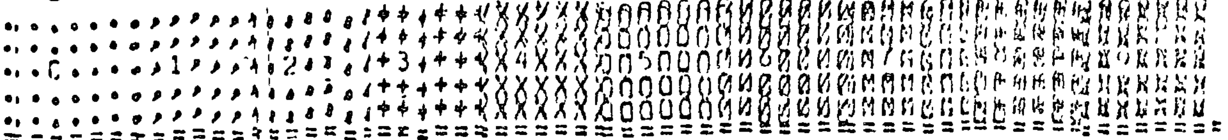


VARIABLES DATO USADAS EN EL MODELO

NO	NOMBRE	PESO
2	PORCIENTO DE PENDIENTE	1
5	PORCIENTO DENSIDAD DE ARBOLES	1
3	ZONAS LLANAS	1
7	ACCIDENTES GEOLOGICOS	1

NIVELES	0	1	2	3	4	5	6	7	8	9
FRECUENCIA	421	0	150	0	103	88	37	15	8	2

SIMBOLOS



```
        *** 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
```

```
                                                0000000000000000
MODELO DE ATRACTIVIDAD # 5                      0000000001111111
                                                1234567890123456

                                          052   . . . . . . , , , X . . . . . .   052
                                          051   . . . . . . , , , , . . . . . .   051
ESTACIONAMIENTO PARA TRAILERS             050   . . . . . . , , + , , . . . . .   050
                                          049   . . . . . . , , , X , . . . . .   049
                                          048   . . . . . , , X + X , . . . . .   048
                                          047   . . . . . . , , + , , , . . . .   047
                                          046   . . . . . . . , , + , , . . . .   046
                                          045   . . . . . . . , , , , , , . . .   045
                                          044   . . . . . . . , , + , + X . . .   044
                                          043   . . . . . . . , X , . + X . . .   043
                                          042   . . . . . . . + 0 + X + X X . .   042
                                          041   . . . . . . , + X , , , X X . .   041
                                          040   . . . . . . , + X , , , + , . .   040
AREA DE ESTUDIO____VALLE DEL PACIFICO     039   . . . . . . . + X , , , + X , .   039
                                          038   . . . . . . . , + , , + , + X .   038
                                          037   . . . . . . . , + X , , , , , .   037
                                          036   . . . . . . . . + X , , , + X .   036
                                          035   . . . . . . . , + , , , , + X .   035
                                          034   . . . . . . , , , X + , , X 0 .   034
                                          033   . . . . . . . + X X X , , + X .   033
                                          032   . . . . . . , , 8 0 0 0 X + + 0   032
                                          031   . . . . . . . + 0 0 0 0 X + 0 +   031
                                          030   . . . . . . , X X X X , 0 R X 0   030
                                          029   . . . . . . , + X + + X X 0 0 X   029
                                          028   . . . . . , , X + + + + X X + ,   028
                                          027   . . . . . , , X + + + + + 0 X ,   027
LABORATORIO DE PLANEACION URBANA          026   . . . . . , , + + + + + + + 0 0   026
                                          025   . . . . . . , , X X X , 0 , 0 0   025
                                          024   . . . . . . . . . 0 0 , , , X +   024
                                          023   . . . . . . . , , , X , , , , ,   023
                                          022   . . . . . . . . . , , , , , , .   022
                                          021   . . . . . . . . . . , + , , , .   021
                                          020   . . . . . . . . . . , + + , , .   020
                                          019   . . . . . . . . . . , , , , , .   019
                                          018   . . . . . . . . . , , 0 X + , .   018
                                          017   . . . . . . . . . , , , + + , .   017
                                          016   , . . . . . . . , , , , , , , .   016
                                          015   . . . , , , , , , X X , , , , .   015
                                          014   . . . , , X X X X + , , 0 0 + .   014
                                          013   . . . . , X 0 0 0 X + 0 0 0 X .   013
                                          012   . . . . , , 0 0 0 0 0 0 0 0 0 .   012
                                          011   . . . , , + , 0 X + + 0 0 0 . .   011
                                          010   . . . , + + , X 0 X X X 0 . . .   010
                                          009   . . . . X , , 0 + X X X . . . .   009
VARIABLES DATO USADAS EN EL MODELO        008   . . . , , , , , 0 0 0 X . . . .   008
                                          007   . . . . , , , , , X 0 X X . . .   007
 NO        NOMBRE                  PESO   006   . . . . , , + , , + 0 X X . . .   006
                                          005   . . . . , X + X , , + 0 0 , . .   005
  2        PORCIENTO DE PENDIENTE      2  004   . . . . , , , + , , , X 0 , . .   004
  5        PORCIENTO DENSIDAD DE ARBOLES 2 003  . . . . . . , 0 , 0 X , X X . .   003
  6        ZONAS LLANAS                2  002   . . . . . . , 0 0 0 0 0 + , . .   002
  4        VEGETACION POR ZONA         1  001   . . . . . . , , , X X X 0 . . .   001
  3        PROXIMIDAD AL AGUA          1
  1        USOS DEL SUELO ACTUALES     1

                                                0000000000000000
                                                0000000001111111
                                                1234567890123456


NIVELES          0          1          2          3          4          5          6          7          8          9
           ================================================================================================================
           . . . . . . . , , , , , , , , , , , , + + + + + + X X X X X 0 0 0 0 0 0 0 0 0 0 0 M M M M M H H H H H X X X X X
           . . . . . . . , , , , , , , , , , , , + + + + + + X X X X X 0 0 0 0 0 0 0 0 0 0 0 M M M M M H H H H H X X X X X
SIMBOLOS   . . 0 . . . . , 1 , , , , , 2 , , , + + 3 + + + X 4 X X X 0 0 5 0 0 0 0 0 0 0 0 M M 7 M M H H H H H X X 9 X X X
           . . . . . . . , , , , , , , , , , , , + + + + + + X X X X X 0 0 0 0 0 0 0 0 0 0 0 M M M M M H H H H H X X X X X
           . . . . . . . , , , , , , , , , , , , # # + + # + X X X X X 0 0 0 0 0 0 0 0 0 0 0 M M M M M H H H H H X X X X X
           ================================================================================================================
FRECUENCIA      416        116         78         73         82         37         19          8          1          2
```

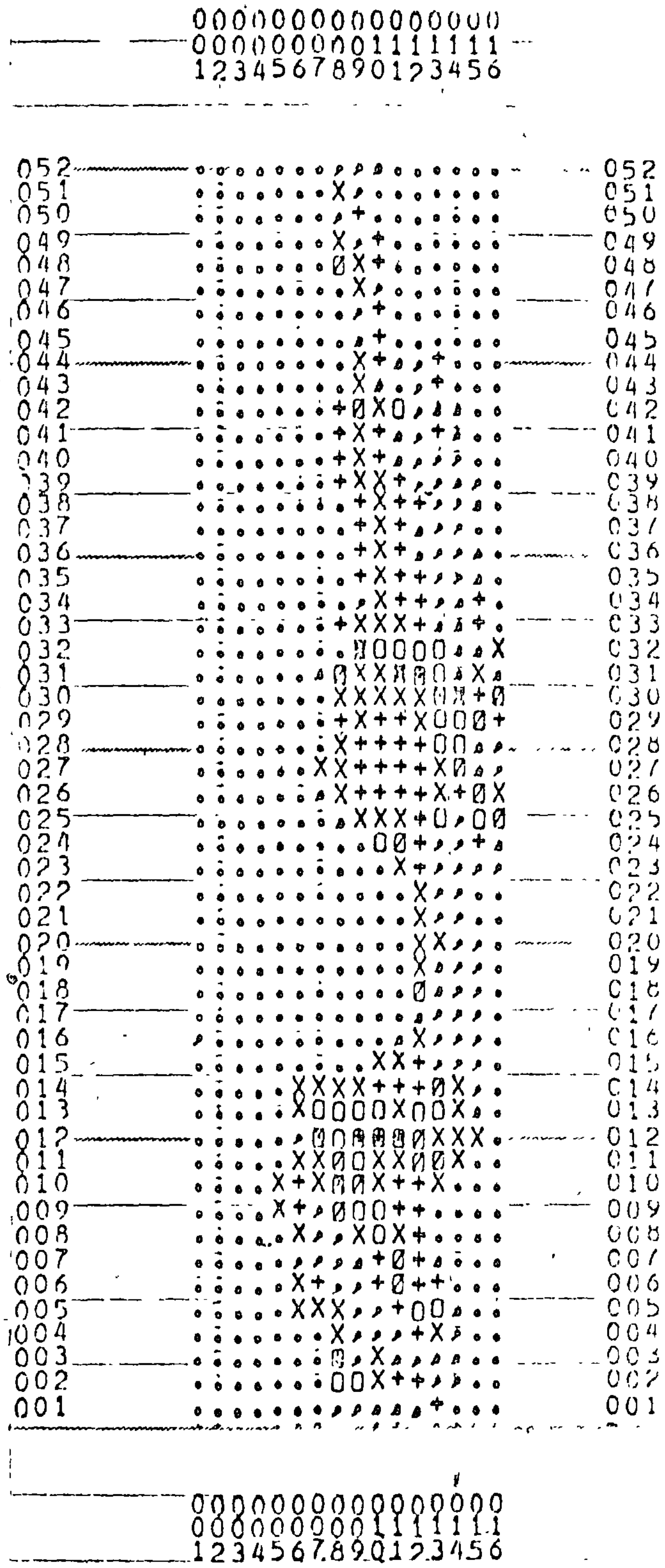
```
               ***MODELO DE ATRACTIVIDAD NUMERO   6***
LOS INDICES DE ATRACTIVIDAD SON PUESTOS EN EL ARCHIVO EN DISCOLAEP/ATTR 6
EL NUMERO DE VARIABLES EN EL MODELO ES  5
ESTAS SON LAS TARJETAS DE LA HOJA DE CODIFICACION USADA EN EL MODELO :
        2          0 0 0 9 0 8 0 4 0 0       2
        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

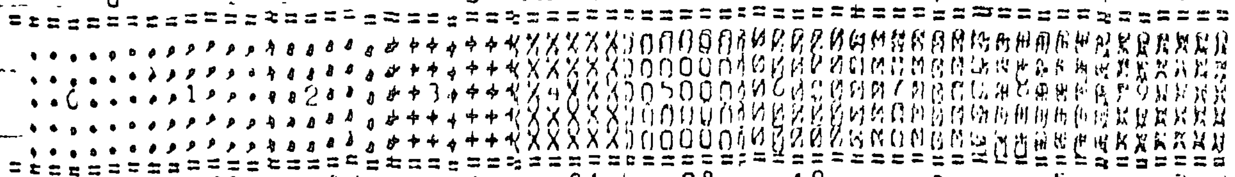


VARIABLES DATO USADAS EN EL MODELO

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
FRECUENCIA	497	73	42	76	81	28	19	3	5	3

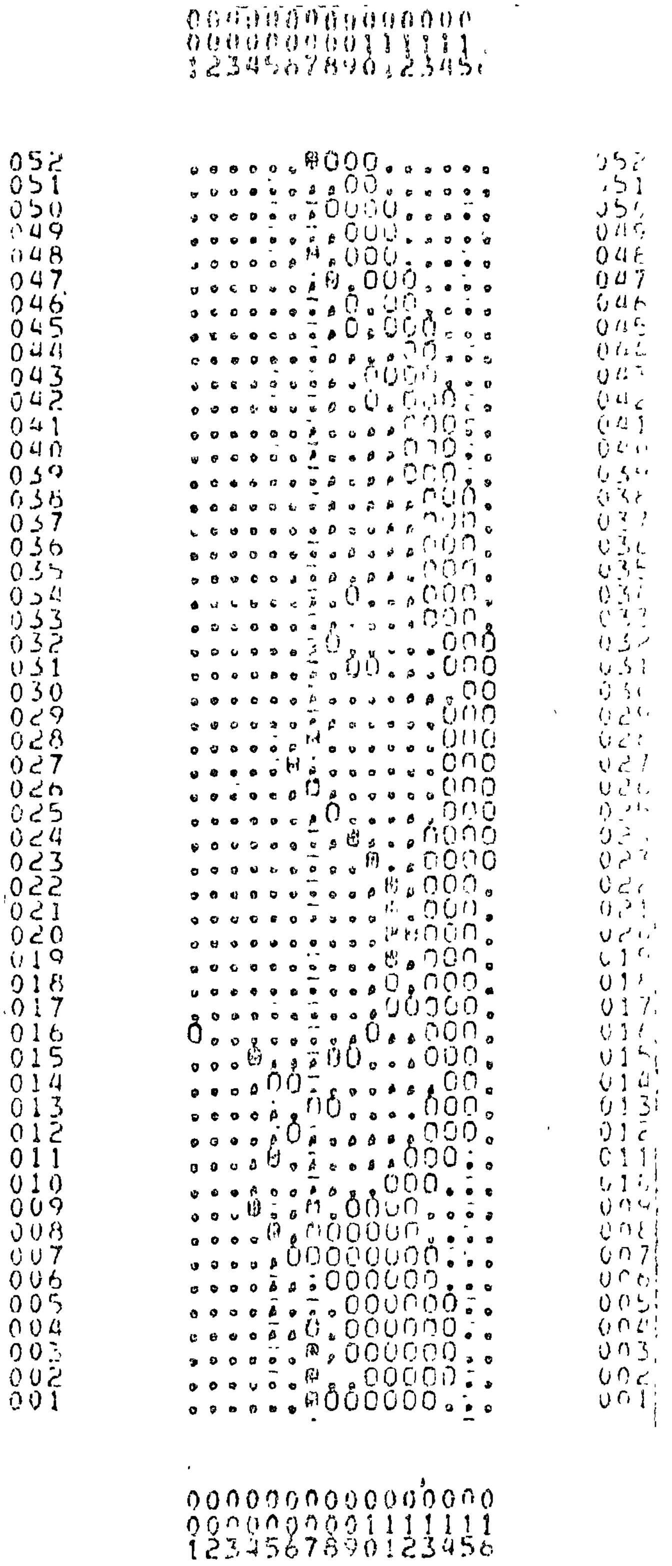
SIMBOLOS



```
                     RUN MAPAS
                     RUN IMPACTO.




@@      @@      @@
@@      @@@    @@@
@@      @@@@  @@@@
@@      @@ @@@@ @@
@@      @@  @@  @@
@@      @@      @@
@@      @@      @@
@@      @@      @@
@@      @@      @@
@@      @@      @@
                  @@@@@@@@     @@@@@@@@     @@@@@@
                  @@@@@@@@@   @@@@@@@@@@   @@@@@@@@
                  @@     @@@  @@      @@  @@@    @@@
                  @@      @@  @@      @@  @@      @@
                  @@     @@@  @@@@@@@@@@  @@
                  @@@@@@@@@   @@@@@@@@@@  @@
                  @@@@@@@@    @@      @@  @@      @@
                  @@          @@      @@  @@@    @@@
                  @@          @@      @@   @@@@@@@@
                  @@          @@      @@    @@@@@@
                                                @@@@@@@@@@    @@@@@@      @@@@@@
                                                @@@@@@@@@@   @@@@@@@@    @@@@@@@@
                                                    @@      @@      @@  @@@    @@@
                                                    @@      @@      @@   @@@
                                                    @@      @@      @@     @@@
                                                    @@      @@      @@       @@@
                                                    @@      @@      @@         @@
                                                    @@      @@      @@  @@@    @@@
                                                    @@       @@@@@@@@    @@@@@@@@
                                                    @@        @@@@@@      @@@@@@

                         CONTORNO IRREGULAR
                         ------------------



            *** NUMERO DEL MODELO DE IMPACTO   1 ***
IMPACTO QUE RESULTA PARA EL PLAN DE EVALUACIONCOLOCADO EN EL ARCHIVO DE DISCO
IMPACTO QUE RESULTA PARA EL MAPA GRID QUE ESTA COLOCADO EN ELARCHIVO DE DISCO
```

```
                        TITULO DEL MAPA
                        ===============


             ODELO DE IMPACTO Ñ 1:    EROSION

             REA DE ESTUDIO    VALLE DEL PACIFICO

             ABORATORIO 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
   7     LOS SIMBOLOS SON
           .,,+X00000123456789          1 CARACTERES EN SENTIDO HORIZONTAL
            .+X0000A
               /X*X
                 +V


  10     EL TEXTO DEL MAPA ES               ===========
                    VARIABLES DATO USADAS EN EL MODELO

                       2    PORCIENTO DE PENDIENTE
                       4    VEGETACION POR ZONA
                       8    SUELOS POR TIPO
  13    LA NUMERACION DE LA MALLA COMIENZA EN      1     52
  14    SE SUPONE QUE LOS DATOS ESTAN PRE-ESCALADOS



LEGENDA: EL SIMBOLO MAS OBSCURO CORRESPONDE AL IMPACTO MAYOR

LOS VALORES DEL IMPACTO PARA LOS 3 GRUPOS DEL USO DEL SUELO SON:

             0           LOS TRES GRUPOS DE US COMPATIBLES
             1           GRUPO DE US 111 MODERADO
             2           GRUPO DE US 111 MODERADO, GRUPO 11 MODERADO
             3           LOS TRES GRUPOS DE US MODERADO
             4           GRUPO DE US 111 SEVERO
             5           GRUPO DE US 111 SEVERO, GRUPO 11 SEVERO
             6           LOS TRES GRUPOS DE US SEVERO
             7           GRUPO DE US 111 TERMINAL
             8           GRUPO DE US 111 TERMINAL, GRUPO 11 TERMINAL
             9           LOS TRES GRUPOS DE US TERMINAL

USO DEL SUELO 1          USO DEL SUELO 11          USO DEL SUELO 111

ACTIVIDADES DE ESPARC.   CAMINOS                   SERVICIO INF. TURISTICA
CAMINATA                 ESTACIONAMIENTOS          CABALLERIZAS
                         ESTRUCTURAS               ESTAC./ TRAILERS
```

```
MODELO DE IMPACTO N 18    EROSION


AREA DE ESTUDIO    VALLE DEL PACIFICO


LABORATORIO DE PLANEACION URBANA



VARIABLES DATO USADAS EN EL MODELO

    2    PORCIENTO DE PENDIENTE
    4    VEGETACION POR ZONA
    9    SUELOS POR TIPO


LEGENDA: EL SIMBOLO MAS OBSCURO
        CORRESPONDE AL IMPACTO MAYOR


LOS VALORES DEL IMPACTO PARA LOS
   3 GRUPOS DEL USO DEL SUELO SON:

1  LOS TRES GRUPOS DE US COMPATIBLES
2  GRUPO DE US 111 MODERADO
3  GRUPO DE US 111 MODERADO, GRUPO 11 MODERADO
4  LOS-TRES GRUPOS DE US MODERADO
5  GRUPO DE US 111 SEVERO
6  GRUPO DE US 111 SEVERO, GRUPO 11 SEVERO
7  LOS-TRES GRUPOS DE US SEVERO
8  GRUPO DE US 111 TERMINAL
9  GRUPO DE US 111 TERMINAL, GRUPO 11 TERMINAL
10 LOS-TRES GRUPOS DE US TERMINAL

USO DEL SUELO                    USO DEL SUELO 11

ACTIVIDADES DE ESPARC.           CAMINOS
CAMINATA                         ESTACIONAMIENTOS
                                 ESTRUCTURAS

         USO DEL SUELO 111

         SERVICIO INF. TURISTICA
         CABALLERIZAS
         ESTAC./TRAILERS
```



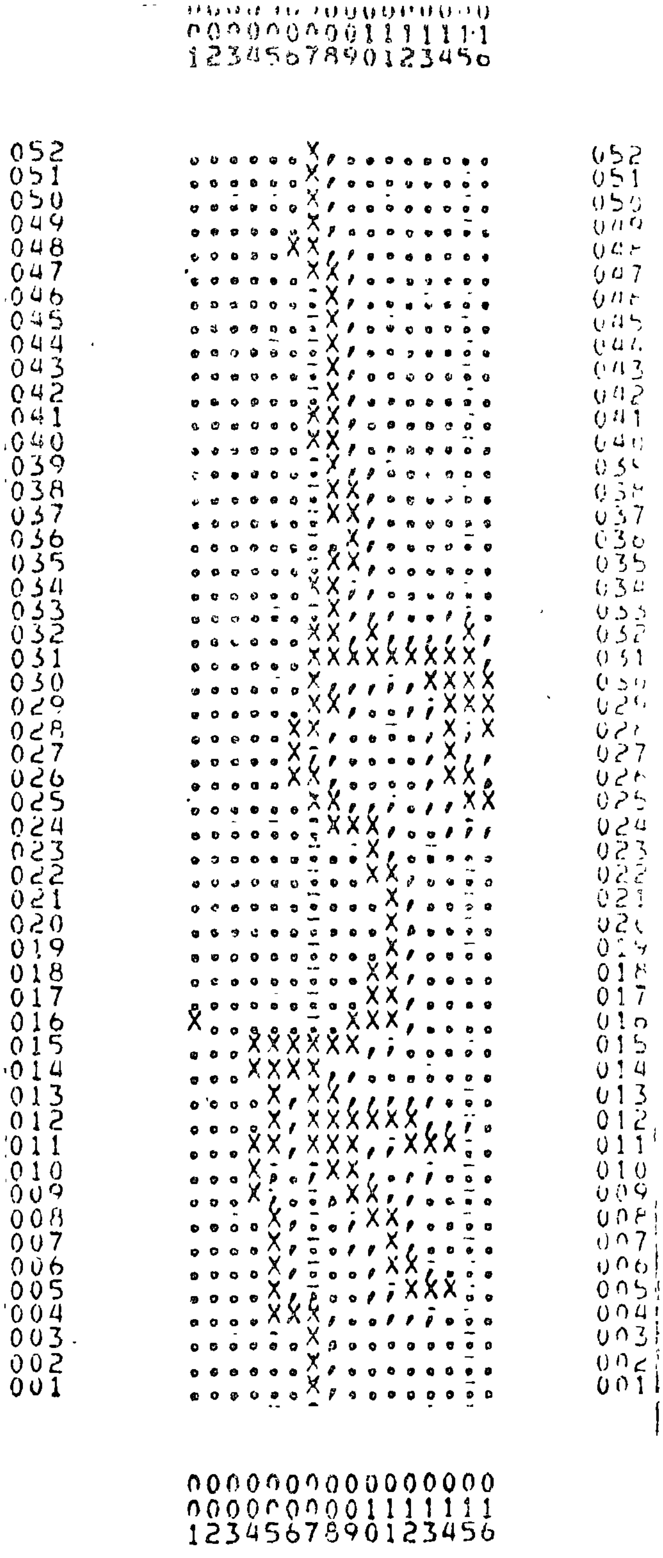
```
NIVELES          0           1            2           5           8
          ==========  ==========  ==========  ==========  ==========
           .........  //////////  ooooooooo   000000000   @@@@@@@@@
           .........  //////////  ooooooooo   000000000   @@@@@@@@@
SIMBOLOS   .........  //////////  ooooooooo   000000000   @@@@@@@@@
           .........  //////////  ooooooooo   000000000   @@@@@@@@@
           .........  //////////  ooooooooo   000000000   @@@@@@@@@
          ==================================================================
FRECUENCIA      509           1          99         203          20
```

```
                 TITULO DEL MAPA
                 ===============



      MODELO DE IMPACTO Ñ 2:      CAMBIO EN LA VISUALIDAD
      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

                           4    VEGETACION POR ZONA
                           5    PORCIENTO DENSIDAD DE ARBOLES
                           2    PORCIENTO DE PENDIENTE



   LEGENDA: EL SIMBOLO MAS OBSCURO CORRESPONDE AL IMPACTO MAYOR
   LOS VALORES DEL IMPACTO PARA LOS 3 GRUPOS DEL USO DEL SUELO SON:
              0           LOS TRES GRUPOS DE US COMPATIBLES
              1           GRUPO DE US 111 MODERADO
              2           GRUPO DE US 111 MODERADO, GRUPO 11 MODERADO
              3           LOS TRES GRUPOS DE US MODERADO
              4           GRUPO DE US 111 SEVERO
              5           GRUPO DE US 111 SEVERO, GRUPO 11 SEVERO
              6           LOS TRES GRUPOS DE US SEVERO
              7           GRUPO DE US 111 TERMINAL
              8           GRUPO DE US 111 TERMINAL, GRUPO 11 TERMINAL
              9           LOS TRES GRUPOS DE US TERMINAL
   USO DEL SUELO 1           USO DEL SUELO 11            USO DEL SUELO 111

   ACTIVIDADES DE ESPARC.    ESTRUCTURAS                 ESTACIONAMIENTOS
   CAMINATA                  VIS                         CAMINOS
   CABALLERIZAS                                          ESTAC./ TRAILERS
```

```
                                                                                0000000000000000
                                                                                0000000001111111
                                                                                1234567890123456

   MODELO DE IMPACTO # 2:
                                                                052             MMMMMMMM+.......    05
                                                                051             MMMMMMMMMM......    05
                              CAMBIO EN LA VISUALIDAD           050             MMMMMMM.MMMM....    05
                                                                049             MMMMMMMMMMM.....    04
                                                                048             MMMMMMMMMMM.....    04
                                                                047             MMMMMMMMMMMM....    04
                                                                046             MMMMMMMMMMMM....    04
                                                                045             MMMMMMMM+M.M....    04
          AREA DE ESTUDIO ___ VALLE DEL PACIFICO                044             MMMMMMMMMMMM....    04
                                                                043             MMMMMMMMMMM+....    04
                                                                042             MMMMMMMMMMM+....    04
                                                                041             MMMMMMMMMMM+....    04
                                                                040             MMMMMMMMMMM+++..    04
            LABORATORIO DE PLANEACION URBANA                    039             MMMMT.MMMMM++.+.    03
                                                                038             MMMMM.MMMMM+++..    03
                                                                037             MMMMM.MMMMMM++M.    03
                                                                036             MMMMMMMMMMMM+MX.
                                                                035             MMMMMMMMMMMM++..     3
                                                                034             MMMMMMMMMMMM++..     3
         VARIABLES DATO USADAS EN EL MODELO                     033             MMMMMMMMMMMMM++.    03
                                                                032             MMMMMMMM.MMMM++.    03
          4    VEGETACION POR ZONA                              031             MMMMMMMMMM..M+.M    03
          5    PORCIENTO DENSIDAD DE ARBOLES                    030             MMMMMMMMMMMMM.+.    03
          2    PORCIENTO DE PENDIENTE                           029             MMMMMMMMMMMMMM.+    02
                                                                028             MMMMMMMMMMMMMMMM    02
                                                                027             MMMMMMMMMMMMMXMM    02
                                                                026             MMMMMMM+MMMMMM..    02
                                                                025             MMMMMMMMMMMM+M+.    02
                                                                024             MMMMMMMMMMMM+M++    02
                                                                023             MMMMMMMMMMMM++M+    02
          LEGENDA: EL SIMBOLO MAS OBSCURO                       022             MMMMMMMMMMMM++M.    02
             CORRESPONDE AL IMPACTO MAYOR                       021             MMMMMMMMMMMM++M.    02
                                                                020             MMMMMMMMMMMM+++.    02
                                                                019             MMMMMMMMMMMMM++.    01
                                                                018             MMMMMMMMMMMM+++.    01
                                                                017             MMMMMMMMMMMMM++.    01
         LOS VALORES DEL IMPACTO PARA LOS                       016             XMMMMMMMMMMM+++.    01
           3 GRUPOS DEL USO DEL SUELO SON:                      015             MMMMMMMMX.XM+++.    01
                                                                014             MMMMMMMMMMXM+++.    01
                                                                013             MMMMMM+MMMMM+++.    01
                                                                012             MMMMMMMM.M......    012
                                                                011             MMMMMMMM+MMM....     11
  1  LOS TRES GRUPOS DE US COMPATIBLES                          010             MMMMMMM+MMMM....    010
  2  GRUPO DE US 111 MODERADO -                                 009             MMMMMMMMMM++....    009
  3  GRUPO DE US 111 MODERADO, GRUPO 11 MODERADO                008             MMMMMMM+++.+....    008
  4  LOS-TRES GRUPOS DE US MODERADO                             007             MMMMMM+M++.++...    007
  5  GRUPO DE US 111 SEVERO                                     006             MMMMMMMMM+.++...    006
  6  GRUPO DE US 111 SEVERO, GRUPO 11 SEVERO                    005             MMMMMMM+MM++++..    005
  7  LOS-TRES GRUPOS DE US SEVERO                               004             MMMMMMMM++M+.+..    004
  8  GRUPO DE US 111 TERMINAL                                   003             MMMMMMMXMMM+++..    003
  9  GRUPO DE US 111 TERMINAL, GRUPO 11 TERMINAL                002             MMMMMMMM++MM++..    002
     LOS-TRES GRUPOS DE US TERMINAL                             001             MMMMMMMMM+++....    001


  USO DEL SUELO 1                     USO DEL SUELO 11                          0000000000000000
                                                                                0000000001111111
  ACTIVIDADES DE ESPARC.              ESTRUCTURAS                               1234567890123456
  INATA                               VIS
  CABALLERIZAS

                      USO DEL SUELO 111

                      ESTACIONAMIENTOS
                      CAMINOS
                      ESTAC./ TRAILERS


      NIVELES                0            2            3            7
                    ==================================================

                     .........    ooooooooo    +++++++++    MMMMMMMMM
                     .........    ooooooooo    +++++++++    MMMMMMMMM
      SIMBOLOS       ....0....    oooo2oooo    ++++3++++    MMMM7MMMM
                     .........    ooooooooo    +++++++++    MMMMMMMMM
                     .........    ooooooooo    +++++++++    MMMMMMMMM
                    ==================================================
      FRECUENCIA        123           11           99          586
```

```
                    TITULO DEL MAPA
                    ===============


MODELO DE IMPACTO Ñ 3:    CONTAMINACION DE LA SUPERFICIE DEL AGUA
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

                        0    PROXIMIDAD AL AGUA
                        8    SUELOS POR TIPO
                        4    VEGETACION POR ZONA



LEGENDA: EL SIMBOLO MAS OBSCURO CORRESPONDE AL IMPACTO MAYOR
LOS VALORES DEL IMPACTO PARA LOS 3 GRUPOS DEL USO DEL SUELO SON:
           0          LOS TRES GRUPOS DE US COMPATIBLES
           1          GRUPO DE US 111 MODERADO
           2          GRUPO DE US 111 MODERADO, GRUPO 11 MODERADO
           3          LOS TRES GRUPOS DE US MODERADO
           4          GRUPO DE US 111 SEVERO
           5          GRUPO DE US 111 SEVERO, GRUPO 11 SEVERO
           6          LOS TRES GRUPOS DE US SEVERO
           7          GRUPO DE US 111 TERMINAL
           8          GRUPO DE US 111 TERMINAL, GRUPO 11 TERMINAL
           9          LOS TRES GRUPOS DE US TERMINAL
ACTIVIDADES DE ESPARC.  ESTACIONAMIENTOS              CABALLERIZAS
CAMINATA                ESTRUCTURAS                   ESTAC./ TRAILERS
                        CAMINOS
                        VIS
```

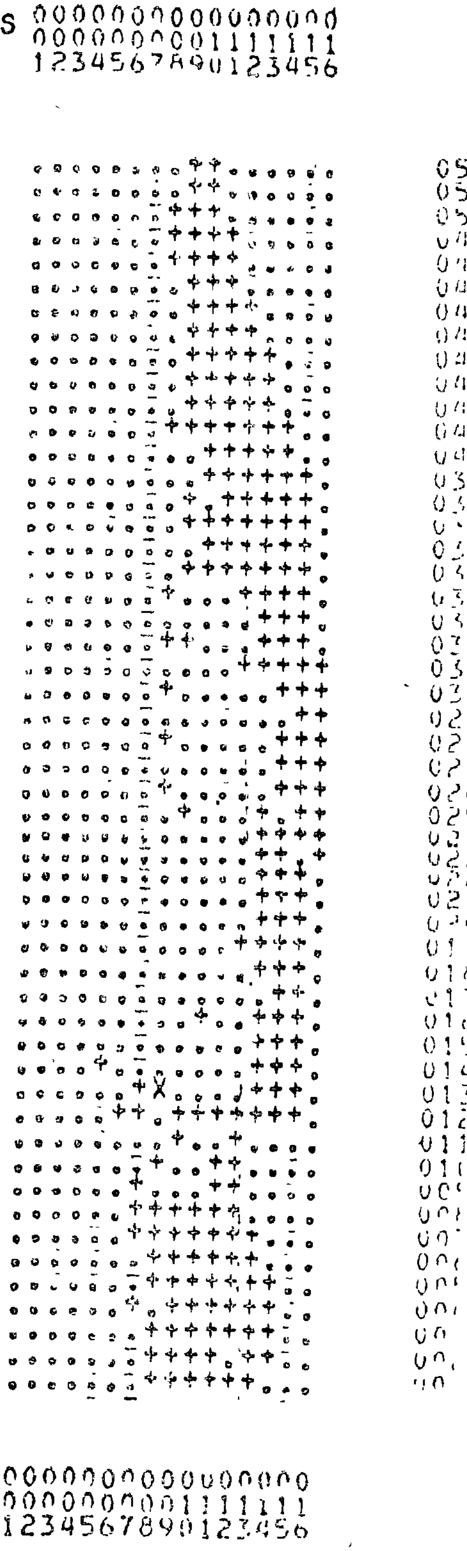
```
              MODELO DE IMPACTO Ñ 3:


      CONTAMINACION DE LA SUPERFICIE DEL AGUA



     AREA DE ESTUDIO     VALLE DEL PACIFICO


             LABORATORIO DE PLANEACION URBANA


     VARIABLES DATO USADAS EN EL MODELO
        9    PROXIMIDAD AL AGUA
        8    SUELOS POR TIPO
        4    VEGETACION POR ZONA


     LEGENDA: EL SIMBOLO MAS OBSCURO

      CORRESPONDE AL IMPACTO MAYOR


     LOS VALORES DEL IMPACTO PARA LOS

      3 GRUPOS DEL USO DEL SUELO SON:


 0  LOS TRES GRUPOS DE US COMPATIBLES
 1  GRUPO DE US 111 MODERADO -
 2  GRUPO DE US 111 MODERADO, GRUPO 11 MODERADO
 3  LOS-TRES GRUPOS DE US MODERADO
 4  GRUPO DE US 111 SEVERO
 5  GRUPO DE US 111 SEVERO, GRUPO 11 SEVERO
 6  LOS-TRES GRUPOS DE US SEVERO
 7  GRUPO DE US 111 TERMINAL
 8  GRUPO DE US 111 TERMINAL, GRUPO 11 TERMINAL
 9  LOS-TRES GRUPOS DE US TERMINAL


    ESTACIONAMIENTOS          CABALLERIZAS
    ESTRUCTURAS               ESTAC./ TRAILERS
    CAMINOS
    VIS

            ACTIVIDADES DE ESPARC.
            CAMINATA
```



```
   NIVELES           0             1            2            4
             ===========================================================
              o o o o o o o o o  / / / / / / / / /  ø ø ø ø ø ø ø ø ø  XXXXXXXXX
              o o o o o o o o o  / / / / / / / / /  ø ø ø ø ø ø ø ø ø  XXXXXXXXX
   SIMBOLOS   o o o o 0 o o o o  / / / / 1 / / / /  ø ø ø ø 2 ø ø ø ø  XXXX4XXXX
              o o o o o o o o o  / / / / / / / / /  ø ø ø ø ø ø ø ø ø  XXXXXXXXX
              o o o o o o o o o  / / / / / / / / /  ø ø ø ø ø ø ø ø ø  XXXXXXXXX.
             ===========================================================
   FRECUENCIA      581            111            6           134
```

```
                         TITULO DEL MAPA
                         ----------------


      MODELO DE IMPACTO Ñ 4:     VULNERABILIDAD A FUEGOS
      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

                         4    VEGETACION POR ZONA
                         3    ORIENTACION
                         5    PORCIENTO DENSIDAD DE ARBOLES



LEGENDA: EL SIMBOLO MAS OBSCURO CORRESPONDE AL IMPACTO MAYOR
LOS VALORES DEL IMPACTO PARA LOS 3 GRUPOS DEL USO DEL SUELO SON:

             0           LOS TRES GRUPOS DE US COMPATIBLES
             1           GRUPO DE US 111 MODERADO
             2           GRUPO DE US 111 MODERADO, GRUPO 11 MODERADO
             3           LOS TRES GRUPOS DE US MODERADO
             4           GRUPO DE US 111 SEVERO
             5           GRUPO DE US 111 SEVERO, GRUPO 11 SEVERO
             6           LOS TRES GRUPOS DE US SEVERO
             7           GRUPO DE US 111 TERMINAL
             8           GRUPO DE US 111 TERMINAL, GRUPO 11 TERMINAL
             9           LOS TRES GRUPOS DE US TERMINAL
USO DEL SUELO 1          USO DEL SUELO 11          USO DEL SUELO 111

ESTACIONAMIENTOS         ACTIVIDADES DE ESPARC.    ESTAC./ TRAILERS
CABALLERIZAS             ESTRUCTURAS               CAMINATA
CAMINOS                  VIS
```

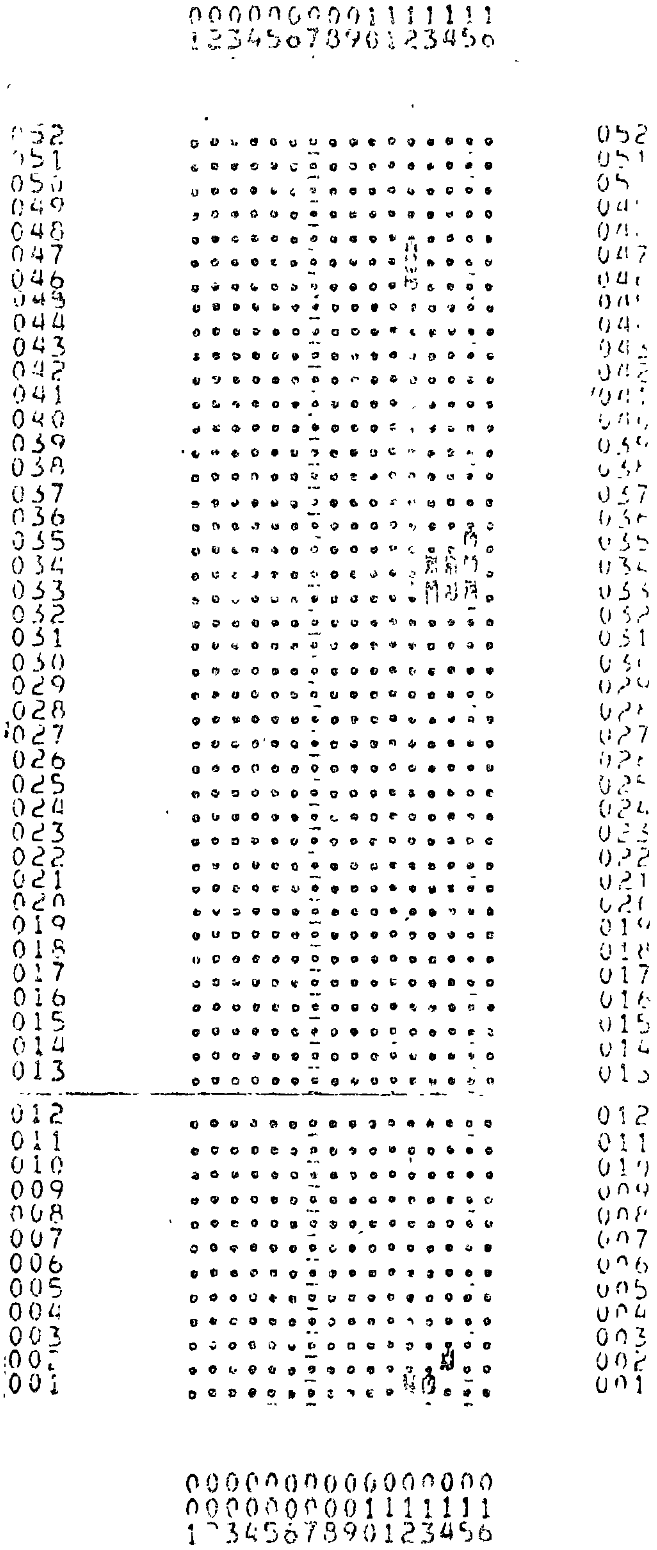
```
                 LABORATORIO DE PLANEACION URBANA

              AREA DE ESTUDIO    VALLE DEL PACIFICO

   MODELO DE IMPACTO N 4:    VULNERABILIDAD A FUEGOS

     VARIABLES DATO USADAS EN EL MODELO :
                                                    052
                                                    051
        4    VEGETACION POR ZONA                    050
        3    ORIENTACION                            049
        5    PORCIENTO DENSIDAD DE ARBOLES          048
                                                    047
                                                    046
                                                    045
                                                    044
     LEYENDA: EL SIMBOLO MAS OBSCURO                043
                                                    042
                                                    041
     CORRESPONDE AL IMPACTO MAYOR                   040
                                                    039
                                                    038
                                                    037
     LOS VALORES DEL IMPACTO PARA LOS               036
                                                    035
                                                    034
                                                    033
     3 GRUPOS DEL USO DEL SUELO SON:                032
                                                    031
                                                    030
                                                    029
  LOS TRES GRUPOS DE US COMPATIBLES                 028
  GRUPO DE US 111 MODERADO                          027
  GRUPO DE US 111 MODERADO, GRUPO 11 MODERADO       026
  LOS TRES GRUPOS DE US MODERADO                    025
  GRUPO DE US 111 SEVERO                            024
  GRUPO DE US 111 SEVERO, GRUPO 11 SEVERO           023
  LOS TRES GRUPOS DE US SEVERO                      022
  GRUPO DE US 111 TERMINAL                          021
  GRUPO DE US 111 TERMINAL, GRUPO 11 TERMINAL       020
  LOS TRES GRUPOS DE US TERMINAL                    019
                                                    018
          USO DEL SUELO 1                           017
                                                    016
          ESTACIONAMIENTOS                          015
          CABALLERIZAS                              014
          CAMINOS                                   013
                                                    012
                                                    011
          USO DEL SUELO 11                          010
                                                    009
          ACTIVIDADES DE ESPARC.                    008
          ESTRUCTURAS                               007
          VIS                                       006
                                                    005
          USO DEL SUELO 111                         004
                                                    003
          ESTAC./ TRAILERS                          002
          CAMINATA                                  001
```



```
  NIVELES          0                 3                  4
           ==========:      :==========:      :==========:

            ..........       +++++++++         XXXXXXXXX
  SIMBOLOS  ..........       +++++++++         XXXXXXXXX
            ....0.....       ++++3++++         XXXX4XXXX
            ..........       +++++++++         XXXXXXXXX
            ..........       +++++++++         XXXXXXXXX
           ==========:      :==========:      :==========:
  FRECUENCIA   598              233
```

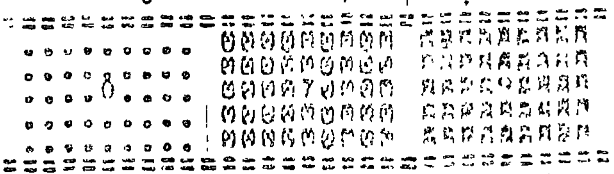
```
                    TITULO DEL MAPA
                    ---------------



        MODELO DE IMPACTO Ñ 5:    DESLIZAMIENTOS DE TIERRA.
        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

                          7   ACCIDENTES GEOLOGICOS
                          2   PORCIENTO DE PENDIENTE
                          4   VEGETACION POR ZONA



    LEGENDA: EL SIMBOLO MAS OBSCURO CORRESPONDE AL IMPACTO MAYOR
    LOS VALORES DEL IMPACTO PARA LOS 3 GRUPOS DEL USO DEL SUELO SON:
              0             LOS TRES GRUPOS DE US COMPATIBLES
              1             GRUPO DE US 111 MODERADO
              2             GRUPO DE US 111 MODERADO, GRUPO 11 MODERADO
              3             LOS TRES GRUPOS DE US MODERADO
              4             GRUPO DE US 111 SEVERO
              5             GRUPO DE US 111 SEVERO, GRUPO 11 SEVERO
              6             LOS TRES GRUPOS DE US SEVERO
              7             GRUPO DE US 111 TERMINAL
              8             GRUPO DE US 111 TERMINAL, GRUPO 11 TERMINAL
              9             LOS TRES GRUPOS DE US TERMINAL
    USO DEL SUELO 1          USO DEL SUELO 11         USO DEL SUELO 111

    ACTIVIDADES DE ESPARC.   CABALLERIZAS             ESTACIONAMIENTOS
    CAMINATA                                          ESTRUCTURAS
                                                      VIS
                                                      CAMINOS
                                                      ESTAC,/ TRAILERS
```

```
MODELO DE IMPACTO N 3:

          DESLIZAMIENTOS DE TIERRA


AREA DE ESTUDIO      VALLE DEL PACIFICO


LABORATORIO DE PLANEACION URBANA




VARIABLES DATO USADAS EN EL MODELO
  7   ACCIDENTES GEOLOGICOS
  2   PORCIENTO DE PENDIENTE
  4   VEGETACION POR ZONA


LEGENDA: EL SIMBOLO MAS OBSCURO
CORRESPONDE AL IMPACTO MAYOR

LOS VALORES DEL IMPACTO PARA LOS
3 GRUPOS DEL USO DEL SUELO SON:


0  LOS TRES GRUPOS DE US COMPATIBLES
1  GRUPO DE US 111 MODERADO
2  GRUPO DE US 111 MODERADO, GRUPO 11 MODERADO
3  LOS TRES GRUPOS DE US MODERADO
4  GRUPO DE US 111 SEVERO
5  GRUPO DE US 111 SEVERO, GRUPO 11 SEVERO
6  LOS TRES GRUPOS DE US SEVERO
7  GRUPO DE US 111 TERMINAL
8  GRUPO DE US 111 TERMINAL, GRUPO 11 TERMINAL
9  LOS TRES GRUPOS DE US TERMINAL
```



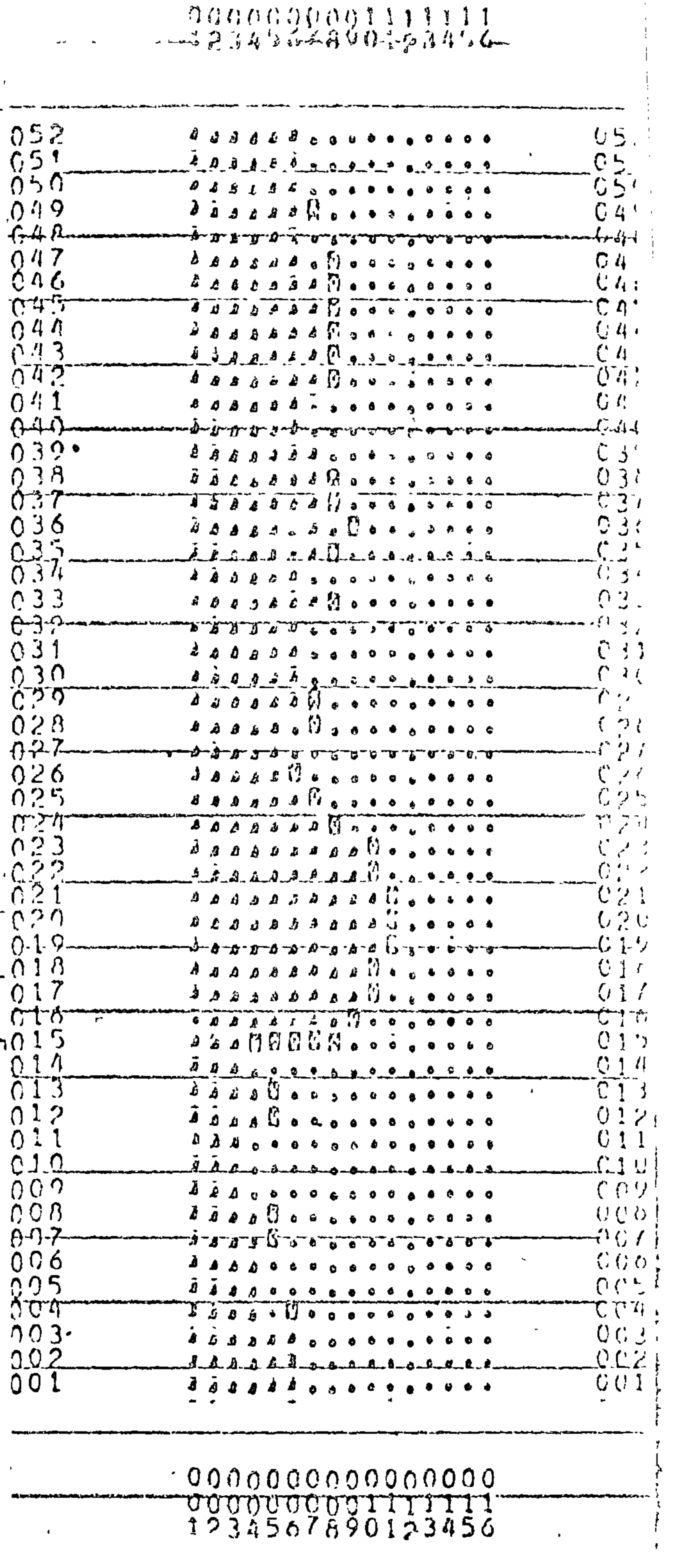
```
USO DEL SUELO 1              USO DEL SUELO 11

ACTIVIDADES DE ESPARC.       CABALLERIZAS
CAMINATA

     USO DEL SUELO 111

     ESTACIONAMIENTOS
     ESTRUCTURAS
     VIS
     CAMINOS
     ESTAC./ TRAILERS
```

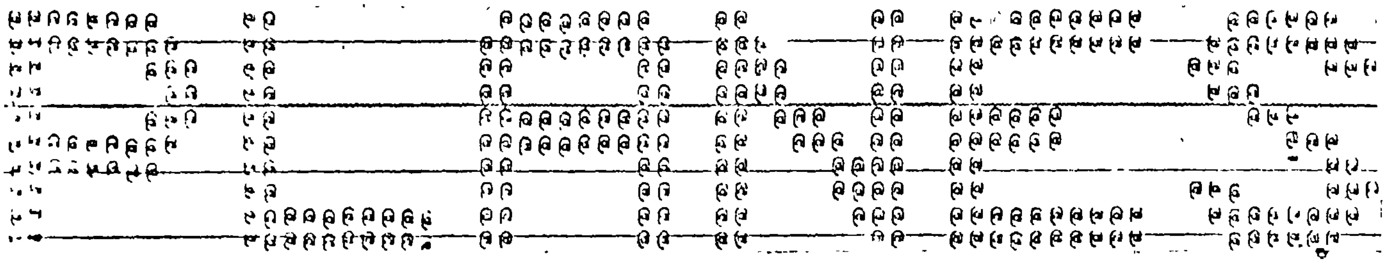
```
                   NIVELES          0          7          9

                   SIMBOLOS

                   FRECUENCIA     820          4          8
```



```
                 TITULO DEL MAPA
                 ===============


       MODELO DE IMPACTO N 6:    BASURA EN EL MAR
       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

                    4    VEGETACION POR ZONA
                    6    ZONAS LLANAS
                    9    PROXIMIDAD AL AGUA



LEGENDA: EL SIMBOLO MAS OBSCURO CORRESPONDE AL IMPACTO MAYOR
LOS VALORES DEL IMPACTO PARA LOS 3 GRUPOS DEL USO DEL SUELO SON:
            0           LOS TRES GRUPOS DE US COMPATIBLES
            1           GRUPO DE US 111 MODERADO
            2           GRUPO DE US 111 MODERADO, GRUPO 11 MODERADO
            3           LOS TRES GRUPOS DE US MODERADO
            4           GRUPO DE US 111 SEVERO
            5           GRUPO DE US 111 SEVERO, GRUPO 11 SEVERO
            6           LOS TRES GRUPOS DE US SEVERO
            7           GRUPO DE US 111 TERMINAL
            8           GRUPO DE US 111 TERMINAL, GRUPO 11 TERMINAL
            9           LOS TRES GRUPOS DE US TERMINAL
USO DEL SUELO 1         USO DEL SUELO 11         USO DEL SUELO 111

ACTIVIDADES DE ESPARC.  ESTRUCTURAS              ESTACIONAMIENTOS
CAMINATA                                         VIS
CABALLERIZAS                                     CAMINOS
                                                 ESTAC./ TRAILERS
```

```
MODELO DE IMPACTO # 6:     BASURA EN EL MAR

     AREA DE ESTUDIO    VALLE DEL PACIFICO

          LABORATORIO DE PLANEACION URBANA
     VARIABLES DATO USADAS EN EL MODELO

       4    VEGETACION POR ZONA
       6    ZONAS LLANAS
       9    PROXIMIDAD AL AGUA

     LEGENDA: ~~EL SIMBOLO MAS OBSCURO~~
        ~~CORRESPONDE AL IMPACTO MAYOR~~

     LOS VALORES DEL IMPACTO PARA LOS
      3 GRUPOS DEL USO DEL SUELO SON:

 1  LOS TRES GRUPOS DE US COMPATIBLES
 2  GRUPO DE US 111 MODERADO
 3  GRUPO DE US 111 MODERADO, GRUPO 11 MODERADO
 4  LOS TRES GRUPOS DE US MODERADO
 5  GRUPO DE US 111 SEVERO
 6  GRUPO DE US 111 SEVERO, GRUPO 11 SEVERO
 7  LOS TRES GRUPOS DE US SEVERO
 8  ~~GRUPO DE US 111 TERMINAL~~
    GRUPO DE US 111 TERMINAL, GRUPO 11 TERMINAL
 9  LOS TRES GRUPOS DE US TERMINAL.

USO DEL SUELO 1                 USO DEL SUELO 11
ACTIVIDADES DE ESPARC.          ESTRUCTURAS
[illegible]
~~[illegible]~~

          USO DEL SUELO 111
          ESTACIONAMIENTOS
          VIAS
          ~~CAMINOS~~
          ESTAC./ TRAILERS

 NIVELES            0          2           7
        .=================================.

 SIMBOLOS

 FRECUENCIA      480         317          35
```





```
CONTORNO IRREGULAR
------------------



TITULO DEL MAPA
---------------
EVALUACION DE LA ATRACTIVIDAD DEL PLAN # 2
AREA DE ESTUDIO     VALLE DEL 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
          .,,+XOOOOO0123456789     Y    1 CARACTERES EN SENTIDO HORIZONTAL
            .+XOOOA
               /X*X
                +V

 10     EL TEXTO DEL MAPA ES          -----------


 13    LA NUMERACION DE LA MALLA COMIENZA EN       1      52
 14    SE SUPONE QUE LOS DATOS ESTAN PRE-ESCALADOS
```

```
TITULO DEL MAPA
EVAL DEL IMPACTO DEL PLAN 1 :CONTAMINACION DE LA SUP. DE AGUA.
AREA DE ESTUDIO    VALLE DEL PACIFICO
LABORATORIO DE PLANEACION URBANA

OPCIONES USADAS PARA ESTE MAPA

 10      EL TEXTO DEL MAPA ES
LOS USOS DEL SUELO SE AGRUPARON EN LOS GRUPOS:
USO DEL SUELO 1            USO DEL SUELO 11              USO DEL SUELO 111

ACTIVIDADES DE ESPARC.     ESTACIONAMIENTOS              CABALLERIZAS
CAMINATA                   ESTRUCTURAS                   ESTAC./ TRAILERS
                           CAMINOS
                           VIS

TITULO DEL MAPA
EVALUACION DEL IMPACTO DEL PLAN # 1 EN: EROSION
AREA DE ESTUDIO    VALLE DEL PACIFICO
LABORATORIO DE PLANEACION URBANA

OPCIONES USADAS PARA ESTE MAPA

 10      EL TEXTO DEL MAPA ES
LOS USOS DEL SUELO SE AGRUPARON EN LOS GRUPOS:
USO DEL SUELO 1            USO DEL SUELO 11              USO DEL SUELO 111

ACTIVIDADES DE ESPRC.      CAMINOS                       SERVICIO INF. TUR
CAMINATA                   ESTACIONAMIENTOS              CABALLERIZAS
                           ESTRUCTURAS                   ESTAC./ TRAILERS

TITULO DEL MAPA
EVALUACION DEL IMPACTO DEL PLAN # 1 EN: CAMBIO EN LA VISUALIDAD
AREA DE ESTUDIO    VALLE DEL PACIFICO
LABORATORIO DE PLANEACION URBANA

OPCIONES USADAS PARA ESTE MAPA

 10      EL TEXTO DEL MAPA ES
LOS USOS DEL SUELO SE AGRUPARON EN LOS GRUPOS:
USO DEL SUELO 1            USO DEL SUELO 11              USO DEL SUELO 111

ACTIVIDADES DE ESPARC.     ESTRUCTURAS                   ESTACIONAMIENTOS
CAMINATA                   VIS                           CAMINOS
CABALLERIZAS                                             ESTAC./ TRAILERS
```

VALUACION DEL IMPACTO DEL PLAN Ñ 1 EN: EROSION

AREA DE ESTUDIO   VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

LOS USOS DEL SUELO SE AGRUPARON EN LOS GRUPOS:

USO DEL SUELO 1

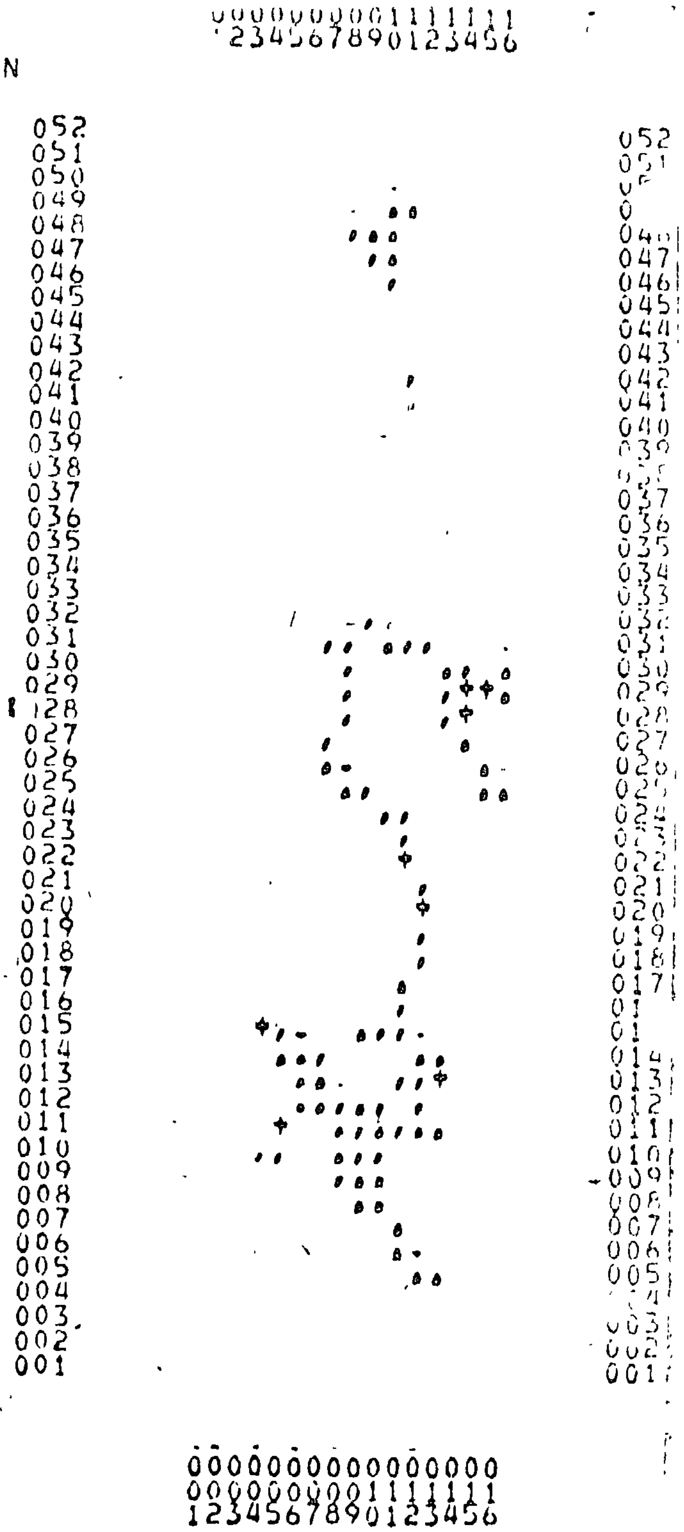
ACTIVIDADES DE ESPARC.
CAMINATA

USO DEL SUELO 11

CAMINOS
ESTACIONAMIENTOS
ESTRUCTURAS

USO DEL SUELO 111

SERVICIO INF. TURISTICA
CABALLERIZAS
ESTAC./ TRAILERS



NIVELES	0	1	2	3	4	5
SIMBOLOS	.	/	'	+	X	0
FRECUENCIA	0	42	38	8	0	0

EVAL DEL IMPACTO DEL PLAN

CONTAMINACION DE LA SUP. DE AGUA.

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

LOS USOS DEL SUELO SE AGRUPARON EN LOS GRUPOS:

USO DEL SUELO 1

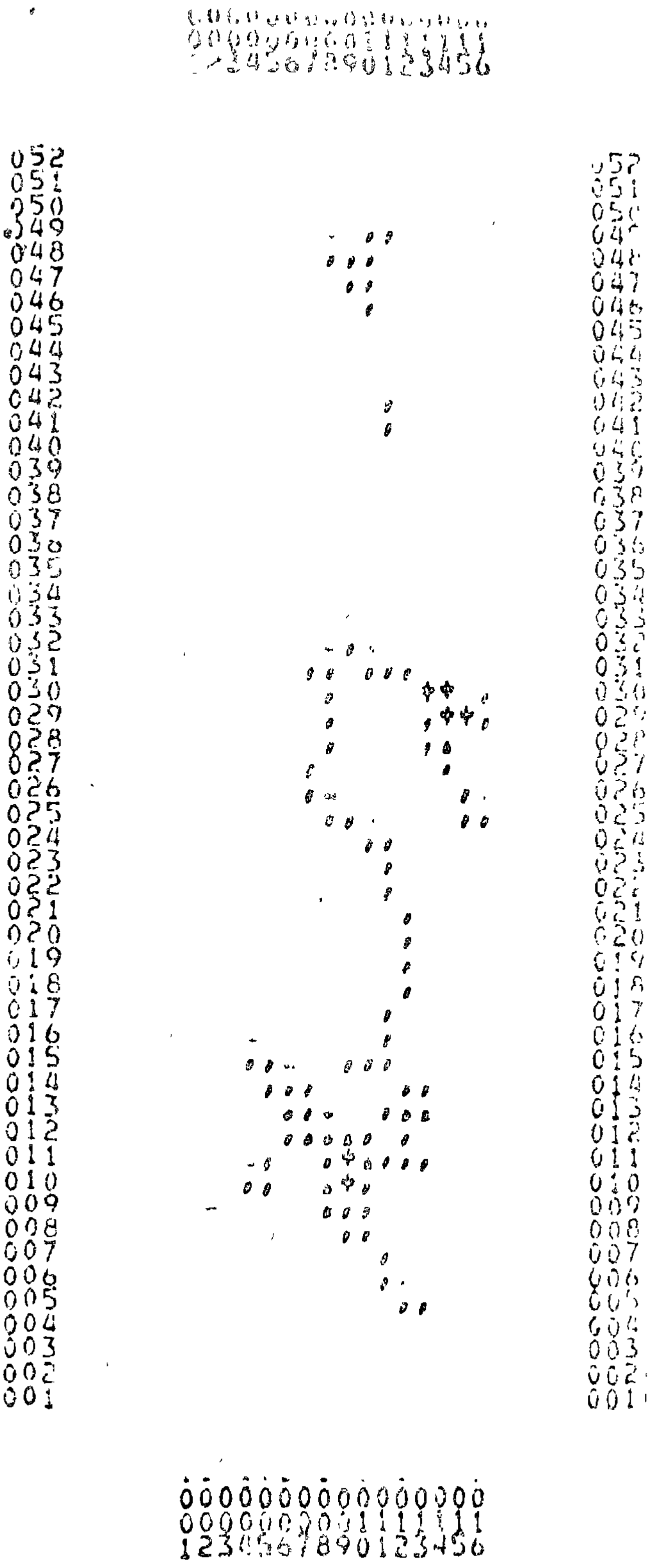
ACTIVIDADES DE ESPARC.
CAMINATA

USO DEL SUELO 11

ESTACIONAMIENTOS
ESTRUCTURAS
CAMINOS
VIS

USO DEL SUELO 111

CABALLERIZAS
ESTAC./ TRAILERS



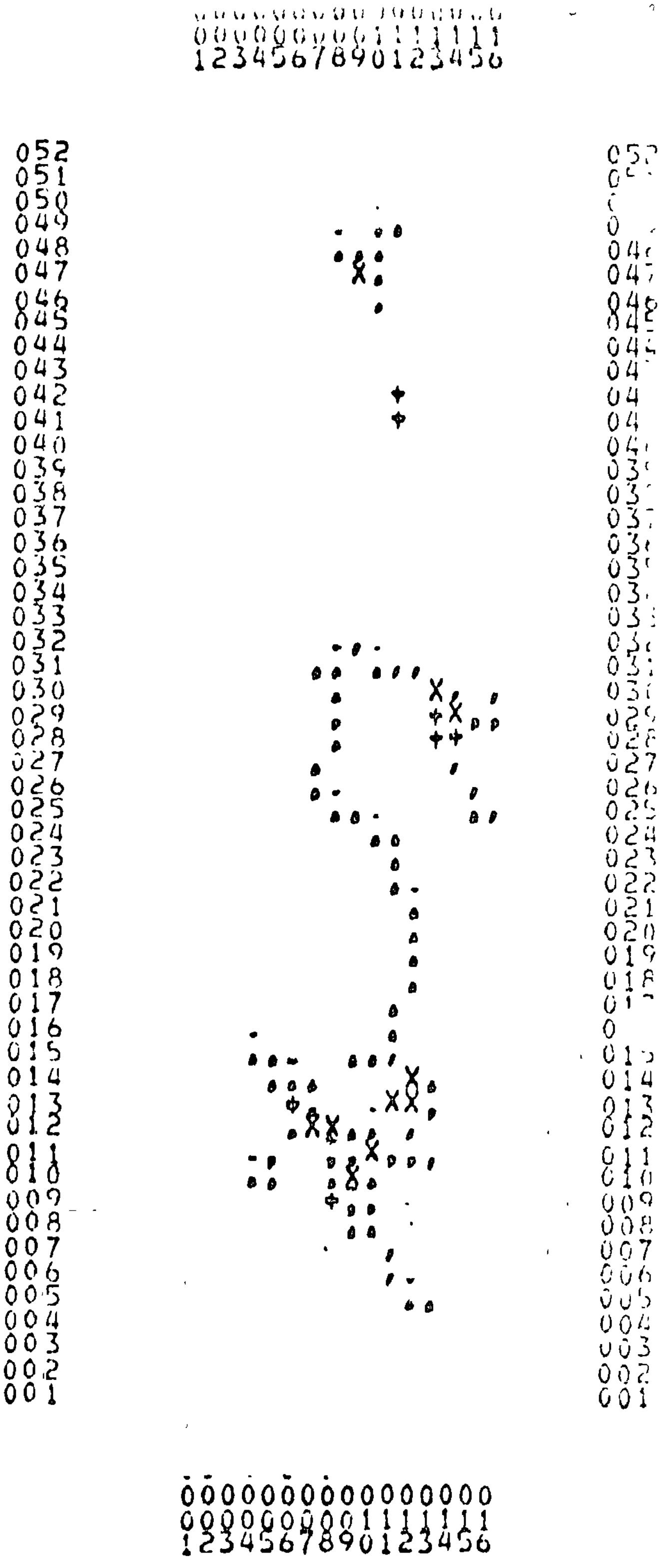
NIVELES	0	1	2	3	4	5
SIMBOLOS	.	'	o	+	X	O
FRECUENCIA	0	73	9	6	0	0

EVALUACION DEL IMPACTO DEL PLAN Ñ 1 EN:

CAMBIO EN LA VISUALIDAD

AREA DE ESTUDIO   VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA



OS USOS DEL SUELO SE AGRUPARON EN LOS GRUPOS:

SO DEL SUELO 1
CTIVIDADES DE ESPARC.
AMINATA
ABALLERIZAS

USO DEL SUELO 11
ESTRUCTURAS
VIS

USO DEL SUELO 111
ESTACIONAMIENTOS
CAMINOS
ESTAC./ TRAILERS

```
NIVELES                1          2            3            4            5
          ======================================================================
          .........  /////////  ,,,,,,,,,  +++++++++  XXXXXXXXX  000000000
          .........  /////////  ,,,,,,,,,  +++++++++  XXXXXXXXX  000000000
SIMBOLOS  ....0....  ////1////  ,,,,2,,,,  ++++3++++  XXXX4XXXX  000050000
          .........  /////////  ,,,,,,,,,  +++++++++  XXXXXXXXX  000000000
          .........  /////////  ,,,,,,,,,  +++++++++  XXXXXXXXX  000000000
          ======================================================================
FRECUENCIA    0          13         58           7          10          0
```

```
     L MAPA
================
   N DEL IMPACTO DEL PLAN # 1 EN: VULNERABILIDAD A FUEGOS
  EL ESTUDIO    VALLE DEL PACIFICO
  ATORIO DE PLANEACION URBANA

   LES USADAS PARA ESTE MAPA
===========================

IO       EL TEXTO DEL MAPA ES            ==========
               LOS USOS DEL SUELO SE AGRUPARON EN LOS GRUPOS:
               USO DEL SUELO 1            USO DEL SUELO 11            USO DEL SUELO 11

               ESTACIONAMIENTOS           ACTIVIDADES DE ESPARC.      ESTAC./ TRAILERS
               CABALLERIZAS               ESTRUCTURAS                 CAMINATA
               CAMINOS                    VIS



ITULO DEL MAPA
================
  ACION DEL IMPACTO DEL PLAN # 1 EN: DESLIZAMIENTOS DE TIERRA
  DE ESTUDIO    VALLE DEL PACIFICO
  ORATORIO DE PLANEACION URBANA

   LES USADAS PARA ESTE MAPA
===========================

IO       EL TEXTO DEL MAPA ES            ==========
               LOS USOS DEL SUELO SE AGRUPARON EN LOS GRUPOS:
               USO DEL SUELO 1            USO DEL SUELO 11            USO DEL SUELO 11

               ACTIVIDADES DE ESPARC.  CABALLERIZAS                   ESTACIONAMIENTOS
               CAMINATA                                               ESTRUCTURAS
                                                                      VIS
                                                                      CAMINOS
                                                                      ESTAC./ TRAILERS



TULO DEL MAPA
================
  UACION DEL IMPACTO DEL PLAN #1 EN: BASURA EN EL MAR
  DE ESTUDIO    VALLE DEL PACIFICO
  ORATORIO DE PLANEACION URBANA

   ILES USADAS PARA ESTE MAPA
===========================

O        EL TEXTO DEL MAPA ES            ==========
               LOS USOS DEL SUELO SE AGRUPARON EN LOS GRUPOS:
               USO DEL SUELO 1            USO DEL SUELO 11            USO DEL SUELO 11

               ACTIVIDADES DE ESPARC.  ESTRUCTURAS                    ESTACIONAMIENTOS
               CAMINATA                                               VIS
               CABALLERIZAS                                           CAMINOS
                                                                      ESTAC./ TRAILER
```

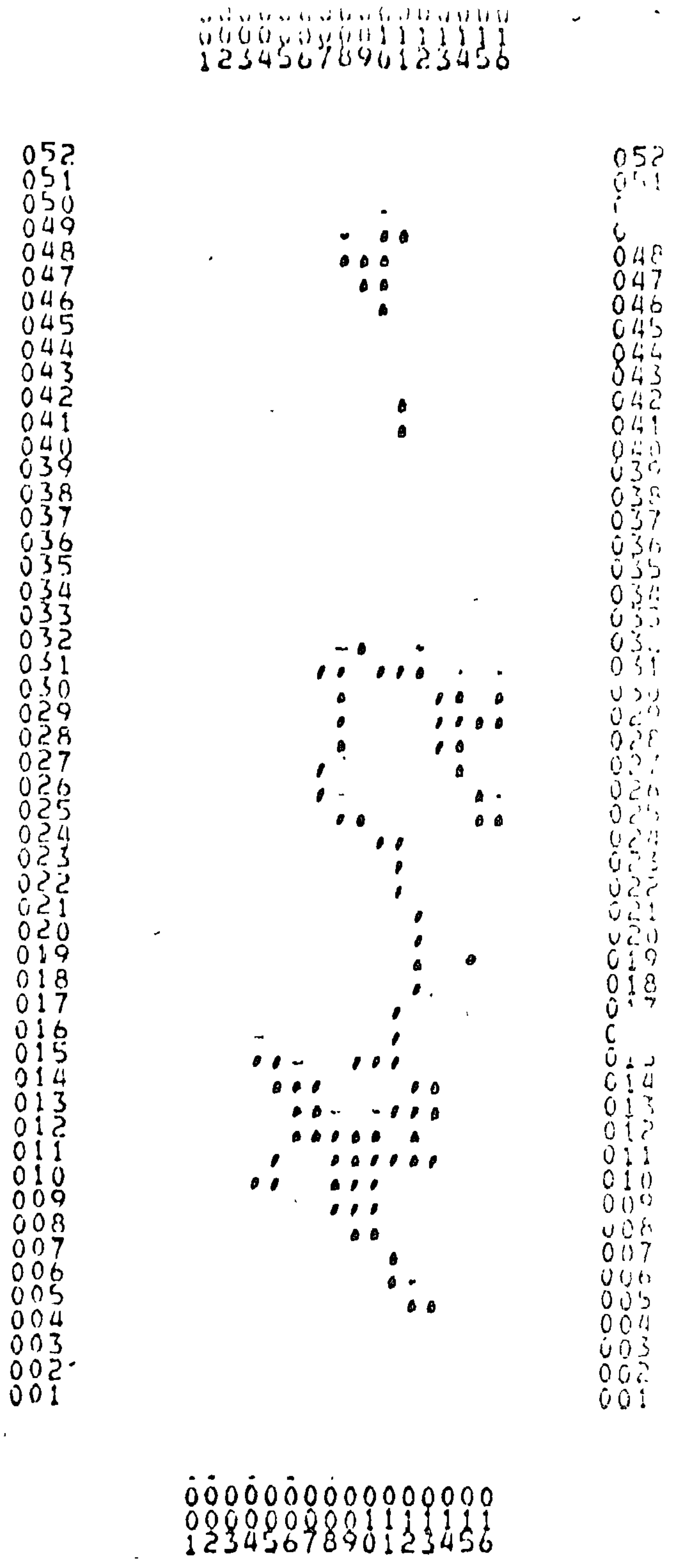
EVALUACION DEL IMPACTO DEL PLAN Ñ 1 EN:

VULNERABILIDAD A FUEGOS

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

LOS USOS DEL SUELO SE AGRUPARON EN LOS GRUPOS:



USO DEL SUELO 1

ESTACIONAMIENTOS
CABALLERIZAS
CAMINOS

USO DEL SUELO 11

ACTIVIDADES DE ESPARC.
ESTRUCTURAS
VIS

USO DEL SUELO 111

ESTAC./TRAILERS
CAMINATA

NIVELES	0	1	2	3	4	5
SIMBOLOS	.	'	•	+	X	0
FRECUENCIA	0	45	43	0	0	0

EVALUACION DEL IMPACTO DEL PLAN 81 EN:

BASURA EN EL MAR

AREA DE ESTUDIO VALLE DEL PACIFICO

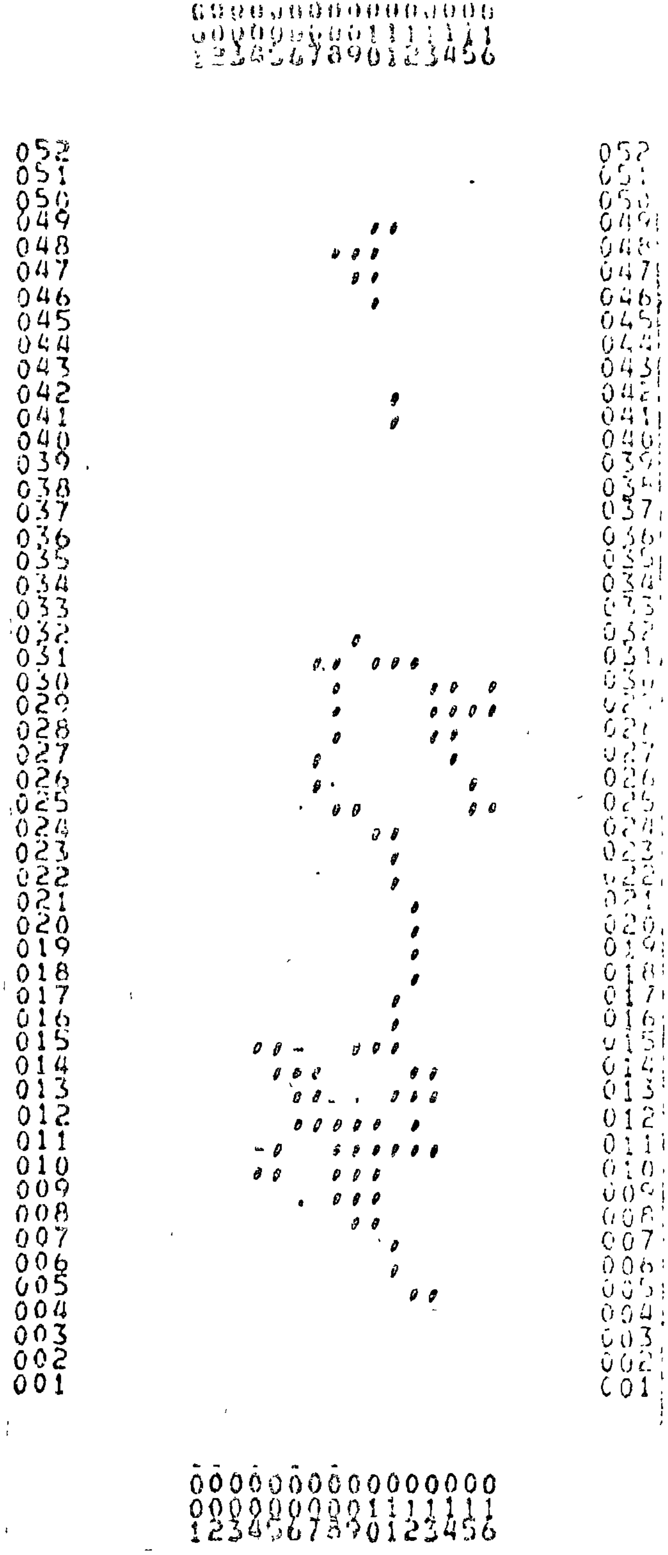
LABORATORIO DE PLANEACION URBANA

LOS USOS DEL SUELO SE AGRUPARON EN LOS GRUPOS:

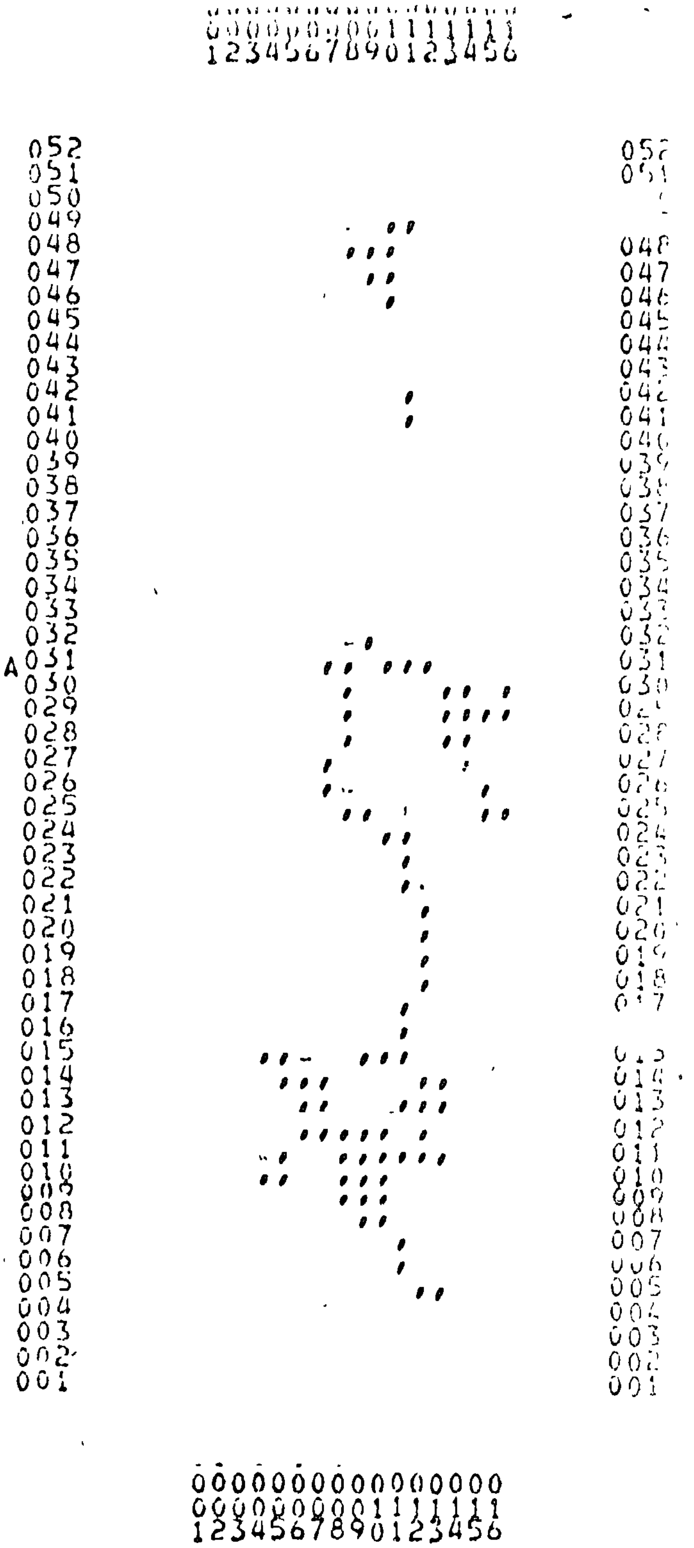
USO DEL SUELO 1
ACTIVIDADES DE ESPARC.
CAMINATA
CABALLERIZAS

USO DEL SUELO 11
ESTRUCTURAS

USO DEL SUELO 111
ESTACIONAMIENTOS
VIS
CAMINOS
ESTAC./ TRAILERS



```
NIVELES           0              1              2              3              4              5
          ======================================================================================
          .........    '''''''''    ooooooooo    +++++++++    XXXXXXXXX    OOOOOOOOO
SIMBOLOS  .........    '''''''''    ooooooooo    +++++++++    XXXXXXXXX    OOOOOOOOO
          ....0....    ''''1''''    oooo2oooo    ++++3++++    XXXXXXXXX    OOOOOOOOO
          .........    '''''''''    ooooooooo    +++++++++    XXXXXXXXX    OOOOOOOOO
          .........    '''''''''    ooooooooo    +++++++++    XXXXXXXXX    OOOOOOOOO
          ======================================================================================
FRECUENCIA        0             88              0              0              0              0
```



```
EVALUACION DEL IMPACTO DEL PLAN Ñ 1 EN:

            DESLIZAMIENTOS DE TIERRA



AREA DE ESTUDIO   VALLE DEL PACIFICO



            LABORATORIO DE PLANEACION URBANA



 USOS DEL SUELO SE AGRUPARON EN LOS GRUPOS:



USO DEL SUELO 1                  USO DEL SUELO 11
CTIVIDADES DE ESPARC.            CABALLERIZAS
CAMINATA


              USO DEL SUELO

              ESTACIONAMIENTOS ,
              ESTRUCTURAS
              VIS
              CAMINOS
              ESTAC./ TRAILERS




NIVELES            0              1              2              3              4              5
          ======================================================================================
          .........   '''''''''   •••••••••   +++++++++   XXXXXXXXX   000000000
SIMBOLOS  ....0....   ''''1''''   ••••2••••   ++++3++++   XXXX4XXXX   000050000
          .........   '''''''''   •••••••••   +++++++++   XXXXXXXXX   000000000
          ======================================================================================
FRECUENCIA         0             88              0              0              0              0
```

EVALUACION DE LA ATRACTIVIDAD DEL PLAN Ñ 2

AREA DE ESTUDIO VALLE DEL PACIFICO

LABORATORIO DE PLANEACION URBANA

LOS USOS DEL SUELO SE AGRUPARON EN LOS GRUPOS:

USO DEL SUELO 1

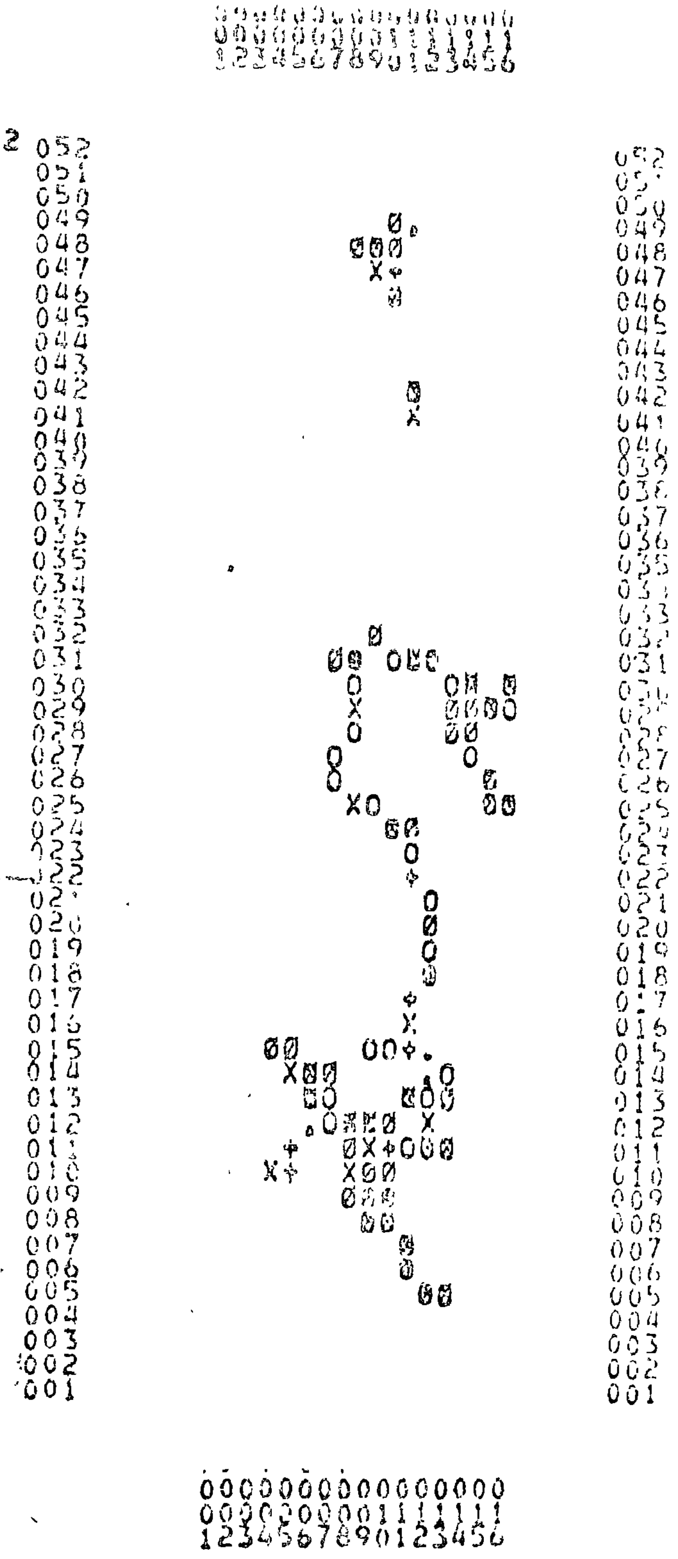
ACTIVIDADES DE ESPARC.
CAMINATA
CABALLERIZAS

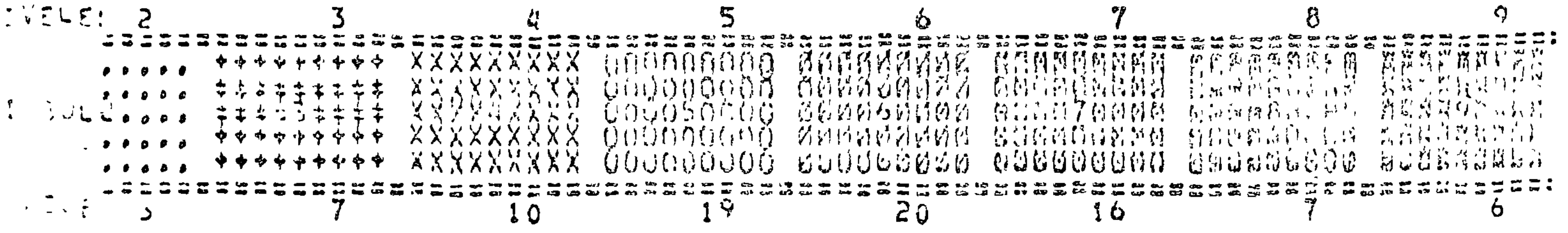
USO DEL SUELO 11

ESTRUCTURAS

USO DEL SUELO 111

ESTACIONAMIENTOS
VIS
CAMINOS
ESTAC./ TRAILERS





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

MODELOS DE SINULACION PARA LA PLANEACION INTREGRAL DEL USO Y DEL TRANSPORTE

TEMA: CLUG
JUEGO DEL USO DEL SUELO.

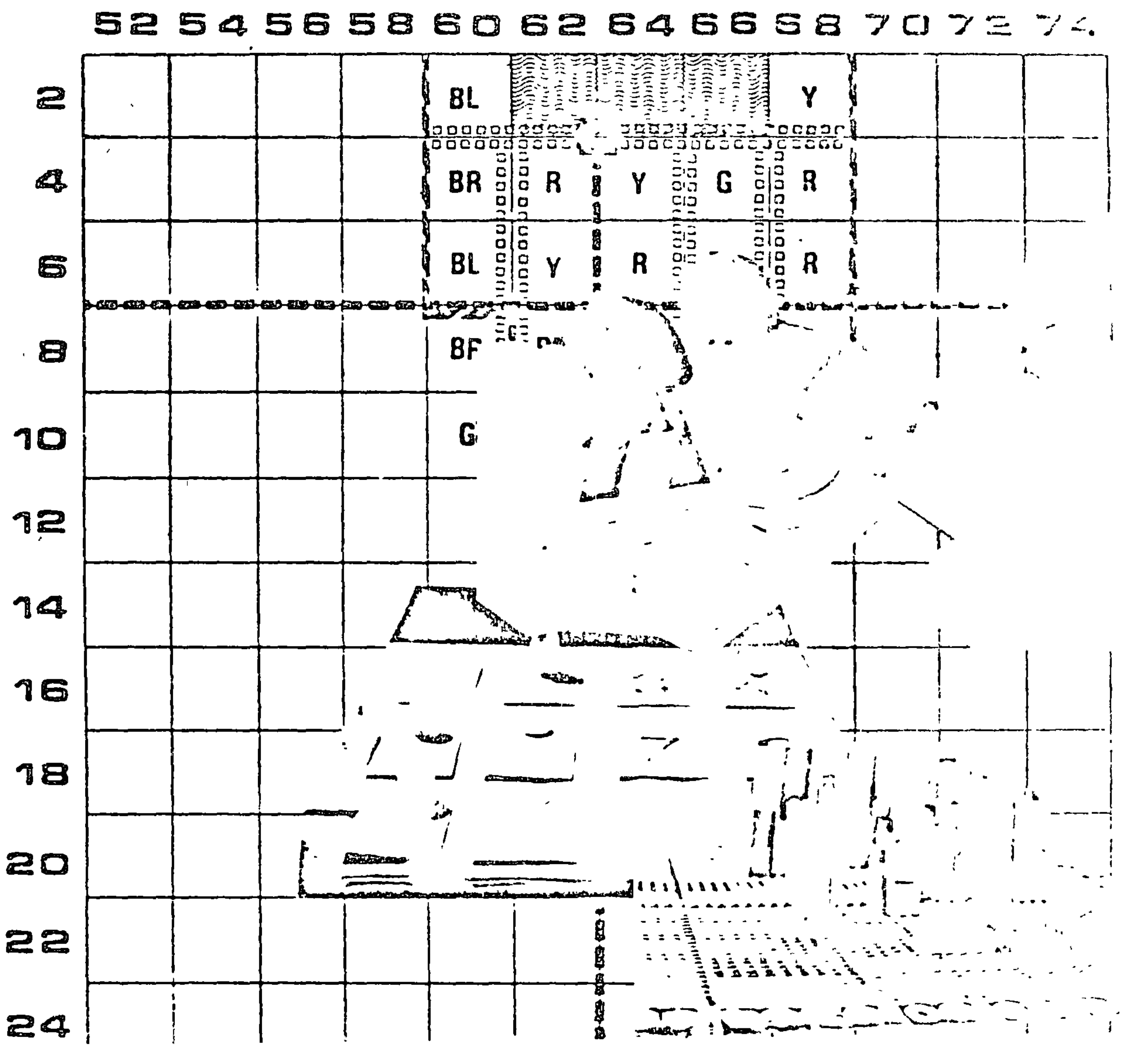
AUTOR: ALLAN FELDT.
COORDINADOR: ARQ. ALEJANDRO VILLANUEVA.

DIVISION DE ESTUDIOS SUPERIORES
FACULTAD DE INGENIERIA UNAM.

SECCION DE PLANEACION

CLUG

JUEGO PARA LA SIMULACION DE SISTEMAS URBANOS.



EL JUEGO BASICO DEL USO DE LA TIERRA DE LA COMUNIDAD

Resumen teórico del CLUG básico

Los teorizantes del urbanismo, dedicados al estudio de la geografía, la sociología y la economía, comprenden la importancia de ciertos principios básicos de la organización de ciudades y zonas; pero la transmisión de estos conocimientos ha sido difícil, debido a los problemas inherentes al empleo de métodos verbales adecuados para describir los procesos dinámicos y fluctuantes, involucrados en todo sistema económico, y sus correspondientes patrones del uso del suelo. A última fecha, muchos teóricos han recurrido al empleo de modelos matemáticos y físicos para lograr una mejor expresión de este proceso básico. Ello ha excluido a mucho urbanista, puesto que no todos los teorizantes dominan las técnicas matemáticas.

Una descripción matemática clara de los fenómenos básicos del crecimiento urbano, dinámico, fue hecha por Brian Berry, * y resume los elementos que el CLUG trata de explicar.

---

* Brian Berry   Fronteras de la investigación en geografía urbana, documento preparado por el Comité en Urbanización del Consejo de Investigación de Ciencias Sociales, 1961. Reproducido con permiso del autor.

ESTRUCTURA ESPACIAL DEL USO DEL TERRENO URBANO

Los estudios más recientes sobre la estructura interna de las ciudades afirman que éstas están sustentadas por actividades básicas, cuya localización se determina exógenamente a la ciudad, a base de ventajas comparativas entre los sistemas económicos, regionales, nacionales e internacionales.

Tales actividades básicas implican normalmente un distrito central de negocios, que es el núcleo, no sólo de la ciudad, sino a la vez, de toda la región tributaria. También se desarrollan varias actividades especializadas, tales como la industria del acero en Chicago, la producción de aviones en Seattle, o empaque de carne en Omaha. Las actividades básicas, aunadas al sistema de transporte, constituyen las características esenciales del patrón urbano.

Este patrón está integrado, en parte, por las viviendas de los trabajadores de actividades básicas y tiene características dinámicas debido al flujo y reflujo diario de quienes acuden a ella a trabajar, así como —más allá de sus límites— al movimiento de bienes y comerciantes que van desde y hacia los sitios de las actividades básicas. Podría haber, por supuesto, algunas motivaciones basadas en la calidad del terreno sobre el que se extiende la ciudad.

Otra caracterización proviene de la orientación de los negocios hacia las actividades básicas y de las actividades terciarias hacia los trabajadores. —

— consumidores — y sus familias. También los viajes de compras originan flujo de clientes. Es entonces cuando aparecen los efectos secundarios: la localización de la vivienda de quienes trabajan en actividades no básicas, trabajadores que acuden adicionalmente a la ciudad, mayor demanda de actividades terciarias, etc., y que forman una cadena, cada vez más compleja, de efectos -- multiplicadores.

Establecer los argumentos simbólicamente:

Definiciones:

A = patrón de actividades básicas locales.
B = sistema de transporte
C = conjunto de sitios urbanos
D = empleados o trabajadores en A
E = patrón residencial de D
F = patrón básico de desplazamiento de los empleados
G = sistema de servicios de negocios
H = sistema de actividades terciarias
I = conjunto de efectos secundarios

Establecido así A,

$$D = f\,(A)$$

$$E = f\,(A,\ D,\ \text{supeditado a } B \text{ y } C)$$

F = f (A, E, supeditado a B)

G = f (A)

H = f (E, D, supeditado a B)

I = f (G, H)

Una especificación precisa de estas relaciones funcionales figura en el trabajo referente a la estructura interna de las ciudades.

CLUG utiliza esas técnicas y postulados estáticos y les da vida en condiciones más reales, dinámicas, con los hechos y complicaciones imprevistas, que se presentan en el crecimiento y evolución urbanos. En operación, CLUG se convierte en un modelo de simulación donde los seres humanos son quienes toman decisiones; esto permite a tales seres, como jugadores, enjuiciar el proceso de crecimiento y decadencia de una ciudad al participar en el proceso. Aunque muy limitado en términos de los elementos del crecimiento urbano que intenta incluir, y aparentemente con una estructura muy simple, el juego genera rápidamente un patrón de desarrollo complejo e imprevisible, el cual es justamente representativo, por lo menos, de algunos de los componentes del crecimiento y sus inter-relaciones, que describió Berry refiriéndose a las ciudades norteamericanas.

El tablero del juego presenta una gran variedad de posibles sitios urbanos (C), divididos entre una matriz de igual área por un sistema coordenado rectangular. Sobrepuesto a sitios de posibles usos del terreno, está el sistema de transportes (B), integrado por caminos principales y secundarios, así como por una terminal, el cual constituye el determinante básico para la localización de las industrias básicas (A).

Estas actividades básicas se hallan representadas por industrias chicas y grandes, que pueden ser " construidas " y operadas por los jugadores. La economía de tal operación, tomando en cuenta el sistema de transporte, hace que las localizaciones tiendan a agruparse tan cerca de la terminal como sea posible. Los trabajadores de estas industrias (D) disponen de unidades residenciales construidas por otros jugadores. El costo del transporte, desde y hacia el trabajo (F), así como el intento de los jugadores de minimizar ese costo dan por resultado un patrón de localizaciones residenciales, que tiende a agruparlas alrededor de las industrias, sujeto esto a la disponibilidad de terrenos, al precio éste, y al sistema de transporte E = f (A, D sujeto a B y C). El número de unidades habitacionales depende del número y tipo de industrias existentes D = f (A).

El sistema de instalaciones para negocios, está representado por un tipo de edificio llamado oficina (G). Las ventajas de este tipo de construcción dependen del número potencial de inquilinos constituido, principalmente por las industrias G = f (A') . Las actividades terciarias son los centros comer ciales (H) que sirven a las unidades habitacionales. El número y localización de estas tiendas depende del volumen y ubicación de las unidades residenciales previamente establecidas, y de acuerdo con las restricciones y facilidades que proporciona un sistema de transporte disponible H = f (E, D sujeto a B).

Al igual que la localización de los negocios y de las unidades de servicios terciarios, surgen nuevas oportunidades de localización, o sea lo que Berry (I) ha llamado efectos adicionales. Estas comprenden la construcción de unidades residenciales para los elementos que trabajan en las tiendas y los negocios que habrán de ubicarse lo más cerca posible de dichas fuentes de trabajo - I = f (G, H). Estas unidades habitacionales crean, a su vez, servicios terciarios adicionales en la misma vuelta o en las subsecuentes. Estos efectos secundarios continúan hasta que, en un momento dado, el sistema llega a estar relativamente bien balanceado.

La evolución de la ciudad en CLUG sigue estas vueltas básicas, además de la construcción de industrias adicionales, dando lugar así a otros ciclos sucesivos de crecimiento y desarrollo. En el curso del juego, el proceso de decadencia de las estructuras más viejas, aunada a posibles dislocaciones debidas a desastres naturales o económicos, así como a cambios en el sistema de transporte, conducen a continuas modificaciones y ajustes en ese proceso.

SINTESIS DEL PROCEDIMIENTO DEL CLUG BASICO

El juego CLUG está específicamente diseñado para proporcionar un conocimiento básico de los factores subyacentes que afectan el crecimiento de una zona urbana. El juego hace énfasis en ciertos aspectos de la economía urbana: la relación que existe entre las industrias básicas y el empleo, la habitación, y los costos de transporte; el desarrollo de servicios comerciales; el financiamiento e instalación de servicios comerciales; el financiamiento e instalación de servicios municipales; y en la ubicación e interdependencia de todas estas actividades en una zona urbana.

Como participante de CLUG, el jugador tiene oportunidad de invertir en tierra, construir edificios de varios tipos, y buscar los medios adecuados para relacionar las inversiones dentro de la economía local y hacer producir su inversión (cuando la inversión ha sido bien planeada). Si bien es de esperarse la competencia entre los compañeros por un lugar en el sistema, existe, desde luego un límite en dicha competencia. El bienestar individual depende, en gran parte, de todos los jugadores. Aunque adversarios, los jugadores aprenderán a cooperar unos con otros, para lograr ciertos propósitos, en bien del desarrollo y de la comunidad.

Un instructor dirige el juego explicando e interpretando las reglas. Representa también a las economías externas a la comunidad local, en la compra de productos manufacturados y la venta de servicios comerciales a los jugadores que no son capaces de obtenerlos a un precio adecuado dentro de la comunidad. Dicho intructor anuncia también aquellas catástrofes inesperadas que pueden dañar la estabilidad de la comunidad a través de la pérdida de sus inversiones. No interviene directamente en el juego, excepto para ordenar eventos y ejecutar reglas. Dentro de los límites de la capacidad humana, opera objetiva y racionalmente en pro de los jugadores y de la comunidad. Como tiene más habilidad que los jugadores novatos, el instructor puede ocasionalmente aconsejar sobre inversiones y toma de decisiones en general. Su consejo deberá ser cuidadosamente considerado, pero aún el más hábil jugador no se halla capacitado para predecir los eventos que pueden ocurrir en el juego. El instructor no es omnipotente, él no controla la dirección o evolución de la ciudad. Esta se halla totalmente en manos de los jugadores, y restringido por las reglas que se expondrán después.

De vez en cuando, suele presentarse en el juego un evento de especial interés y que tiene aplicaciones importantes en algún fenómeno similar de la vida real, o en teoría. En tal caso, el instructor puede interrumpir el juego para explicar los detalles de lo que ha ocurrido y mostrar su relación con los eventos de la vida real. Si bien tales interrupciones pueden retrasar el proceso de crecimiento en la comunidad, constituyen buenas oportunidades para aprender más acerca de la manera en que operan los sistemas urbanos, tanto en el juego como en la vida real.

El instructor es usualmente auxiliado por un contador, quien lleva el registro de las propiedades, estima los valores de las construcciones y terrenos, la edad de las construcciones y el estado financiero de la comunidad, incluyendo impuestos vencidos. Este contador representa al tesorero público, y debe de estar abierto a la inspección por parte de los jugadores durante el juego. Para entender la medida y significación de ese registro público se requiere, sin embargo, cierta experiencia.

Las reglas que siguen proporcionan el marco básico del juego, en su forma más simple. No obstante que al principio parece difícil comprenderlo, casi todos los jugadores aprenden rápidamente, en el curso de las dos horas iniciales del juego. La misma serie de pasos ocurre en cada vuelta. Después de algunas vueltas, se familiarízan con los pasos y los números y cantidades empleadas resultan pronto conocidas a todos los jugadores. Como no se ha adquirido - práctica en la operación del juego, solamente una o dos vueltas se realizan usualmente durante las primeras dos horas del juego. Después de este período inicial, el juego se desarrolla a razón de 30 minutos, o menos, por vuelta.

La velocidad con que puede completarse un ciclo, está determinada en gran parte, por la eficiencia de los jugaodres para tomar decisiones, tanto en los - - equipos individuales como en la comunidad en su conjunto. Como el juego contiene numerosas decisiones basadas sobre mucha información, los jugadores podrán elegir la información útil y así alcanzar práctica como ejecutivos urbanos. Deberán aprender a distinguir las decisiones importantes de las triviales y evaluar la información recomendable para tomar rápidamente decisiones, siempre anticipándose a las decisiones complementarias, que deberán hacerse por otros jugadores.

Una racionalización completa es casi imposible, aún en este simple juego, y los jugadores deben aprender a vivir en un mundo donde las presiones del tiempo y la falta de información los forzan a establecer una conducta sólo parcialmente, racional o satisfactoria.

CLUG se juega más efectivamente con 3 o 5 equipos de una a tres personas cada uno. Es posible usar más equipos o más personas por equipo, pero los procedimientos contables se vuelven muy complejos. Además, es difícil que un equipo llegue a una decisión cuando cuenta con más de 3 integrantes, a menos que esté muy bien organizado internamente. Una determinación clara de las funciones dentro de cada equipo ayuda a tomar decisiones rápida y fácilmente; los jugadores deben ser estimulados a elegir un líder, que será el portavoz y voto representativo de cada equipo.

El juego básico de CLUG, descrito en las siguientes reglas, es lo suficientemente complejo como para satisfacer a la mayoría de los jugadores en 5 ó 10 vueltas de juego. Es posible extenderse más aún, pero el aprendizaje adicional disminuye mucho a partir de la décima-quinta vuelta, a menos que se introduzcan reglas o modificaciones adicionales.

El conjuto de experimentos diseñados de acuerdo con las reglas básicas permite explorar, sistemáticamente, los efectos de cierta clase de fenómenos urbanos. Una vez aprendido el juego básico, la mayor parte de los jugadores está en posición para realizar sus vueltas, con propósitos experimentales, escogiendo a uno de los jugadores como instructor, a otro como contador y haciendo los demás los papeles que requieren los experimentos en particular. Una - lectura rápida a la sección correspondiente del Manual del Instructor — que puede ser facilitada por el instructor — aclara los pocos detalles aquí faltantes que pudieran surgir durante el juego.

Este juego intenta contestar las dudas particulares y problemas del crecimiento urbano. Cada experimento va acompañado de lecturas cuidadosamente seleccionadas para relacionar el fenómeno del juego con la vida real. La -- combinación de dos o más experimentos es posible, si bien en este caso, la estructura total del juego puede resultar demasiado compleja para ser debidamente comprendida. Se recomienda especial cuidado al intentar desarrollar combinaciones de experimentos, empleando el tiempo adecuado para probarlas y revisarlas con - todo, diseñar nuevas modificaciones al juego puede ser muy instructivo, y los - esfuerzos en este sentido deberán ser estimulados. Con todo, modificar un sistema todavía tan simple como el CLUG resulta algo difícil. Cambios que parecen muy simples en las reglas, a menudo se producen profundos y desorganizados efectos en la operación del juego. Descubrir esos efectos y aprenderse sus implicaciones constituye una experiencia valiosa. La simple lectura de las --- reglas siguientes no permite entender completamente su significa- - - - - -

do, y la trascendencia de cada uno de los detalles indicados. Deben leerse inicialmente para tener un conocimiento general del flujo del juego y de la clase de factores que lo afectan. Los números y cantidades dadas en las reglas deberán siempre hallarse siempre disponibles como referencias durante el juego. De esta manera, sólo se requiere tener un conocimiento general de las tablas empleadas y de su contenido general.

REGLAS DETALLADAS DEL CLUG BASICO

La simple lectura de estas reglas es suficiente para proporcionar al jugador una base confiable para empezar a jugar. El significado y finalidades específicas de muchas reglas irán aclarándose a medida que el juego progrese. - Como cualquier juego, más que leer acerca de él, la mejor manera de entender el CLUG es jugarlo. Aunque aparentemente difíciles, las reglas son repetitivas, por lo que su comprensión es sumamente fácil.

Dinero CLUG

Al principio del juego básico, cada equipo dispone de $ 100 000.00 en - efectivo. Este dinero puede emplearse en comprar tierra, eregir construcciones, cubrir impuestos y hacer los pagos necesarios a otros jugadores. En cada vuelta entra al juego dinero adicional, cuando el instructor hace los pagos a los dueños de las industrias. Una parte de este dinero circula hacia otros jugadores en forma de pago de empleos y pagos a las tiendas y oficinas por servicios diversos. A su vez, sale dinero del juego cuando los jugadores pagan servicios por transportación, compran bienes o servicios en economías externas a la ciudad, pagan servicios municipales a través de impuestos y hacen pagos para la construcción o renovación de edificios. En el juego básico no está permitido que el instructor realice préstamos, si bien la comunidad, conjuntamente puede contraer deudas hasta por el límite permitido. Los préstamos entre equipos son válidos siempre que sus términos no interrumpan el juego. Los jugadores pueden esperar usualmente un diez por ciento de ganancia en una buena inversión en la ciudad, aunque, de hecho, logran ganancias más altas o más bajas de acuerdo con las inversiones y administraciones, buenas o desacertadas, que realicen.

Tipos de uso del suelo

En el CLUG existen tres tipos básicos de uso del suelo: industrial, comercial y residencial. Existen niveles de densidad industrial: gran industria - (GI), y pequeña industria (PI). En las etapas de construcción de cualquier - - vuelta, una industria pequeña puede transformarse en grande pagando la diferencia de su costo inicial de construcción, y cambiando el tipo de sus edificios en el tablero. La designación de las edificaciones de estos tipos de uso del suelo, así como sus costos iniciales de construcción aparecen ilustrados en la Figura 1.

En terrenos comerciales se construyen 3 tipos básicos de edificios: - - tienda local (TL), grandes almacenes (GA) y oficina (O), que también aparecen ilustrados en la Figura 1. Una tienda local proporciona bienes y servicios tales como abarrotes, productos farmacéuticos, gasolina, artículos domésticos, y auxilios médicos elementales, etc. Cada unidad residencial del tablero deberá comprar una cantidad determinada de bienes en cada tienda local y en cada - vuelta, o adquirirlas del instructor. Los grandes almacenes ofrecen un rango -

más especializado de bienes y servicios, de menos frecuente adquisición, como son joyas, muebles, automóviles, servicios médicos especializados, etc. Cada unidad residencial del tablero deberá adquirir una cantidad normal de estos bienes en los grandes almacenes o del instructor, en cada vuelta.

Finalmente, las oficinas proporcionan una gran variedad de servicios - contables y administrativos a las tiendas e industrias. Todas las industrias y - las tiendas deberán pagar una cantidad normal de estos servicios en cada vuelta a las oficinas construidas o al instructor. No se permiten cambios de un tipo de uso del suelo comercial a otro, a menos que el edificio original sea completamente demolido y se construya un nuevo edificio en su lugar.

El suelo residencial se emplea para unidades habitacionales, las que pueden pertenecer a 4 niveles distintos de densidad: R1, R2, R3 o R4, (ilustrados en la Figura 1). Cada unidad habitacional incluye una unidad para los empleados y sus familias, los cuales deberán hacer adquisiciones normales en las tiendas locales y centrales, en cada vuelta. A una unidad residencial de alta densidad - (R4) corresponden 4 unidades para empleados, y 4 grupos potenciales de clientes de las tiendas locales y centrales. La densidad del terreno residencial puede ser incrementado en las etapas de construcción de cada vuelta, pagando la diferencia de costo entre los dos tipos de edificios, y cambiando, naturalmente, - las formas que lo representan en el tablero. El costo de construcción para cada unidad residencial es algo mayor en las de alta densidad, que en las de baja, por lo exclusivo del lugar, y el costo de los servicios públicos. Esto es resultado en parte, del hecho de que en el juego básico se considera que todos los residentes pertenecen a un mismo nivel socio-económico y viven en un tipo de habitaciones.

Empleo

Todas las industrias y comercios deben emplear unidades residenciales para poder trabajar. Una industria grande, trabajando a plena capacidad, puede emplear hasta 4 unidades residenciales; una industria pequeña emplea hasta dos unidades residenciales; las tiendas locales y centrales y las oficinas emplean, cada una de ellas, una unidad residencial, cuanto están en operación. Los diferentes niveles de salarios quedan establecidos en el juego básico. El dueño de una unidad residencial empleada por una industria o comercio recibe un pago de $ 6 000.00 por cada grupo de empleados, en la etapa de pago de empleos de cada vuelta. Este dinero procede del equipo al que pertenece la fábrica o tienda. Los contratos de empleo son negociados entre los dos equipos —los dueños de los empleos y los de los empleados—, deben ponerse marcas codificadas en ambos edificios. Todos los contratos de empleo afectan a los edificios con marcas y su efecto dura hasta la siguiente vuelta divisible entre cinco (cinco, diez, quince, etc.). En ese tiempo, en la etapa de alquileres, pueden negociarse nuevos contratos, de común acuerdo entre patrones y empleados.

Comercio

Como ya fue dicho, el suelo empleado para usos comerciales ofrece un determinado volumen de bienes y servicios a sus clientes potenciales. Los usuarios llamados a comprar estos bienes y servicios prefieren a aquellos vendedores que ofrecen mejores precios y tienen localización más ventajosa.

Una vez que cada edificio residencial ha convenido en comprar en determinada tienda local, gran almacén u oficina, en particular, los convenios y los precios regirán hasta la siguiente vuelta divisible entre cinco.

Los precios de los servicios y bienes normales los fija el dueño de cada tienda u oficina, quien se ve restringido por tres factores básicos. El primero es que esos bienes y servicios estén al alcance de los jugadores de economías - externas, por ejemplo el instructor. El instructor vende también sus bienes y servicios a un precio fijo, a cualquier jugador que los solicite: $ 2 000 para la canasta de la tienda local; $ 1 000 para la canasta de los grandes almacenes; y $ 4 000, $ 2 000 o $ 1 000 para las oficinas, dependiendo éstos de si el comprador es una industria grande, pequeña o tienda, respectivamente. La segunda - restricción en el costo de transportación para la clientela de una tienda, en cada vuelta. El instructor ofrece sus bienes y servicios a precios que incluyen la entrega en la casa del cliente; así los costos de éste no se aumentan con los relativos a transportación. Todos los dueños de tiendas u oficinas tienen, sin embargo, que considerar el costo de transportación de sus clientes potenciales, -

viajando desde y hacia la tienda, para establecer sus propios precios. Así, su precio es siempre menor que el de las economías externas, para compensar esos costos de transportación. La tercera restricción a los precios consiste en que solamente puede establecerse un precio en cada tienda local, almacen u oficina. Este precio, deberá aplicarse a todos los clientes por igual, y deberá sostenerse hasta la siguiente vuelta divisible entre cinco. Así pues, el comerciante, para la fijación de precios deberá considerar tanto el desarrollo presente como el futuro, y su probable localización con respecto a las unidades comerciales. Los valores usados en las tres secciones precedentes, están resumidas en la Tabla 1.

El Tablero

El tablero CLUG mide 36 x 36 cm y está dividido en 144 manzanas, cada una de las cuales representa un predio. En CLUG básico, las líneas delgadas representan los caminos secundarios mientras que las líneas gruesas —dobles- representan los caminos principales. El sistema de coordenadas está diseñado de tal manera que los números pares designan las columnas o filas de manzanas, o de cuadras; y los números impares que intervienen representan las líneas horizontales y verticales que separan a las manzanas. Así, una pareja de números pares, tales como 8-66 indica un predio, mientras que una pareja de números impares tales como 7-67, indica una esquina particular de un predio. Similarmente, una combinación de un número impar y otro par, para designar coordenadas tales como 9-66, indica un segmento a lo largo de la orilla de un predio determinado. Aunque puede introducirse cualquier configuración topográfica en el tablero, la Figura 2 proporciona un modelo usual, que puede ser empleado en la mayoría de los juegos básicos. Consiste en un camino principal y horizontalmente a lo largo de la línea 39, desde la 5 hasta la 25, y verticalmente a lo largo de la línea 9, desde la 27 hasta la 51. Están indicados algunos - - otros perímetros importantes en el juego básico: un lago, o bahía, aparece en las manzanas 2-38, 2-40 y 4-38, 4-40. Incluye también esta Figura una terminal marítima en la 3-38 y una central de servicios públicos en la 3-40. La terminal proporciona servicios de embarque a las economías externas y es el punto desde el cual todos los productos industriales deberán ser embarcados para su venta. La central de servicios públicos es el punto hacia el que convergen todas las líneas de servicios. Esta planta proporciona un gran número de servicios, que utilizan los ductos y líneas de servicios construidas durante el juego. Además de los servicios de agua y drenaje, con la central y sus líneas, se presta un gran número de servicios municipales, como son los de policía y bomberos, electricidad y calefacción, recolección de basura, bibliotecas, escuelas, etc. Las líneas de servicios se colocan a lo largo de las calles, por mayoría de votos de los equipos. Se considera que un predio dispone de servicios y está en aptitud de desarrollarse, cuando a su lado corre una línea de servicios.

Transportación

Un factor importante en muchas decisiones del juego, es el costo de - - transporte requerido por ciertos usos del suelo; este costo es factor determi--

nante en la localización de las construcciones. Los costos de transportación se liquidan mediante pagos hechos por los jugadores al instructor. Son cubiertos por las industrias, grandes y pequeñas, al adquirir productos que llegan por mar a la terminal; por los residentes, al dirigirse a su trabajo, y a efectuar compras en las tiendas; así como por las industrias y las tiendas al transportar el material desde, y hacia, las oficinas. Su costo se incrementa con la distancia, y es menor cuando el movimiento se efectúa a lo largo de un camino principal que cuando se requiere emplear caminos secundarios. Para el juego básico, la relación de eficiencia entre un camino principal y otro secundario es de 1 a 2. Los costos se incrementan en razón del número de cuadras recorridas, ya sea en un camino principal o secundario (la unidad de distancia es una cuadra). Todos los movimientos deberán verificarse en línea recta, a lo largo de las calles. Estas se cuentan desde la esquina más próxima de la manzana de origen, a la más cercana de la manzana a la que se dirige. La distancia viajada entre 2 predios contiguos es cero, y el costo de transportación es cero también.

El costo unitario de un viaje depende de la frecuencia y volumen de cada clase particular del mismo. El costo es muy alto tratándose de embarques de productos manufacturados, y relativamente bajo en los viajes de compra y trabajo. Este factor de peso para una clase particular de transporte, se denomina "peso asociado", y aparece detallado en la Tabla 2. El peso asociado se agrega al número de manzanas recorridas y se multiplica por 1, o por 2, ya sea que el transporte se realice por un camino principal o secundario, respectivamente. Así, para trasladarse al trabajo, desde una residencia, cuando el sitio de empleo involucra dos unidades de caminos secundarios y dos de caminos principales, se calcula como sigue: dado un peso asociado de 300 para traslado desde la residencia al trabajo, el costo deberá ser 2 veces 300 para cada unidad de camino secundario recorrido; más una vez 300 para cada unidad de camino recorrido; y 1 800 por cada unidad residencial en dicha localización. El costo de transportación resultante lo pagan al instructor, en cada vuelta los equipos que hacen uso del terreno. El costo sólo es reducible disminuyendo la distancia de recorrido o, en algunos casos, disponiendo de mejores medios de transporte.

Impuestos y finanzas de la comunidad

Las erogaciones que ocasiona la comunidad, se pagan en forma de impuestos, en cada vuelta, y corresponden al número de líneas de servicio construidas y de unidades residenciales que existen en el tablero. El costo inicial de construcción de cada segmento de línea de servicios es de 2 000, y los costos de operación y mantenimiento posteriores a su construcción cuestan a la comunidad mil por vuelta. Los equipos jugadores deciden, por mayoría de votos, la construcción y localización de las líneas de servicios. No puede construirse ningún edificio en una manzana que no tenga líneas de servicio a lo largo de uno de sus lados. Los costos se computan por unidades. Toda línea de servicio deberá, por supuesto, unirse al sistema conectado finalmente con la planta de servicios. Además del costo de los servicios públicos, cada unidad residencial paga 1 000 por servicios sociales en cada vuelta. Este costo, como el de la cons trucción de los servicios y los costos de mantenimiento de los mismos, deberán ser pagados por la comunidad independientemente de los impuestos.

En el juego básico, se paga el impuesto predial sobre terrenos o edificios. La tasa de impuesto se fija por mayoría de votos de los equipos. Esta tasa, multiplicada por el valor catastral de las propiedades de un equipo, determina el monto que deberá ser pagado en impuestos en la vuelta siguiente. En las vueltas sucesivas, la tasa puede bajar o elevarse de acuerdo con las necesidades financieras de la comunidad y sus planes de crecimiento.

La comunidad puede tener un superávit o déficit financiero en cualquier vuelta, y este superávit o déficit se acumula en las vueltas sucesivas. Sin embargo, el déficit no podrá exceder el 10% del valor catastral de las propiedades de la comunidad, en cualquier momento. La comunidad pagará una tasa de interés básico del 10% sobre su déficit. El superávit acumulado no ocasiona interés. Si se excede el límite del 10% no se autorizará la construcción de servicios hasta que el déficit sea previamente pagado por la comunidad, es decir, hasta que la deuda de la comunidad corresponda a cero. El control contable de las propiedades, impuestos, superávit o déficit, deudas extraordinarias corresponderá al contador de la comunidad. La principal labor de los jugadores será establecer las tasas tributarias, pagar sus impuestos y cuidar que no se acumule ningún déficit que perjudique la capacidad financiera y crediticia de la comunidad.

Depreciación

El valor inicial de las construcciones se deprecia a razón de un 5% por vuelta. Cada 5 vueltas, los dueños de los edificios deberán ser informados sobre la vida útil y estado de depreciación de los mismos, y se les preguntará si desean,o no, renovar total o parcialmente sus construcciones. La renovación se realiza mediante un pago al instructor del 5% de los costos de depreciación acumulados en las vueltas. Esto renueva al edificio y prolonga efectivamente su vida útil. Los pagos de renovación podrán ser hechos cada 5% que se deprecie — la depreciación equivalente a una vuelta —. Así el edificio puede regresar a su edad cero, o ser parcialmente renovado por el equivalente a una vuelta o dos, dependiendo del valor y del uso que se dé a la estructura y al terreno. Después de haber tomado la decisión de renovar, o no, una construcción particular, cada equipo deberá tirar un par de dados para cada edificio, las caras que hacen perder el uso del edificio son proporcionales al estado de depreciación. Cuando un edificio esté "perdido", se le eliminará efectivamente, y no podrá ser usado durante 5 vueltas. Después de esas cinco vueltas, el edificio estará de nuevo apto para ser renovado pudiendo ser jugado con los dados nuevamente. Sin embargo, el edificio será entonces 5 vueltas más viejo, y sus probabilidades de ser eliminado son mayores. En este punto el edificio podría ser renovado para reducir la probabilidad de perderlo una vez más.

Toda construcción tiene, cuando menos, un 5% de probabilidades de perderse, no obstante que sea completamente renovado en cada vuelta. Esta proba

11. PAGO DE IMPUESTOS.- El contador informa a cada equipo de los impuestos que debe la comunidad, y el instructor los cobra . Asimismo informa a la comunidad del monto total recabado por impuestos, del monto de los gastos originados, del interés pagado sobre deudas, y del déficit o superávit de la comunidad. Los jugadores discuten y votan sobre la tasa de impuestos que regirá en la siguiente vuelta.

Preparación para el juego

Anexo a este manual vienen, por duplicado de las tablas de información para el jugador, y que contienen toda la información básica necesaria para el juego CLUG. Vienen en forma de hojas desprendibles, de manera que puedan emplearse fácilmente durante el juego. Para la mayoría de los jugadores es suficiente una lectura de las reglas ya enunciadas junto con el contenido de las tablas, es suficiente para contestar la mayoría de las dudas que surgen durante el juego. Se proporciona en duplicado de esta tabla, para casos de pérdida o mutilación de las mismas.

También se adjuntan algunos cupones para que el jugador registre en ellos que sus solicitudes o licitaciones sobre terrenos y los contratos de compra con los otros jugadores. Al llenarlos, el vendedor de un terreno indica las coordenadas numeradas del predio solicitado, junto con el color del equipo, y el número de la vuelta. El precio ofrecido deberá anotarse redondeándolo en centenas, puesto que ésta es la más baja denominación del dinero de CLUG.

Un contrato de venta es un acuerdo escrito entre el dueño de una propiedad residencial y el dueño de una tienda local o gran almacén, en que aquél se obliga a hacer sus compras en dicha tienda, a un precio previamente acordado, hasta la siguiente vuelta divisible entre cinco. Al llenar estas formas deben indicarse las coordenadas de la unidad residencial que corresponden al cliente y el color del equipo dueño de esa unidad residencial. También deben quedar indicadas las coordenadas de la tienda, el precio acordado (usualmente se reporta un precio por cada unidad residencial) y el color del equipo dueño de la tienda. Este acuerdo de venta es usualmente llenado por el comprador y queda en poder del dueño de la tienda, de manera que éste pueda llevar un registro de lo que cada equipo le debe durante cada vuelta.

También se incluyen 2 copias del registro de impuestos CLUG, y de la forma del estado financiero de la comunidad. Una de cada una de estas formas quedará en poder del contador, para cada experimento, o corrida del juego. Este necesitará también una copia del registro general de las propiedades que contienen el Manual del Instructor.

Después de jugar 5 ó 10 vueltas del juego básico CLUG, la mayoría de los jugadores tiene ya una idea clara de la secuencia del juego y está capacitado para jugar por sí mismo. Una cuidadosa revisión del Manual del Instructor puede ayudar en esta etapa de entrenamiento. Dependiendo esto del interés de los jugadores puede considerarse necesario, ya sea repetir toda la carrera o bien sólo de partir de la 5a. o 6a. vuelta. En este punto, es a menudo conveniente que los equipos eli jan entre sí un gerente de juego quien hará el papel del instructor. El gerente, a su vez, deberá elegir un contador que registra las operaciones sobre propiedades, los gravámenes y los impuestos. El equipo al que pertenecen el gerente y el contador elegidos recibe, usualmente, 6 mil por vuelta de los fondos de la ciudad. El gerente puede ser reelegido cada dos vueltas con facultades para nombrar un nuevo contador.

Otras operaciones, más elaboradas y reales del gobierno de la ciudad pueden introducirse en el juego CLUG básico, particularmente si se quiere poner en práctica otras formas de planeación o zonificación. En este caso, el experimento 1 proporciona los mecanismos más apropiados.

Contador: estudiante, o profesor adjunto, que ayuda a conducir el juego, llevando un registro de la propiedad de los terrenos, la construcción de edificios, el valor catastral, el estado financiero de la comunidad, etc.

Edad de las construcciones y edificios: estimación sobre la vida útil de los edificios basados en la diferencia entre el número de la vuelta que se juega y el de aquélla en que fue construido el edificio.

Valor catastral del terreno y edificios: estimación del valor comercial actual del terreno y edificios, el del terreno se fija con base en el precio de compra del propio inmueble y de las propiedades adyacentes. El valor de un edificio se establece tomando como base su costo inicial de construcción, menos el 5% de depreciación por cada vuelta.

Peso asociado: magnitud que refleja la interacción entre dos usos cualquiera del terreno, en una sola vuelta. Refleja también el volumen y frecuencia de los movimientos que se efectuan, desde y hacia el terreno, en una unidad de tiempo.

Grandes almacenes: representan en el juego CLUG, el uso del terreno destinado a establecimientos comerciales altamente especializados, y que generalmente se instalan en los grandes centros de población, ya que requieren de un volumen de clientes mucho mayor que el de una tienda local.

Déficit o superávit de la comunidad: estado financiero de la comunidad al final de cada vuelta. Existe superávit cuando el monto de los impuestos colectados para una vuelta, excede del correspondiente de los gastos de la comunidad en dicha vuelta. El déficit resulta cuando los gastos son mayores que los impuestos cobrados. Tanto el superávit como el déficit, si existen, se acumulan al déficit o superávit proveniente de las vueltas anteriores, para actualizarlo.

<u>Límite del pasivo</u>: el pasivo máximo en que puede incurrir una comunidad en términos de su déficit acumulado. Usualmente, este límite lo fijan las autoridades de la ciudad, como un porcentaje del valor catastral de todas las propiedades durante los últimos años. En CLUG, dicho límite es el 10% del valor catastral correspondiente a la vuelta anterior, de los inmuebles de toda la comunidad.

<u>Contratación de empleos</u>: corresponde al período en que los propietarios de industrias o tiendas solicitan empleados, y en que, a la vez, los jugadores que poseen residencias, buscan empleo. Ambos sectores celebran acuerdos sobre la ocupación del personal, los que regirán hasta la próxima vuelta divisible entre cinco . Sin embargo, en ciertos experimentos, este término es menor.

<u>Mercados externos</u>: sitios que quedan fuera de la comunidad CLUG y del tablero y que adquieren productos manufacturados a las industrias construidas durante el juego. Salvo en un experimento, estos mercados son ilimitados e inmodificables, y no se ven afectados por las acciones de los jugadores.

<u>Gran industria</u>: propiedad destinada a la producción de bienes que se venden, incluso, a las economías externas.

Director del juego: instructor, profesor adjunto, o estudiante que dirige las actividades del juego. Anuncia; por ejemplo, los pasos sucesivos y el estado financiero de la comunidad.

Tienda local: propiedad destinada al comercio no especializado, y que realiza ventas al menudeo de los productos de consumo local, y que requiere una clientela menor que la de los grandes almacenes.

Caminos principales : carreteras con altas especificaciones, y usualmente con -- accesos controlados en los puntos que designa el director del juego - o las instrucciones del experimento, y que tiene, por lo general, un valor pesado de 1/2. Se usa sólo en contados experimentos.

Valor pesado: cifra que representa, en términos relativos, la eficiencia y costo de transporte a lo largo de un tipo de camino. Así los números 1 y 2 constituyen la relación entre el valor pesado de un camino principal y otro secundario, en la mayoría de los experimentos. El valor pesado, multiplicado por el peso asociado, nos da el costo de transporte a lo largo de una distancia unitaria del CLUG.

Oficina: representa el terreno utilizado en servicios administrativos, contables y directivos, prestados a otros negocios.

Predio: porción de terreno correspondiente a una manzana.

Pequeña industria: el terreno ocupado por una industria de importancia, dimensiones y eficiencia menores que las de la gran industria

Paso: declaración que hace el jugador de un equipo, rehusándose a construir o demoler, en esa vuelta.

Caminos principales: aquéllos que tienen, en general, accesos a todos los predios, y localizados por líneas particularmente señaladas en el tablero. Usualmente tienen un valor pesado de 1.

Reconstrucción: renovación de un edificio construido en vueltas anteriores, pagando al Banco o al Director del Juego una suma equivalente a su deterioro, reduciendo así las probabilidades de perder el edificio al jugar los dados.

Unidad residencial: concepto que representa tanto las habitaciones como la fuerza de trabajo que las habita. Pueden construirse para densidades que fluctúan entre R1 y R4.

Caminos secundarios: denominación que se da a las vías urbanas y caminos de segundo orden, y que sirven, principalmente a las zonas residenciales, para sus compras locales. Tienen un valor pesado de 2, 3, y 4, según los fines del experimento.

Fijación de precios: en este paso, los propietarios de las tiendas y oficinas recientemente construidas anuncian los precios de sus almacenes, mismos que regirán hasta la siguiente vuelta divisible entre cinco. El monto de dichos precios se fija tomando en cuenta la competencia — existente o potencial — y la magnitud del mercado futuro al que puedan atraer dichos precios.

Pasos del juego: secuncia de los acontecimientos que ocurren dentro de una vuelta del juego. Todas las actividades de una vuelta deben coincidir con el paso correspondiente.

Equipo: uno, o más jugadores que operan conjuntamente, y que tienen los mismos intereses económicos y políticos dentro de la comunidad.

Terminal: el punto hacia el cual se remiten los productos manufacturados por la grande y pequeña industria, para ser transportados y vendidos en el exterior. En algunos experimentos, un tablero puede tener más de una terminal.

Costo del transporte: el costo inherente al transportar de personas o mercancía, de un terreno a otro durante una vuelta. Se calcula de acuerdo con el tipo de transporte requerido por cada uso del terreno, en relación con el correspondiente peso asociado, considerando el número de unidades de distancia recorridas. Este se multiplica por el valor pesado del camino que se emple para cada transporte.

Tasa de impuestos: cifra que que, multiplicada por el valor catastral de los predios de un equipo, determina el monto de los impuestos que éste deberá pagar en cada vuelta. Usualmente se representa como un porcentaje, o como una cifra determinada por cada millar del valor catastral. Por ejemplo, una tasa de impuestos del 5% equivale a 50 por cada millar del valor catastral de un inmueble.

Distancia unitaria: la distancia existente entre dos esquinas de un predio, siguiendo las líneas del tablero.

Líneas de servicios públicos: líneas de transmisión o conducción construidas para la comunidad, con los fondos públicos, y que proporciona los servicios indispensables para el desarrollo urbano, como son: suministro de agua, drenaje, policía, protección contra incendios, etc. Estas líneas de servicio estarán conectadas con la central correspondiente.

Central de servicios públicos: las instalaciones, de propiedad municipal, que -- proporcionan los servicios públicos a las líneas respectivas. En algunos experimentos, las centrales de servicios son construidas por particulares.

DIVISION DE ESTUDIOS SUPERIORES
FACULTAD DE INGENIERIA, UNAM
SECCION DE PLANEACION

CLUG

JUEGO PARA LA SIMULACION DE SISTEMAS URBANOS

TIPO DE INVERSION	COSTO DE CONSTRUCCION	GANANCIA POR AÑO	NO. DE UNIDADES DE EMPLEADOS NECESARIOS	SALARIOS POR AÑO	PRECIOS TOPE SUPER MERCADO	PRECIOS TOPE GRAN ALMACEN	PRECIOS TOPE OFICINAS
Gran Industria	96 000	48 000	4	24 000	—	—	4 000
Pequeña Industria	48 000	22 000	2	12 000	—	—	2 000
Supermercado	24 000	*	1	6 000	—	—	1 000
Gran Almacén	24 000	*	1	6 000	—	—	1 000
Oficina	36 000	*	1	6 000	—	—	—
Residencia Sencilla	12 000	6 000	-	—	2 000	1 000	—
Residencia Doble	30 000	12 000	-	—	4 000	2 000	—
Residencia Triple	48 000	18 000	-	—	6 000	3 000	—
Residencia Cuadruple	72 000	24 000	-	—	8 000	4 000	—

El Ingreso por TL, GA y O dependen del número de clientes obtenidos y el precio cargado. Un ingreso bruto de $10 000 $15 000 puede ser supuesto antes de que esas unidades operen.

CLUG

DIVISION DE ESTUDIOS SUPERIORES
FACULTAD DE INGENIERIA, UNAM
SECCION DE PLANEACION

JUEGO PARA LA SIMULACION DE SISTEMAS URBANOS

COSTOS DE TRANSPORTACION POR CUADRA AL AÑO

VIAJES DESDE:	VIAJES HACIA:				
	CENTRAL	OFICINA	EMPLEO	COMPRAS (TL)	COMPRAS (GA)
Industria Pesada (GI)	$4 000	$ 400	-	-	-
Pequeña Industria (PI)	$2 000	$ 200	-	-	-
Tienda Central o Supermercado	-	$ 100	-	-	-
Residencias por unidad (RI)*	-	-	$ 300	$ 200	$ 100

* Para R2, R3 o R4, multiplicar por 2, 3 o 4, respectivamente.

CLUG
COMMUNITY FINANCIAL STATUS

Round Number	Total Community Assessed Value	Utility Costs @ $2,000/new @ $1,000/old		Social Service costs @ $1,000 per residential unit		Debt Service @ 10 % previous deficit	Total Community Costs	Total Taxes Raised	Round Surplus or Deficit	Cumulative Surplus or Deficit	Debt Limit @ 10% of Assessed Value	Debt as % of Debt Limit
		# new # old	cost cost	# units	cost							
1												
2												
3												
4												
5												
6												
7												
8												
9												
0												
1												
2												
3												
4												
5												
6												
7												
8												

DIRECTORIO DE ASISTENTES AL CURSO: MODELOS DE SIMULACION PARA LA PLANEACION INTEGRAL DEL USO DEL SUELO Y EL TRANSPORTE (DEL 23 AL 28 DE ENERO DE 1978)

	NOMBRE Y DIRECCION	EMPRESA Y DIRECCION
1.-	ING. ABEL AGUILAR ARELLANO Pascual Luna No. 21 Tezoyuca Edo. de Méx. Tel.	S.A.H.O.P. Insurgentes Sur # 1443-2o. Piso México 12, D.F. Tel. 563-80-76
2.-	ARQ. FCO. JAVIER AGUILAR LOPEZ Nicolás de León No. 125-A-5 Jardín Balbuena México 9, D.F. Tel. 552-55-55	DELEGACION VENUSTIANO CARRANZA Av. Fco. Del Paso y Troncoso Esq. F.S.T. De Mier Col. Jardín Balbuena Tel. 768-32-77
3.-	ING. FRANCISCO JOSE ALVAREZ CASO Ricardo Polmerin No. 11 Col Guadalupe Inn México 20, D.F. Tel. 651-43-96	
.-	ING. ENRIQUE BUENDIA PINEDA Benito Juárez No. 3 Col. Sn. Pablo Tecalco Edo. de Méx. Tel.	S.A.H.O.P. Insurgentes Sur # 1443-2o. Piso México 12, D.F. Tel. 563-80-76
5.-	LIC. ARTURO CALDERON ZUGASTI Av. Lomas Anáhuac Col. Lomas Anáhuac Edo. de Méx. Tel. 511-24-14	DIRECCION DE DOCUMENTACION Y ANALISIS DE LA S.P.P. Av. José Vasconcelos # 221-5o.P. Col. Sn. Miguel Chapultepec México 18, D.F.
6.-	ARQ. LETICIA CAVAZOS GARZA Monte Albán No. 199 Col. Narvarte México 12, D.F. Tel. 521-80-72	DIAMAT ARQS. CONSULTORES, S.C. Emerson No. 150-202 Col. Polanco México 5, D.F. Tel. 250-07-67
7.-	ANGEL CERVANTES VELAZQUEZ Paseo del Faisan No. 194 Col. Lomas Verdes Naucalpan, Edo. de Méx. Tel. 572-72-50	INST. DE ACCION URBANA E INTEGRACION SOCIAL Parque de Orizaba No. 7 Naucalpan, Edo. de Méx. Tel. 576-58-39

DIRECTORIO DE ASISTENTES AL CURSO: MODELOS DE SIMULACION PARA LA PLANEACION INTEGRAL DEL USO DEL SUELO Y EL TRANSPORTE (DEL 23 AL 28 DE ENERO DE 1978)

	NOMBRE Y DIRECCION	EMPRESA Y DIRECCION
8.-	ARQ. EFREN DE GORTARI RABIELA Vieña No. 271-4 Col. Coyoacán México 21, D.F. Tel. 534-13-53	S.A.H.O.P. Dr. Barragán No. 779-8o. Piso Col. Narvarte México 13, D.F. Tel. 590-44-01
9.	ING. ARQ. P. ALBERTO FRIAS RODRIGUEZ Av. Alfonso No. 86-10 Col. Alamos México 13, D.F. Tel. 538-29-13	FIDEICOMISO LAZARO CARDENAS Miguel Angel de Quevedo No. 276 Col. Coyoacán México 13, D.F. Tel. 554-11-45
10.-	SALATIEL GARCIA CHAVEZ Calle 311 No. 320 Altos Col. Nva. Atzacoalco México 14, D.F. Tel. 781-79-14	DELEGACION VENUSTIANO CARRANZA Fco. del Paso y Troncoso Esq. F.S.T. De Mier Col. Jardín Balbuena México 9, D.F. Tel. 768-32-77
11.-	ING. FERNANDO GIFFARD SANCHEZ Privada Díaz Mirón No. 22 Col. Nextitla México 17, D.F. Tel. 547-08-50	SECRETARIA DE COMUNICACIONES Y TRANSPORTES Centro S.C.O.P.
12.-	ING. HECTOR A. GONZALEZ COSIO ZENTENO Av. 11 Poniente No. 704 Puebla, Pue. Tel. 42-13-70	S.A.R.H. 3 Sur No. 1501-6o. Piso Puebla, Pue. Tel. 43-80-27
13.-	ING. PABLO GUZMAN ZACAPA Félix Berenguer No. 126-1er. Piso Col. Lomas Virreyes México 10, D.F. Tel 540-72-20	INARCO S.C. INGS. Y ARQS. CONSULTORES Félix Berenguer 126-1er. Piso Col. Lomas Virreyes México 10, D.F. Tel. 540-72-20
14.-	ING. ROBERTO HEATLEY CORTES Augusto Rodín No. 358-Depto. 307 Col. Nochebuena, Mix. México 19, D.F. Tel. 598-16-35	S.E.F.I. Palacio de Minería Tacuba Núm. 5 Mezzanine Col. Centro México 1, D.F. Tel. 512-33-53

DIRECTORIO DE ASISTENTES AL CURSO: MODELOS DE SIMULACION PARA LA PLANEACION INTEGRAL DEL USO DEL SUELO Y EL TRANSPORTE (DEL 23 AL 28 DE ENERO DE 1978)

	NOMBRE Y DIRECCION	EMPRESA Y DIRECCION
15.-	ARQ. EDUARDO HERNANDEZ GONZALEZ Misión Sn. Javier No. 81 Col. Jardines de Bulevares México, D.F. Tel. 562-18-00	INFONAVIT Barranca del Muerto No. 280 México 20, D.F. TEL.
16.-	GUSTAVO HERNANDEZ LOZANO Sacrificios No. 12 Col. Tlalpam México 22, D.F. Tel 573-28-31	S.A.R.H. Paseo de la Reforma No. 107 México, D.F. Tel. 566-06-88 Ext. 132
17.-	ARQ. MANUEL LAMA GUAGNELLI Cholula No. 107 Col. Hipódromo Condesa México 11, D.F. Tel. 515-57-28	U. N. A. M. Ciudad Universitaria México, D.F. Tel. 399-27-95
18.-	ING. ERNESTO LARA ROMERO Amores No. 1432-8 Col. Del Valle México 12, D.F. Tel. 575-71-41	S.A.H.O.P. Insurgentes Sur 1443-2o. Piso México 12, D.F. Tel.(575-71-41) 563-80-76
19.-	MAT. YOLANDA LAZO TISCAREÑO Zempoala No. 295-204 Col. Narvarte México 12, D.F. Tel. 563-80-76	S.A.H.O.P. Insurgentes Sur 1443-2o. Piso México 12, D.F. Tel. 598-00-23
20.-	ARQ. CARLOS BENJAMIN LUNA Los Pinares No. 118 Col. Centro America San Salvador Tel. 257-38-82	GARCIA Y ASOCIADOS 2 Ave. Norte, Edificio Murillo San. Salvador El Salvador Tel. 223-722
	Q. ALB ERTO MARTINEZ GUERRERO Juan Ruiz de Alarcón No. 3 Col. Circ uito Dramaturgos Satélite Edo. de M éx. Tel. 562- 14-88	CENTRO REGIONAL DE CONSTRUCCION ESCOLARES PARA AMERICA LATINA CARIBE Auditorio Nacional México, D.F. Tel.

DIRECTORIO DE ASISTENTES AL CURSO: MODELOS DE SIMULACION PARA LA PLANEACION INTEGRAL DEL USO DEL SUELO Y EL TRANSPORTE (DEL 23 AL 28 DE ENERO DE 1978)

	NOMBRE Y DIRECCION	EMPRESA Y DIRECCION
22.-	ARQ. JUDITH LEONOR PADLOG Rincon Del Bosque No. 8 Puebla, Pue. Tel. 428693	UNIVERSIDAD AUTONOMA DE PUEBLA Calle 4 Sur No. 104 Puebla, Pue. Tel.
23.-	ING. JOSE FRANCISCO PONCE CORDOVA Sur 77 No. 254-1 Col. Sinatel México 13, D.F. Tel. 581-08-08	ESTUDIOS DE INGENIERIA Y PLANEACION, S.A. Paseo de la Reforma No. 2165 Col. Lomas Chapultepec México 10, D.F. Tel. 596-29-77
24.-	MA. NELLY RANGEL GUTIERREZ Fausto Nieto No. 29 Col. P. E. Calles México 17, D.F. Tel. 547-17-30	DELEGACION VENUSTIANO CARRANZA Fco. Del Paso y Troncoso Esq. F.S.T. De Mier México 9, D.F. Tel. 768-32-77
25.-	EDUARDO REED M. Mazatlán 61-1 Col. Condesa México 11, D.F. Tel. 553-60-93	MONUMENTOS COLONIALES Plaza del Convento Churubusco Col. Churubusco México, D.F. Tel.
26.-	ARQ. ALEJANDRO RESENDIZ BALLESTER Donatello 25 - 304 Col. Mixcoac México 19, D.F. Tel 563-80-95	S.A.H.O.P. Insurgentes Sur 1462-3er. Piso México 19, D.F. Tel. 563-80-95
27.-	ALEJANDRO F. RODRIGUEZ SOLANO Miraflores No. 39-Depto. 25 Col. Porteles México 9, D.F. Tel. 723-31-43	S.T.C. METRO Delicias No. 67-5o. Piso Col. Centro México 1, D.F. Tel. 521-86-20
8.-	ING. IGNACIO ROSAS IBARRA 17 Poniente No. 1719 Puebla, Pue. Tel. 438139	UNIVERSIDAD AUTONO 4 Sur No. 104 Puebla, Pue.

DIRECTORIO DE ASISTENTES AL CURSO: MODELOS DE SIMULACION PARA LA PLANEACION INTEGRAL DEL USO DEL SUELO Y EL TRANSPORTE (DEL 23 AL 28 DE ENERO DE 1978)

	NOMBRE Y DIRECCION	EMPRESA Y DIRECCION
29.-	ING. ENRIQUE SALCEDO MARTINEZ Monte Alban No. 37 Col. Narvarte México 12, D.F. Tel. 519-66-47	S.A.H.O.P. Liverpool No. 3 Col. Juárez México 6, D.F. Tel. 592-22-39
30.-	ING. JORGE SILVA MIDENCES Bolivar No. 745-104 Col. Alamos México 13, D.F. Tel. 579-79-90	S.A.H.O.P. Dr. Barragán No. 778-8o. Piso Col. Narvarte México 13, D.F. Tel. 590-49-01
31.-	ING. ARQ. MORTIMER H. TAPPAN COPPEL Av. Insurgentes Sur No. 1677-110 Col. Sn. José Insurgentes México 20, D.F. Tel. 524-23-69	BANCOMER, S.A. San Juan de Letran 13-13 Piso Col. Centro México 1, D.F. Tel. 521-79-82
32.-	ING. MAURA J. TREVIÑO T. Arteaga No. 1526 Ote. Col. Centro Monterrey, N.L. Tel. 754-66-8	UNIVERSIDAD AUTONOMA DE NUEVO LEON C.U. San Nicolas de las Garzas Nuevo León Tel. 52-28-78