Environment and Planning B: Planning and Design, 1992, volume 19, pages 663-688

Urban modeling in computer-graphic and geographic information system environments

M Batty
National Center for Geographic Information and Analysis, State University of New York at Buffalo, 301f Wilkeson Quad, Buffalo, NY 14261-0001, USA
Received 6 July 1991; in revised form 5 May 1992

Abstract. Computer graphics and its broader application as visualization are at last beginning to have an impact on the ways analysts and model builders articulate and communicate their intellectual understanding of urban systems. Geographic information systems (GIS) lie in the vanguard of these developments and there is presently much effort in linking these systems to traditional spatial models so that the power of GIS databases and display capabilities can be harnessed. To demonstrate the import of these ideas, a prototypical visual environment has been developed for such modeling and in this paper it is shown how it can be implemented. The elements of such visualization are first defined in terms of the model-based processes which characterize applications, emphasizing data exploration, model calibration, prediction, and prescription. These processes are then concentrated through model inputs, outputs, and causal structures, illustrating the operation of various model sectors through different modes of spatial statistical and graphic analysis. These ideas are implemented through windowing systems which mix text, numerics, and graphics, and are illustrated using an elementary model of residential location of Melbourne, Australia. The operation of the model in a UNIX environment which is accessed through proprietary windows-based software is demonstrated, thus providing a platform to discuss the problems of linking conventional model-based techniques to graphics software in general, GIS in particular. The paper is concluded with a sketch for future work.

1 Introduction: spatial analysis and urban modeling

The ideas presented in this special issue of *Environment and Planning B* represent a number of very different traditions involving spatial representation, analysis, and display, but they represent a unity of purpose in their concern for tracing and exploiting various new and powerful computer environments which are likely to become the predominant way of using computers within the decade. In particular, the use of computer graphics in spatial representation is an obvious prelude to analysis, and the emergence of computer-aided design (CAD) and geographic information systems (GIS) in many different problem domains now provides a framework in which traditional spatial analytic techniques can be researched and applied. However, it is already clear that the goals for such development differ widely in the manner in which such systems might be exploited. Although there may be an apparent coincidence in purpose between the various analytic traditions in their use of computer graphics and related techniques of visualization, the focus of research using such systems may be very different. It is important to be clear about these issues before we embark upon ways of exploiting such systems for what might be very different purposes.

Two traditions which define the somewhat broad domain of spatial analysis are based upon the development of models for conventional scientific use or for applied problems of design. A more cogent way perhaps of making this distinction between science and design is to classify such methods as dealing primarily with understanding and explanation in contrast to forecasting and decisionmaking. Traditional methods of spatial analysis largely comprise the former with the major mode of analysis being based on statistical theory. Problems of model specification through parsimonious structures and model estimation based on consistent statistics dominate this field, although more recently there have been approaches such as those of Anselin (1988) based on spatial econometrics. Nevertheless, such approaches are strongly inductive with much emphasis upon defining the correct model. These are nicely portrayed in the recent texts by Cressie (1991) and Haining (1990).

In contrast, there are many methods of spatial model building which are not primarily focused on better explanation but are orientated towards guiding policy planners and decisionmakers through appropriate organizations for generating technical advice and policy. Such models may be based on some acceptable explanation of the system of interest, but invariably such models do not have the parsimony of statistical models nor is their development geared towards better explanation; their focus is on the prediction of impacts and the design of optimal plans (Harris, 1967). Usually such models are estimated to some database and at that point overlap with spatial analytical models of the inductive type, but it is rare that the performance of the model at this stage makes any difference to its eventual application as a forecasting device. Moreover, such models are often deductively specified and contain structures which are not able to be compared against real world conditions in any way. Recent expositions are provided by Bertuglia et al (1990) for temporal dynamic models and by Putman (1992) for static models.

There are of course other traditions in spatial analysis and modeling. There is, for example, considerable work in spatial interaction modeling which has evolved from both traditions, such as that portrayed in Fotheringham and O'Kelly (1989), although efforts to develop an appropriate statistical theory for such models have been developed in the tradition of econometrics at the level of individual rather than macro behavior (Ben Akiva and Lerman, 1985). There have also been a variety of optimization models coming from traditions in operations research and involved with facility-location problems. These models have been generalized to the wider context of decision support (Densham, 1991), but the link to traditional spatial analysis is considerably stronger than in the design and planning contexts alluded to above. Last there are models in the tradition of mainstream microeconomics, such as those involving location and bid-rent theory and in these contexts the emphasis is not upon spatial analysis but upon rigorous specification in terms of the theory of individual preferences and the equilibrium of resultant markets (Anas, 1987).

All these traditions are being affected by the revolution in computing, especially in terms of the potential use of computer graphics and the development of user-friendly interfaces involving text as well as graphics. For example, the very terminology of such methods—spatial decision-support systems, expert systems, computer cartography, GIS, and so on—indicates the important influence of computation as an organizing theme. Nevertheless, the responses to these challenges by analysts and modelers can be quite different. For example, Openshaw (1990), in speculating about the development of spatial analytic techniques which are associated with understanding and explanation, argues that such techniques can represent quite straightforward additions to the functionality of many propriety GIS. Ding and Fotheringham (1992) have demonstrated how standard spatial autocorrelation functions can be easily and effectively linked to such systems [in this case using the example of ARC-INFO (Esri Inc., Redland, CA)]. In contrast, it is much more difficult to embed conventional urban models within the framework of GIS and there is a tacit acceptance in the field that, wherever such coupling is possible, it is likely to be loose. Harris (1989) argues that metasoftware needs to be produced for 'planning-support systems' which act as an umbrella for many useful but diverse and difficult-to-connect software. Examples of the limitations of standard software are indicated in the context of expert systems by Batty and Yeh (1991) and for GIS by Couclelis (1991).

Although developments in graphics might seem somewhat different with respect to these two approaches, increases in computer power have had much the same effect on such models when it comes to increasing the speed and feasibility of application. There are many examples now of cases where parallel and supercomputers are being used to aid spatial modeling and analysis, although in several of these applications the data are so extensive and the processing so intensive that visualization of some variety is necessary to evaluate model outputs. Visualization is not simply a matter of absorbing data in a more manageable or intelligible form, because new patterns in the data can be identified spatially and hence new insights into real world processes generated. Moreover, the power to visualize also reinforces the increasingly popular notion that model outputs from computers must be evaluated visually, for many statistical techniques do not effectively communicate the quality of model prediction and estimation. In short, the notion that 'seeing is believing' is slowly returning to modeling, but this time it is in terms of computer graphics.

In this paper, I will concentrate on one of the main traditions, that based on urban (or land-use-transportation) modeling, and I will demonstrate how graphics in general can be used to visualize model processes, structures, inputs, and outputs. The visual environment to be developed is hardly a GIS, although it has some GIS-like features. Furthermore, although my concern will be upon developing graphic interfaces for models whose eventual use is in a highly applied, decision-specific context, visualization also provides opportunities for applied models to be subjected to standard spatial analysis. In this sense, the development of these new computing environments is drawing together subfields which originally were somewhat separate. In fact the environment to be created provides several possible lines of direction for theoretical and applied research. Not only will the framework provide an intelligible way in which experts and decisionmakers can communicate with the model and with one another, it also provides insights into the performance and operation of the model and gives some idea of how the mechanics of the model are executed. This is besides the major focus of such a system which is in enabling predictions and designs to be visualized. Such environments are so rich in possibilities for enhancing analysis and modeling that they will be dealt with in some detail before the formulation of the model is described in later sections.

My choice of model is one of the simplest possible. The model which I will use to demonstrate the development of graphic environments for urban analysis and forecasting is based on a singly constrained residential location model of the gravity type in which there is a single parameter controlling the effect of travel cost or distance. However, even with such a simple model which links two activity sectors—demographic and economic—in just one way, the graphic environment to be created is still rich in possibilities and I will exploit this in three distinct ways. First, I will sketch four processes of model building and application based on data *exploration*, model *calibration* (or parameter estimation), *prediction* (for forecasting), and *prescription* (or design). The model or its data can be subjected to each one of these processes which involve exploring the inputs and outputs of the model, and the graphic environment will be organized to reflect these. Second the inputs and outputs of the model and their causal effects will be examined in terms of the activity sectors of the model. In the model of residential location we can view its operation and performance from the residential or from the employment sector and these will provide different windows on the world of the model. Third, I will identify a series of different modes of analysis each of which can be applied to inputs and outputs. In particular, we will examine ways of aggregating or disaggregating the model, various methods of traditional statistical distribution and model performance, the use of effective graphical analyses of performance and model structure, the use of text modes to help the progress of the user through the model, and possible linkages to other standard sets of software giving the user access to the wider world of analysis and display.

Once the form of the visual environment has been developed, I will outline the prototype, which is based on three of the four model-building processes (presciption being excluded at present), two activity sectors (employment and population locations), and the five modes of analysis indicated in the previous paragraph. I will then present a full run of the model and illustrate this with a literal survey of the process of visualization based on the inputs and outputs of the model and the means used for display. But before this I must elaborate some of the ideas which inform the construction of graphic environments and relate these to the sorts of environment provided by GIS so that these possibilities can be evaluated in a somewhat wider perspective.

2 Computer-graphic and GIS environments

Computers have been used for graphics from the very beginning. There is a story, apocryphal perhaps, that in the late 1940s, researchers at Maurice Wilkes's lab in Cambridge University programmed the EDSAC machine to simulate the motions of a highland dancer. Of less dubious report, a young Jay Forrester who directed the Whirlwind computer project at MIT demonstrated how his machine could be programmed to simulate a bouncing ball and to play "Jingle Bells" on Ed Morrow's Christmas show in 1952. In fact, you can see a clip from this program in the Computer Museum on Boston's South Wharf. But, despite these early successes, the gestation period for computer graphics was long and arduous and, two decades later, little progress had been made. It was not until computer memory became widely available thanks to the microrevolution that computer graphics really took off. The direct correspondence of computer memory with the storage of pictures provided the impetus and, once this link had emerged in the late 1970s, dramatic progress was made.

During the last decade, the way we interact with computers has begun to change, in that graphics-based interfaces are now becoming standard in many systems. These were pioneered by Xerox PARC in the 1970s, then by Apple in their Macintosh computers and, more recently, most workstation interfaces are being developed using the windowing interface called X. However, in the 1990s, a major spin-off from computer graphics will involve the development of 'visualization' tools. Already researchers are displaying large quantities of data by means of graphics and are finding insights into their pattern and structure which hitherto were inaccessible. For example, in studies as diverse as turbulence, image processing in medicine, and the teaching of mathematics, these new tools of visualization are becoming the norm. Moreover, in our domain, techniques of computer mapping have now reached the level of widespread accessibility and one of the main reasons why GIS have taken off so rapidly in the last five years has been the availability of hardware in the form of cheap graphics-based workstations and the existence of much more user-friendly graphics interfaces (Dangermond, 1990).

In this paper, we will explore how visualization might generate appropriate mediums for developing some of these linkages but primarily for providing a suitable environment for urban models which focus on land use and transportation. Visualization is much more than simply generating graphics from large data sets which were hitherto too unmanageable to explore. Visualization as I will define it encompasses the entire process of interacting with data in its numeric, textual, and graphic forms but through visual interfaces of the kind which are becoming widely available on micros as well as on workstations. Visualization is not just confined to data.

It can be used to explore how models work, through diagrammatics and visual animation, thus augmenting the user's understanding of the model. Moreover, it is not too ambitious to suggest that visualization represents both a new way of doing science and new ways of generating designs and policies which characterize the goals of planning and management.

GIS are functionally extremely simple with their focus mainly on ways of storing and retrieving digital map data and linking such digital data to attributes associated with various aggregations of the basic set of digitized points defining the map. In fact, their development is much closer to the norms that have guided the development of urban models than those which have dominated the development of spatial analysis. In short, there is little theory to GIS except that associated with the optimal representation of spatial data structures, and the various functions which are incorporated into GIS have been added in an ad hoc manner as and when the need has arisen in practical problem-solving contexts. There is, however, one feature of GIS which shows how close such systems are to decision contexts while at the same time showing how problematic their extension is to incorporate functions involving spatial modeling and analysis. A central function of GIS involves the overlaying of digital data as carried out in many physical design-based operations, particularly in land-based resource analysis and planning. In fact, in ARC-INFO such overlaying is referred to as spatial modeling. The real power of such systems resides in their display facilities—in mapping—and the structure of most GIS is not easily adaptable to functions which involve numerical and statistical modeling. This can of course be done, as Ding and Fotheringham (1992) have shown, but the process is laborious.

There are several developments in urban modeling which have anticipated the link between GIS and model functions. In the Chicago Transportation Study in the mid-1950s, the need to plot traffic flows and desires between origins and destinations led to the development of a graphics device called a 'Cartographatron' which essentially was a purpose-built oscilloscope capable of displaying large sets of origin-destination data involving nearly five million trips (Creighton, 1970). The development of computer mapping was motivated by practical uses and the development of line printer plotting systems of various urban model data such as SYMAP (Baxter, 1976) and Forrester's (1969) 'Dynamo' time-series plots were quite widely known in the 1960s and 1970s. Early GIS for urban models were developed alongside such models, but with little of the fanfare accompanying their use today.

GIS are being coupled to analytic and modeling techniques in diverse ways, but two stand out on a continuum which spans the range from loosely coupled to strongly coupled systems. Simply linking inputs and outputs from one set of models to another set of GIS is the strategy of loose coupling (Brail, 1990) and some fairly ingenious schemes can be developed which make explicit use of the power of several systems (Ferreira and Evans, 1992). In the development of spatial decision support systems, Densham (1991) has shown how loosely coupled systems can be employed to advantage, although the software used embodies fairly standard graphics interfaces which ease the transition from one system to another and GIS which are better tailored to specific transportation and related planning problems.

Optimal schemes of coupling are only likely to exist for purpose-built GIS which incorporate spatial analytic and modeling techniques explicitly. An early scheme is contained in the UDMS (Urban Data Management Software) developed for planning in developing countries by Robinson and Coiner (1986) and systems in which such models are used frequently as the main analytic purpose of the GIS have been pioneered by Birkin et al (1990). There are GIS which are embedded within urban model frameworks such as those developed by Echenique (1983) and others reported in Webster et al (1988). Standard databases and spreadsheets also have the potential to contain analytic functions, such as those presented in Klosterman et al (1992) indicate. Last, we should note the suggestions of Harris (1989) in his development of planning support systems which represent wider frameworks within which diverse sets of software can be meshed. These ideas are taken further by Harris and Batty (1992), but in some senses are already being implemented by Ferreira and Evans (1992).

In fact, the growth of these types of information system in urban planning is nothing short of remarkable given the checkered experience with computer land-use and transportation models in the 1960s and 1970s. In one sense, though, this can be put in perspective, for the basic needs of planners involve information in numeric and map form. However, important problems in the use of GIS are emerging, in that available systems are difficult to link to the multitude of models and their software which are used at various stages of the planning process. Furthermore, there is a dearth of work in this area: there are difficult theoretical problems of embedding such systems into planning—GIS is not planning nor is planning GIS—and there are some technically very tricky problems in establishing appropriate interfaces between GIS and related software as Harris and Batty (1992) allude to and as Ferreira and Evans (1992) demonstrate.

Experience with the use of proprietary GIS in which analytic models and functions might be embedded is still fairly rare, but what limited experience there is shows that embedding and extending such functions is quite problematic although it can be done. In essence, the reasons for such embedding are largely those of expediency in that, if the data are stored in such a GIS to begin with or if the only means of displaying it is through the GIS, then such coupling might be attempted. With ARC-INFO, for example, the existence of a macrolanguage within the overall shell enables different functions of the GIS to be accessed in any order within the prespecified constraints required for valid processing. Moreover, users can escape from the shell and call other standard software or purpose-built programs, although the graphics interface in ARC-INFO is cumbersome to say the least. Control over colors, plotting position, window placement, and a host of other important issues relating to the interface are dubious and leave little flexibility for the production of good-looking designs. Such embeddings are slow to operate and the enormous overheads which such systems assume mean such developments are really only feasible where very large and very fast computers can be dedicated to these purposes.

I will continue this discussion by introducing a purpose-built graphics environment which has been constructed to visualize the inputs, outputs, performance, structure, and operation of the simplest of urban spatial interaction models. As described earlier, the environment and its potential for visualization will be based on three major facets of the model: the processes associated with building and applying the model, the relationships between the sectors of the model, and the various modes of analysis which can be applied to the model inputs and outputs. This environment will not give us anything like the sorts of functions which exist in terms of proprietary GIS, but most of these are not required in the kinds of application I envisage. A major feature of this design will be the extensive use of windows in several different senses which will be gradually introduced as the design unfolds. I will now deal with these components in turn, eventually assembling them into the graphics environment and demonstrating how insights and ideas can be communicated through such media.

3 Processes of model building and application

I have already implied that the development of urban models is based more on processes that involve forecasting and design—prediction and prescription—than upon the search for a better understanding, and hence explanation. However, there is a certain asymmetry here between spatial analysis and urban modeling in that the kind of modeling that can benefit from spatial analysis for the process of model building usually begins, like spatial analysis, with understanding and explanation. On the other hand, spatial analysis rarely involves prediction and prescription, and this narrowing of domain often focuses the links between spatial analysis and GIS somewhat more clearly than the relationship between the GIS and urban modeling. Nevertheless, models represent some embodiment of theory which is being tested through data and, in this sense, the process of model building follows the scientific method in which hypotheses within the model are continually being refined and retested. However, when models are used to aid the planning process, their use is for prediction—for sketching out the consequences of action, for impact analysis—or for devising some optimal prescription or plan of action. This second use, which is the ultimate goal for using models in planning, may be broadened even further when models are used to guide management, inform implementation, and provide a framework for monitoring the plan.

It is this confusion or concatenation of science with design that led so many modeling projects in the 1960s to come to grief. Then computer models were quite new in most fields, computers were rudimentary, and the theory behind such models was hardly developed. Model builders were effectively engaged in the scientific method, but within a context dominated by the pragmatics of urban management—hardly a recipe for the successful application to real problems. The experience then was salutary (Lee, 1973), although much of the reawakening of interest in urban modeling during the last decade has been occasioned through developments in computation rather than in any clarification of these issues (Batty, 1989; Boyce, 1988). Although the first process concerns the exploration of data, the models being used have been preselected and their exploration is simply a means of providing a context for their fine-tuning through calibration, with some minimal possibilities for altering model structure. In this context, then, models are chosen a priori and the process of their testing is one which is geared primarily to their use in planning, not their use in science.

There are several ways of identifying the processes which characterize the model-building environment as will become clear when I sketch the example. Here I will develop the four noted earlier—*exploration*, *calibration*, *prediction*, and *prescription*—and deal with these in turn. Exploration involves searching through the data to examine its inherent order and pattern, to learn about this pattern, and to engage in exploratory data analysis in the sense used by Tukey (1977) but always with some well-defined model in mind. This process does not involve a mindless search through data, but a highly structured search with appropriate spatial and statistical analysis leading to a subprocess in which a particular model is selected from a family, representing a variation on some general theme. In fact in this context, this process will not be extended to the search for a best model type in the inductive manner of spatial statistics, although the implications of this first phase of the model-building process are that some choice of best model might be made, either within this phase or by cycling around it through the process of calibration.

Calibration involves the statistical fine-tuning of the particular model selected to the given data. I call this process calibration rather than estimation, for the term gives more of a sense of fine-tuning in contrast to estimation, which is often used in more inductive scientific contexts. In principle, it is possible to change the structure of the model at this stage if calibrated predictions are compared with observed data, but in practice this would require the model user to iterate the process of exploration and calibration through reselection of the model, additional data collection, and so on. Although this is the implication of the process, the particular interface that has been designed does not allow for such a possibility for it is assumed that any user who wishes to alter the structure of the model or even model type would do this 'off-line'. In fact, although data have rarely been explored in terms of visual pattern prior to the time when cheap computer graphics emerged, it is the calibration stage which is most problematic; in this context, there is some doubt that the traditional means of evaluating the performance of such models through numerical measures of the goodness of fit have ever provided the most appropriate criteria. Models which fit the data rather well numerically have often turned out to be extremely poor when their predictions are mapped spatially, displaying many types of bias which conventional statistics do not reveal.

Once the model has been calibrated it can be used either for prediction or for prescription or both, depending on the broader planning context. Hitherto models in planning have mainly been used for prediction. Through changes in the input data which might incorporate some ideas about a planned or unplanned future, models have been used to generate consequences and impacts. Furthermore, models have been and can be used to give some structure to the process of managing a spatial problem or plan rather than generating the plan itself. This latter use—prescription—involves using the model to identify an optimal or best plan, subject of course to certain objectives and goals which define the evaluative phase of the planning process (Harris, 1967). In fact, models have rarely been used in this way in strategic land-use planning, notwithstanding the many designs for such models which exist in the literature. In any application, these four processes would run alongside the planning process, being coupled to this more general process in various ways; these are fairly obvious and have been widely discussed for the last twenty years. An example is given in Batty (1978) and various extensions to these notions are provided by Harris and Batty (1992) in their discussion of the role of models in planning-support systems.

These four processes structure the way the user interacts with the model in the order given above. Each process can be interrogated by the user in a standard way, thus providing a comparable set of graphic, numeric, and statistical outputs which can be accessed for comparative purposes at any stage. Each process acts upon the same model structure but from different perspectives. The data used by the model are preselected and provide the base comparison with respect to the model calibration. Predictions and prescriptions also produce comparable outputs and in the visual interface which is anticipated, comparisons of inputs and outputs in map form for all these processes are possible. In fact, in the example to be presented later, such comparisons emerge directly because they are traced out and held for each process as soft copy (on the screen). Although we have mainly concentrated so far on visualizing model outputs in map form, extending the visualization to other types of schematic will be an important priority in further development of the environment.

It is necessary, however, to develop models for which these four processes can be unambiguously presented. The first three processes are pragmatically necessary in that any model requires data against which its calibrations and predictions must be judged. However, only a subset of model types can be interpreted in terms of optimization. One of the triumphs of spatial interaction modeling during the 1970s was the emergence of model structures which could be interpreted as a direct consequence of some optimizing scheme. Entropy maximizing was linked to more general nonlinear optimization, subject to constraint and to special case linear programming problems which emerged as extreme points in the array of possible model structures and types (Wilson et al, 1981). These too were linked to developments in disaggregate travel-demand modeling where model derivations using random utility maximization were shown to be consistent with those produced through entropy maximization (Anas, 1982; Ben Akiva and Lerman, 1985; Fotheringham and O'Kelly, 1989). The many developments in integrated land-use-transport modeling served to link these various processes in operational terms, thus establishing these four processes as essential to the comprehensive application of any such model (de la Barra, 1989; Putman, 1983).

The optimization paradigm must be developed a little more deeply in this context. If the model structure can be treated as an optimization problem, then the potential exists to view this optimization as either an actual or an ideal process. In short, the optimization might be positive in the sense often used in economic maximization, for example in utility theory, or it might be normative in the sense used in operations research where problems involve generating better, and hence more optimal, allocations of resources. If as in spatial interaction the optimization can be seen as either actual or ideal, then the prospect exists for each of the four processes to be treated in terms of optimization. Clearly, entropy-maximizing models can be calibrated in relation to their optimality conditions which are usually the model equations, or by maximizing entropy. Prediction is a straightforward extension of the calibration problem, whereas prescription might involve changing the model constraints in some fashion and developing designs through direct optimization. Moreover, considerable insights into such models can arise if their duals are formulated and solved and this, as Wilson et al (1981) show, can also lead to faster methods of solution. I will sketch these ideas more formally in a short appendix in which I present the structure of the model used.

Although we will be concerned here with static models which operate at cross sections in time and which in terms of predictive or prescriptive use assume the existence of an immediate equilibrium, great strides have been made in the field of dynamics during the last decade. Dynamic processes open up a completely new dimension of model testing as well as use and application, and could be used as yet another basis for structuring the visual interface. As we shall see, the interface is based on spatial comparisons through map forms, but dynamic models open up the possibility of spatial animation in presentation and communication. These ideas will not be taken further, for the interface here simply requires one-shot predictions and descriptions to represent an 'instant' dynamics; but in future work, real-time dynamic models will represent yet another way of enriching our visual environment.

4 Representing urban models as interacting sectors of urban activity

The way models influence their users is directly in terms of their structure. Urban models have become increasingly modular as partial models based on single subsectors of the urban system have been developed. Comprehensive urban models now involve stitching partial models together in some sense. The mechanisms of the stitching are often ingenious and can be so complicated as to dwarf the sectoral structure of the overall model, although there is fairly general consensus as to the types of sector which need to be represented in such models. In essence, the urban system can be divided into two broad domains, that of demography which is driven by natural processes of reproduction and aging as well as migration, and that of economy which represents the way production and consumption—demand and supply—condition the markets of the local economy.

These domains can be cross classified and subdivided into several activity sectors, sometimes envisaged as land-use sectors. The residential location and housing sector is one of the most important, and the breakdown of industries into basic and nonbasic, which are exclusively associated with primary or secondary, and tertiary activity, respectively, still constitutes an important means of distinguishing these sectors.

Transportation as a sector is often implicit and is usually relegated to the background context of spatial interaction. The best-developed partial models based on single sectors are those involving retail location which link population demand to the supply of retail goods in shopping centers, and residential or housing-market models, although more complex, have advanced considerably in the last decade. Models of industrial location are less well-developed and other conceptualizations of urban activity systems based on different typologies are beginning to suggest themselves.

A major problem is the changing definition and conceptualization of the urban system in terms of the way activities combine and interact. The basic-nonbasic split is no longer exclusively associated with primary-secondary-tertiary classification. Each of these groups of industries has its own basic and nonbasic functions and these are also changing as automation is reducing the demands on labor and making industries more capital intensive. The service sector is elaborating into other distinct types such as quaternary or education services and quinary or home-based services. Moreover, the distinctions between various types of public and private sector are also generating new economic distinctions between various activities.

There have been many attempts during the last two decades to develop urban models which treat two or more sectors of the urban system. The original Lowry model is the example par excellence which led the way, although advances on the original structure have been much limited by the availability of data and other resources (Batty, 1976; Putman, 1992). However, in general there are at least three types of linkage between sectors which need be represented in an urban model. First, there are functional linkages based on patterns of economic dependence, best seen in the traditional multiplier relations characterizing input-output models. Second, there are in explicitly spatial models interactions or flows of people and goods across space. In fact, these are usually the physical manifestation of functional linkages. The third linkage involves those difficult-to-measure intangible links which can be broadly classed as communications which enable goods to flow and people to move and represent the mechanisms which clear markets, fix prices, and determine subsidies. At any cross section in time, it is assumed that these interactions are stable in forming the glue which holds together urban activities in some observable equilibrium.

The model which will ultimately be developed is a static model structure based on multiple sectors whose functional and spatial linkages are separable; the functional linkages do not manifest themselves spatially but only between sectors whereas the spatial links are uniquely determined across space. Of course, spatial interactions which do exist are also influenced through functional linkages between sectors, but these are not treated simultaneously. Another important issue involves the exogenous or independent variables which drive the model. In fact, a model will be adopted in which there is an exogenous as well as an endogenous part for every activity sector, thus enabling some part of every activity to control and be controllable as well as admitting some uncertainty into the predictive base of the model.

In essence, the model which it is hoped will ultimately be built is an extended spatial input-output structure predicated on the existence of multiple urban activity sectors linked by economic multipliers across sectors and by spatial interactions across space. There have been several attempts in the past to develop such models, although many of them have been somewhat inconsistent (Macgill, 1977). Furthermore, this model incorporates notions based on an approach to modeling through activity-commodity relations (Broadbent, 1973) and also relates to the integrated input-output modeling developed by Batey and Madden (1981). The model structure to be developed here subsumes the nonlinear relations characteristic of spatial interaction into the linear framework of input-output relations.

This characterizes the model in its predictive form. In its prescriptive form, the model can be formulated as a nonlinear program incorporating spatial interaction hypotheses subject to the input-output relations, which are linear constraints. A detailed statement of the model is provided in Batty (1986) and an early version was applied to Melbourne, Australia (Batty, 1983). Residential activity and its interaction to the employment sector in the Melbourne example are the sample data set used below to illustrate the operation of a pilot model within the visual environment being built.

The extended model has not yet been built, although its theory has been clearly articulated. However, the existence of multiple sectors has major implications for the visual interface and these have not yet been explored. In essence, it is the intention to design the interface so that the workings and predictions or prescriptions generated by the model can be viewed through separate 'windows' on each sector. Each sector has implications for all other sectors as well as being influenced by all others, and it is these specific relations that the interface will be designed to capture. A good understanding of the model structure and the effects of changes in each sector and their repercussions on others can be controlled through each of these windows and these windows can then be compared and contrasted to provide an overall picture. Thus, for each window on each sector, the four processes of exploration, calibration, prediction, and prescription can be examined and related across sectors and across processes. In essence, then, what will eventually be designed is a visual environment, where the activity sectors and their interactions through which inputs and outputs of the model might be examined, can be continually defined and redefined. Only in a very restricted sense can we demonstrate the environment here in a model based on a single interaction between two sectors, one of which is assumed as given. This does, however, provide an inkling of the idea that such a model can be viewed through different perspectives on the mechanisms used in its operation.

The notion of manipulating contrasting processes and sectors which form the basic building blocks of the interface is reminiscent of the sorts of inversion which characterize many problems which can be characterized in primal or dual terms. Although this analogy does not completely follow through, even in terms of the predictive and prescriptive processes, the notion of examining a model in terms of its input, outputs, and causal processes from diverse yet complementary controlled perspectives is attractive. It also suggests some of the potential which visual interfaces are able to bring to the process of model building and the use of models in the wider design and planning process. As will be shown in the appendix, many spatial interaction models can be represented in their primal and dual forms and this represents a switch in perspective already partly embodied in the calibration phase of the model operation.

5 Statistical and visual modes of analysis

So far I have conjured up an image of a visual environment which is dominated by modeling and planning processes and interlinked activity sectors, all of which create considerable potential for tracing effects and making comparisons of inputs and outputs which hitherto have not been possible. In fact, I have not indicated the most usual of processes concerning the development of urban models which involve various standard analytic and exploratory procedures applicable to all models. The most obvious outputs that provide the raw material of windowing systems involve graphics, in this context maps of various spatial distributions, point patterns, flows, and so on. If the model were dynamic in the temporal sense, then time patterns and profiles would provide standard output with the added prospect of space-time graphics.

The main outputs from the system will involve graphics, but in any interpretation these will require numeric, statistical, and textual supplements which generate further windows on the model. Text is straightforward and can be built into the database directly, but statistical and spatial outputs must be developed in numerous ways, usually on the fly. To exploit these possibilities, it is worth thinking of the relevant window on the model sector and process in terms of layers. Once a graphic, statistic, or numeric has been displayed, the user can then bring out another layer associated with this artifact in an effort to extract a more meaningful interpretation of the issue in question. There are several ways of generating such analyses and I will list these before I elaborate upon them: these involve spatial aggregation, standard statistical distributions, statistical measures and new graphical representations, text screens, and access to standard packages outside the shell within which the model and its interface reside.

The first layer of analysis concerns spatial aggregation or disaggregation. A major concern in spatial analysis is the effect of scale on interpretation. By building in a zoom capability which can enable both aggregation and disaggregation, some measure of the robustness of the model to scale change can be developed. This is more than simply showing some phenomena in more detail. It can involve actual aggregation or its reverse involving averaging or decomposition, or it might even involve running the model at different scales in the first place. The essence of such analysis concerns the extent to which the model is invariant to scale, and an assessment of the consequences for interpretation without judging in advance the relative merits or otherwise of invariance. Moreover the process of zooming in or out might be used to suppress or enhance detail in the interests of legibility. This is one of the most interesting features of the graphic environment, in that for any reasonably sized problem measured, say, in terms of the number of zones, it is essential to aggregate the inputs and outputs to generate meaningful comparisons on screen. Thus, without explicit intention, the environment necessarily provides some base on which to explore the modifiable areal unit problem and ecological fallacies such as those discussed by Fotheringham and Wong (1991). Furthermore, the whole question of actually running the model at different levels and the stability of their predictions is broached through this perspective.

The basic mode of representation in these types of model is through the thematic map, although other forms of graphic are useful. In particular, histograms and frequency profiles and all the sophisticated battery of new statistical graphics can be brought to bear on much of the data. The possibility of overlaying the spatial and the nonspatial and tracing directly the links between the two is a popular technique at present and it is clear that a variety of new and useful graphics is beginning to emerge as researchers find new and ingenious ways to explore and present observed and predicted data. Conventional statistical measures are usually part and parcel of these sorts of interface, and text screens can provide useful help facilities for the model processes as well as the processes required of the user in operating the visual interface. It is through this mode of analysis that the problems of choosing appropriate numerical, statistical, and graphical goodness-of-fit measures can be resolved and a balanced assessment of performance provided.

There is the possibility of using other standard software for analytic purposes. For example, the sorts of mapping capability in standard packages such as ARC-INFO could well be invoked in model structures such as these. Access to statistical packages such as MINITAB (Minitab Inc., University Park, PA) are possible if some serious and extensive analysis of any numeric data associated with the model is required. Leaving the shell in which the model resides is no problem on most computer systems and in most languages, but the benefits of engaging the greater computational power of other standard programs must be matched against their comparative slowness due to the enormous overheads which such systems carry. Frequently the data set of an urban model is so specific that invoking the resources of a GIS to plot one map, say, is hardly ever worth the time, especially if it simply requires the relevant piece of code to be written.

I have said nothing yet about how the user might interact with this visual environment. There are various sensory devices available covering all types of exotica such as the electronic glove-type systems favored in the Media Lab (Brand, 1987). Here, however, such interactions will be restricted to the keyboard and the mouse. The mouse is used to activate each independent window which is opened by the program and the keyboard is used for most data entry and response. However, there are hopes of extending the mouse input, output, and choice mechanisms by means of pulldown menus, methods for visually changing data and parameter values, and methods for selecting ways of storing the visual output from the screen for printing or later inspection and study. To summarize then, these screen-based processes dominate the analysis to some extent in that much of the output centers around the map, although the ways in which outputs from different processes and different sector windows can be compared enriches the analysis considerably over traditional model building.

Before the prototype visual environment is demonstrated, we should once again note the various goals which are assumed in this project. The issues just sketched as being important to the graphic environment tend to overwhelm the effort because they provide so many possibilities and problems for further research that they can obfuscate the purpose of the project. In some senses, this is quite intentional because we are building a system which not only has its own problems of design but also enables us to address a range of other problems through its very existence. For example, questions of display ranging from choices of layout to color to representation suggest several thorny research problems, as do questions of packing the maximum amount of information on the screen, the order in which maps and other forms are displayed, the contrasts between numeric and graphic display and so on and so forth—all these issues are raised by this framework; and this is before substantive questions of model building, urban theory, and spatial analysis are raised. Moreover, the model is ultimately being designed for a region (Buffalo, NY) whose data problems are perverse in that the database will cross an international border. The substantive problems of the region such as cross-border flows due to differential spatial competition and free trade zoning and location would all be relevant and hence must be captured in any interface developed.

6 Prototyping the visual environment

An appropriate way of describing the environment being created is to list the structures already defined and to examine ways in which the processes involved might fit together. In essence, I have defined *processes* (of model building), *sectors* (defining the urban activity types of the comprehensive model in question), and *modes of analysis* (of the model inputs, outputs, and causal structures). Four processes ordered from exploration, calibration, and prediction, to prescription can be developed for different model sectors based on the main urban activities of residential, industrial, services, retailing, education, leisure, and transportation. There are at least five modes of analysis which can be applied to each process sector: namely, spatial analysis through aggregation, spatial statistics, statistical graphics, text-based explanatory systems ranging from tutorials to help screens, and formal links to other software packages accessible outside the shell of the visual environment.

The least problematic of these features relate to the modes of analysis which can be developed through a series of layers or overlay windows with respect to any of the processes or sectors under scrutiny. It is the process and sector characterizations which pose trickier problems of design. It is possible to organize the environment with processes at the basic level and sectors branching off from each of the processes or with the sectors forming the base upon which the various processes are developed. For example, the interface may be structured in terms of the four processes, with some part of the interface showing all these processes at any one time but with the analysis of different model sectors being displayed with respect to the process in question. Alternatively, the sectors might form the base, with different processes characterizing the various sectors at any instant of viewing. In fact, a more ambitious scheme would be to combine these two general approaches so that the user might move logically from a view of sectors through processes to processes through sectors.

To illustrate these points, a couple of simple examples are in order. Imagine exploring the data of a two-sector urban model, first in terms of, say, the retail sector, then, say, its industrial sector. Other processes such as calibration can be examined in turn for the two sectors in sequence. Previous processes can be recalled by the user at will and some index of progress through the sequence of the four processes is likely to remain as a visual trace on the screen, reminding the user of his or her position in the overall model-building process. It is possible, however, at least in principle, to juxtapose this approach with the other at any point. Imagine during the prediction process that the user, in examining the retail sector, wishes to recall the way the previous processes of data exploration and calibration have impacted on the sector in question. It should be possible to shift viewpoint from the 'process-then-sector' approach to the 'sector-then-process' approach but within the logical limits established by the order of the overall structuring of processes. It would not be possible suddenly to examine, say, calibration if the user had not yet reached this process. Such limits would be obvious constraints on the manipulation of process and sector analyses.

In fact, this ambitious scheme has not yet been invoked. Eventually it is intended to develop a multisectoral model of the Buffalo-Niagara urban region which will involve the four processes of exploration, calibration, prediction, and prescription. So far only the kernel of this idea has been implemented and this is for an urban region in another continent, Melbourne, Australia, for which there is a small and manageable data set (Batty, 1983). The model is a single-sector residential location model. In strict terms, this model involves two sectors because it is based on simulating the journey to work from residential to employment locations. The employment sector, however, is not modeled as the location and size of employment are exogenous to the model and the residential sector is the object of simulation. The formal model is outlined in the appendix.

Only three of the four processes of model building are represented in the prototype and the way these are interrelated is shown in the flow chart in figure 1. Note how each of the three processes accesses common numeric, statistic, and graphic output routines. The prescriptive-optimization-design process has not been developed so far, largely because this is the least used of these processes, although the theory does exist and can be easily implemented. With respect to the various modes of analysis, these are much more straightforward. So far, the idea of zooming in and out on detail in the spatial form of input and output can be demonstrated (although there is no actual aggregation or disaggregation of the data because the example used here is too small), and statistical graphics in the form of frequency distributions have been developed for certain spatial predictions. Text screens, numeric inputs and outputs, and standard spatial statistics are a feature of the prototype and these form the common routines indicated in figure 1. So far, there is no ability to escape from the shell to other software systems which might augment the analysis, although, once again, these pose no particular problems in principle and escape to the operating system and other purpose-built program modules is used frequently. The interface is already rich in possibilities and, even in the prototype, it has been necessary to contain the range of possible options.

The model to be described is based on data for Melbourne, Australia, which has been divided into eight zones reflecting a dominant central business district (CBD) around which there are three inner suburbs and four outer suburbs of fairly low residential density. There is some cross commuting, but the dominant pattern in the journey to work is from the suburbs to the CBD. A singly constrained residential location model has been developed in which the interactions from the origins are constrained to reflect predetermined totals of employment. The model uses a negative exponential function of travel time for its deterrent effect, the parameter of which is the subject of the calibration. The user has a choice of specifying the model with upper bounds on the amount of employment attracted to residential locations based on the method developed by Walsh and Gibberd (1980), and of running the model with or without residential attractors. The goodness-of-fit criteria used to calibrate the model are based on the absolute difference between the observed and predicted mean trip time or the absolute difference between the observed and predicted values of the system entropy (Batty, 1976). The parameter values determined by these two criteria are formally equivalent (Champernowne et al, 1976), each being the dual of the other in the entropy-maximizing formulation. The method of calibration is based on dichotomous search across the parameter

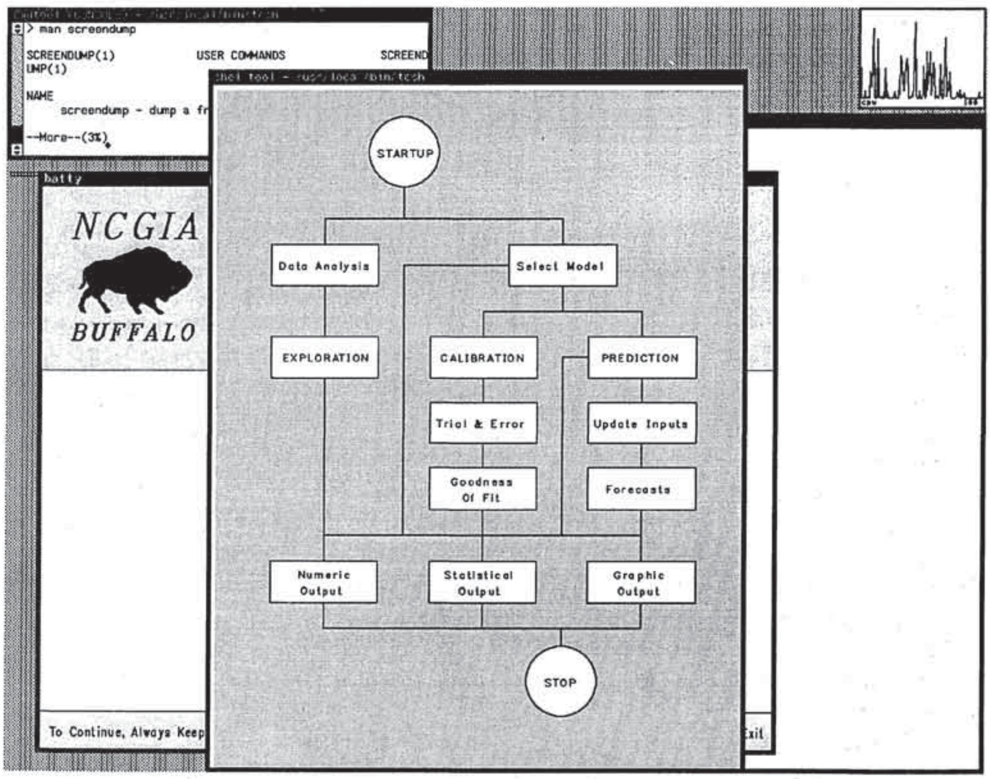


Figure 1. The processes of exploration, calibration, and prediction in flow chart form.

space, which is an approximation to search by golden section or Fibonacci search (Batty, 1971). An outline of the model is given in the mathematical appendix.

As implemented, the program sequences the three processes of exploration, calibration, and prediction in the order shown in figure 1. It is possible to miss any of these stages, but the user always finds himself or herself having to pass by each stage before moving to the next. It is not possible at any stage except the last, which is prediction, to return to an earlier stage, but it is possible to return from a predictive run of the model to make another predictive run, and so on. There is nothing intrinsically difficult about returning to earlier stages and this detail has simply not been incorporated into the structure so far, nor was it in the original program developed in Melbourne in 1982. As noted already, a prescriptive process based on optimizing the model has not been developed here, although the model is highly suited to such optimization. In fact, the use of the entropy function in the optimization at the calibration stage is akin to the sort of optimization which might be developed in the prescriptive stage as indicated in the appendix; this extension is very definitely on the future research agenda for this particular prototype.

The program is written in Fortran 77 as implemented on Sun workstations and consists of around 3000 lines of code, the model itself being some 10 lines, indicating that most of the programming is concerned with input, output, and model-building processes. The graphics language used is CGI, which was developed specifically for Sun workstations, although it is eventually intended to rewrite the program and its more elaborate versions in the more flexible C language. At present the program runs in the Sunview environment, making extensive use of the window capability in that medium, although CGI graphics are only available under this windowing system, that is between Sunview and the operating system. Consequently, it is not possible to make complete use of Sunview because, when a window is opened and subsequently closed, the graphics in previous windows, which are covered, are lost and thus have to be redrawn. Only when the model is reprogrammed in C can the power of Sunview be completely utilized, although by that time, the model will be running under Open Windows or X. The machines on which the program is run are Sparc Stations and in Buffalo, at least, it is not possible to run the program on earlier workstations such as 3/60s. For the most effective use, a high-resolution 19-inch color monitor is required because, as the subsequent figures show, the whole screen is packed with visual detail.

It is necessary to say something about the use of windows in Sunview. As has already been implied there are two types of window which are used in the program. The first type involves windows which simply display some graphic or numeric which the user wishes to examine in a different way and is thus created by the program itself. This is the way layers of windows are added to any program which requires convoluted paths to be set up within the program itself. The second and more usual type of window associated with workstations, however, represents a window in which a different program can be run. Windows in Sunview, for example, correspond to different devices, running different programs and in this sense, windows represent the visual interface to multitasking. This means that whenever a Sunview window is required, the user must intend to run another program within it.

The way these types of window are used here, for example, is to begin with the shelltool, which is the window opened up on entry into Sunview, and then to spawn other windows from within them. Thus the program in each window is spawned from the program running in the previous window and in this way, there is a strict hierarchy of processes in operation. Note that a process is now defined (as by Sun Microsystems) as a particular operation on a device within a window (in this case it being an executable program). There are six windows which are chained from each other, the last being the model itself. The windows are defined and chained as follows from a fixed starting window (a):

$$\text{explanation (b)} \rightarrow \text{logo (c)} \rightarrow \text{flowchart (d)} \rightarrow \text{mouse (e)} \rightarrow \text{model (f)},$$

and the way these are positioned on the screen is shown in figure 2.

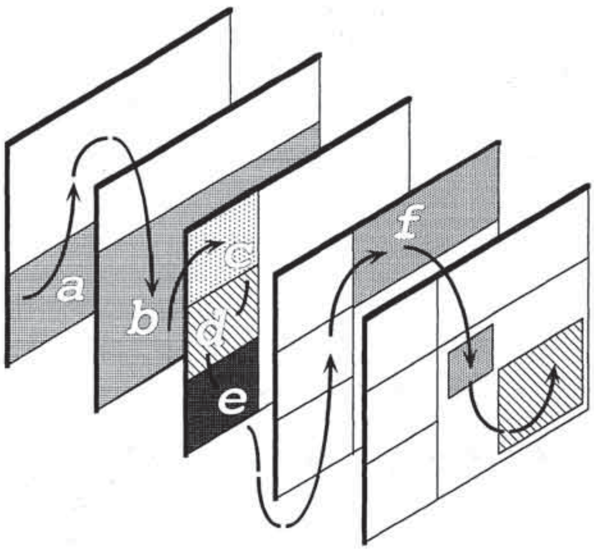


Figure 2. The windowing systems and the layout of the full screen.

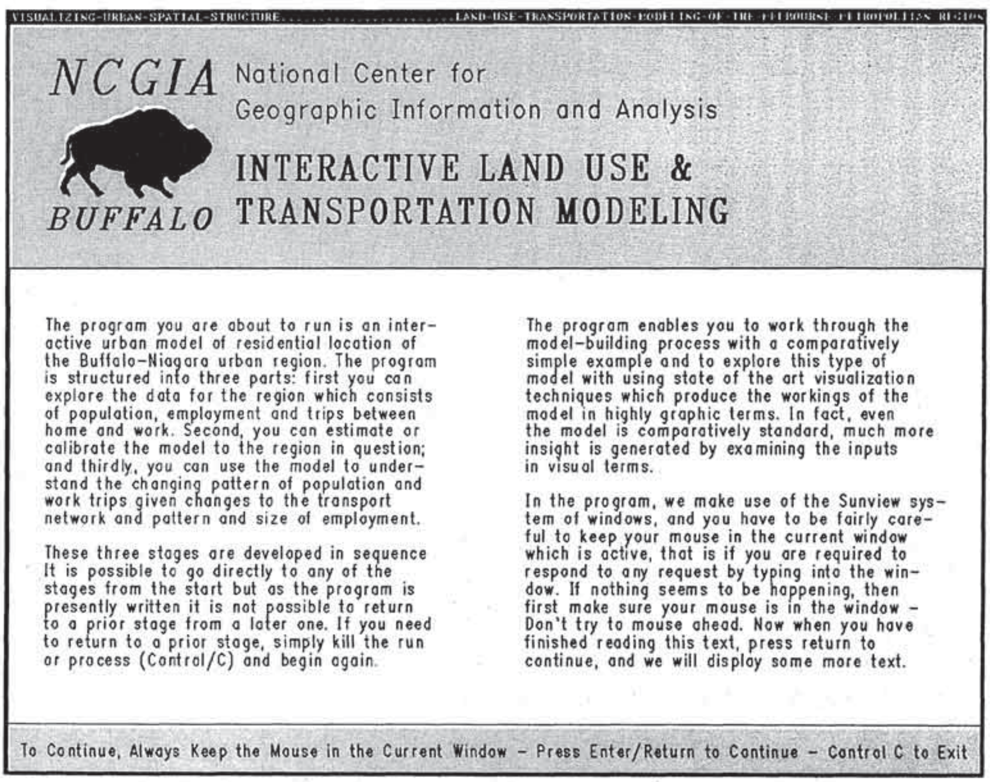


Figure 3. A screen of explanatory text.

Explanation consists of two full screens of text which describe the content of the program to the user and also alert the user to any idiosyncracies of the application. The logo appears in the explanation, but is then scaled down and rerun. A flowchart of the model is displayed (as in figure 1) which shows the way the model is structured around the three processes of exploration, calibration, and prediction, and also the way in which textual, graphic, and numeric outputs are common to all three processes. A mouse—literally—appears in the next window to remind the user that at all times a computer process will only run if the window is active and this involves locating the mouse in that window. Finally the main program is called. As yet alternating text windows have not been set up but these will eventually be incorporated. In fact, the model window is referred to as a graphics tool in Sunview and this splits the screen into a smaller text screen at the top of the window and a larger, more or less square, graphics screen at the bottom. The mouse has to be in the top, text window whenever the user wishes to respond to the prompt made by the program, usually asking for data, decisions on what options to choose, or simply rolling the program onto the next stage. These windows and the way they are called are shown in figure 2 with each window defined from a to f. Figure 3 shows part of the text of the explanation (b). Figure 4 shows the logo (c), flowchart (d), mouse (e), and layout of the model window f, prior to the main model being run.

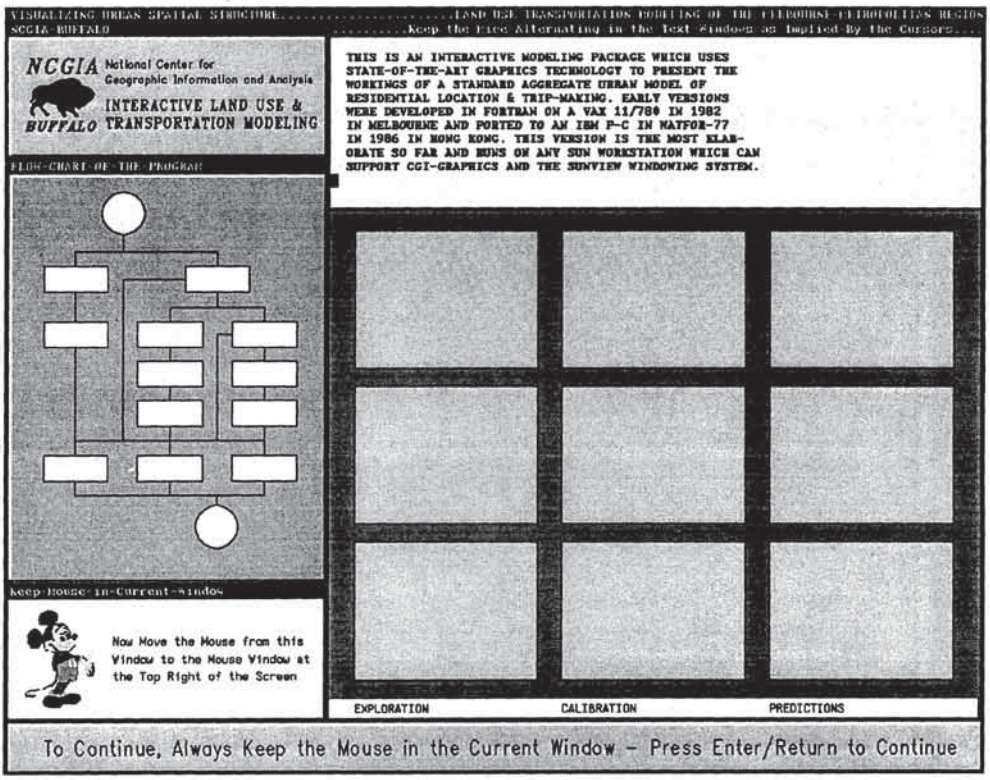
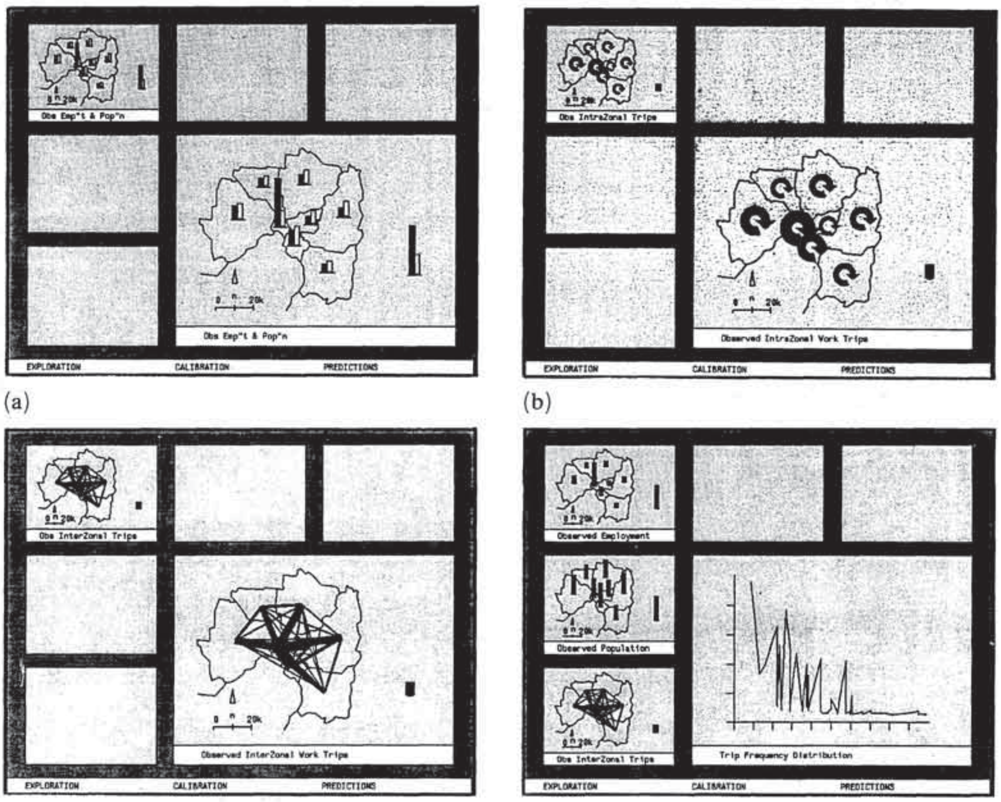


Figure 4. The complete configuration of the window systems on the full screen.

7 A run through the prototype

I will now conclude the paper with a description of a typical run of the model which is the subject of the last Sunview window, defined as f in figure 2, and is illustrated as the large text-graphics window in figure 4. When this window is entered, the layout of the window is fixed, first in terms of the text and graphics panel sizes. Then the graphics area is divided into nine smaller windows, so arranged that they fall into three vertical panels, the first being related to data exploration, the second to calibration, and the third to prediction. It is in these panels that comparable graphics are displayed enabling the outputs of each process to be compared. The screen layout at this stage is that shown in figure 4. In essence, the user begins with some general explanation of what the program is doing and this is presented in the upper text window of the graphics tool. The user can at any point choose to pass by these explanatory text screens. Then the data are read into the model from files and the user is prompted to explore them or not. Whatever decision is made, the model sets up a panel of three small windows (which have already been laid out) and if the user chooses to pass by this stage, three maps are printed in each of the three exploratory data-panel windows; these illustrate the distribution of the origin, destination, and flow data for the residential location model and these are judged as being the most important elements of the spatial model, remaining as the visual trace of this process.

If the user enters the data-exploration stage, then he or she has a choice of nine maps to explore. Whenever a map is printed, the user is asked whether or not the map is required in greater detail. The zoom capability is then initiated and a larger version of the map plotted. In figures 5(a), (b), and (c), typical samples are shown of these graphics which deal with the location of activities (by bar graph), intrazonal flows (by circular arrows), and interzonal flows (by desire-line diagrams). The zoom capability is also shown in these figures. For example, the flow matrix



(a) (b) (c) (d)

Figure 5. The process of data exploration: (a) bar graphs of locational data, (b) intrazonal trips, (c) interzonal trips, (d) the final visual trace of this process.

based on the journey to work is shown in figure 5(c) and an enlarged version of the map is also displayed in another area of the screen. If the flow matrix is selected for plotting, then the user is also prompted as to whether or not a trip-frequency distribution is required which is plotted in the enlarged window. These operations clearly invoke different types of window in that these windows are generated from within the model, and simply represent ways of developing more detail. Once the user has finished with this stage, three maps are plotted—namely the origins, destinations, and flow data linking these locations. This represents the most important common graphic output which ultimately can be compared across the three processes and for the data-exploration stage, it is illustrated in figure 5(d).

If the calibration stage is entered, the user embarks on a structured parameter search across the space of values, and a graph of the process is displayed in the top window of the calibration panel. This shows progress towards locating the best parameter value or, at the end of the process, the complete search. During this stage, the user has to grasp detailed numeric and textual data which are displayed in the text window as well as interpreting the graph of search illustrated at the top of the calibration panel. This is shown in figure 6 where I have also shown the final panel of maps for the data exploration, an enlarged graph of the observed trip-frequency distribution displayed from the earlier data exploration phase, and the graph of the search for the best parameter during the calibration phase. When the calibration is finished, the user can explore how good the fit is visually by displaying various maps of the outputs as shown in figures 7(a) and (b). Once this stage has been concluded, the calibration panel is filled in with three comparable maps to those in the exploration panel and these are shown in figure 8. The user then moves on to consider the final process—prediction.

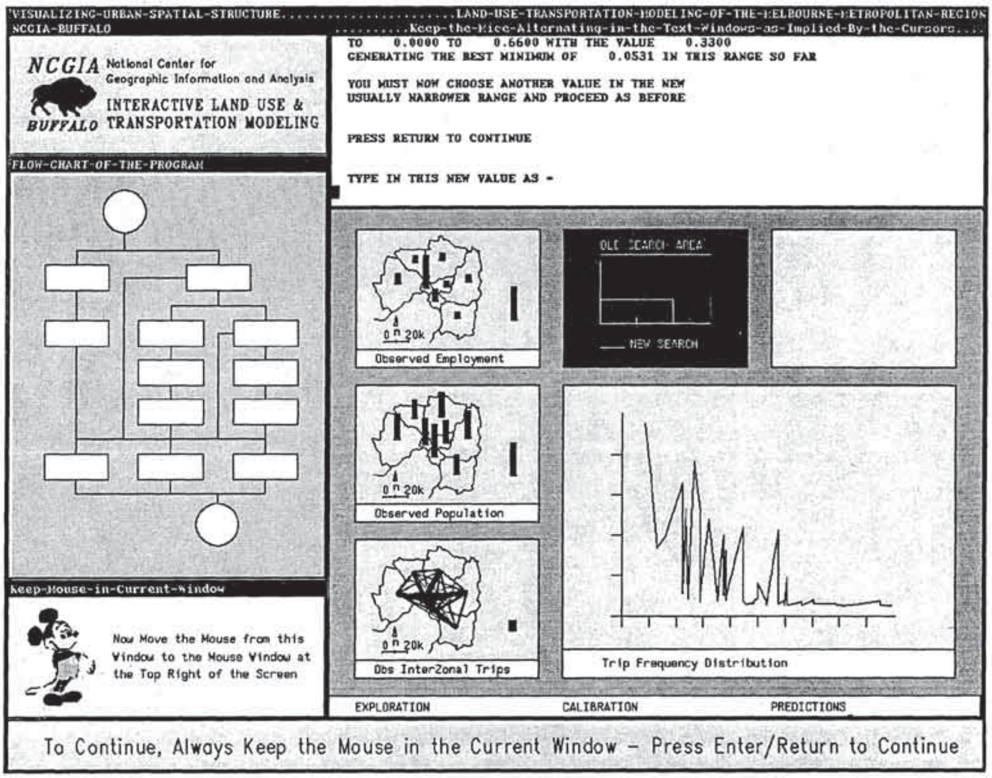
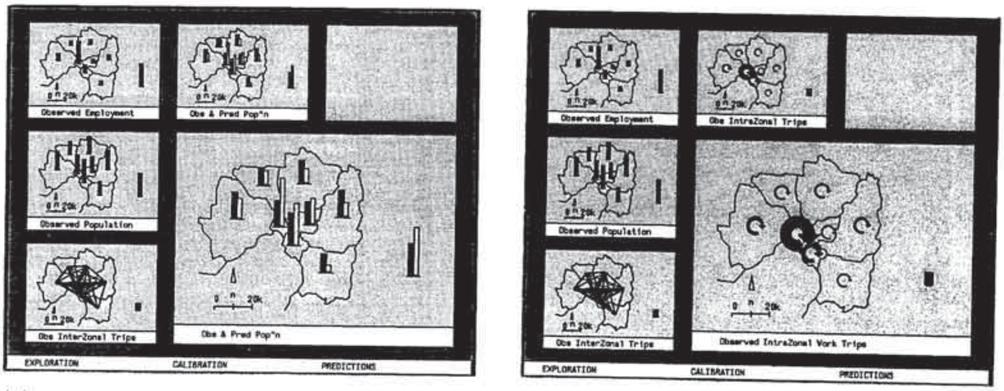


Figure 6. Progress in calibrating the model using dichotomous search.

The user is first prompted to engage in prediction or not. If not, the program finishes and the user can initiate a cascading collapse of the window system back to the initial Suntool. If the user wants to predict with the model, it is possible to alter any of the input data in a clear and structured way, thus enabling the user to test the impact of certain changes on the distribution of the residential population and the flow patterns associated with the journey to work. The same process as before is then followed for the display of spatial data. Figure 9 shows an example of the screen layout as it is in the predictive process where the user is examining the intrazonal flow of trips in aggregate and disaggregate form. Once the predictive



(a) (b)

Figure 7. Assessing the calibration: (a) locational bar graphs, (b) intrazonal trips.

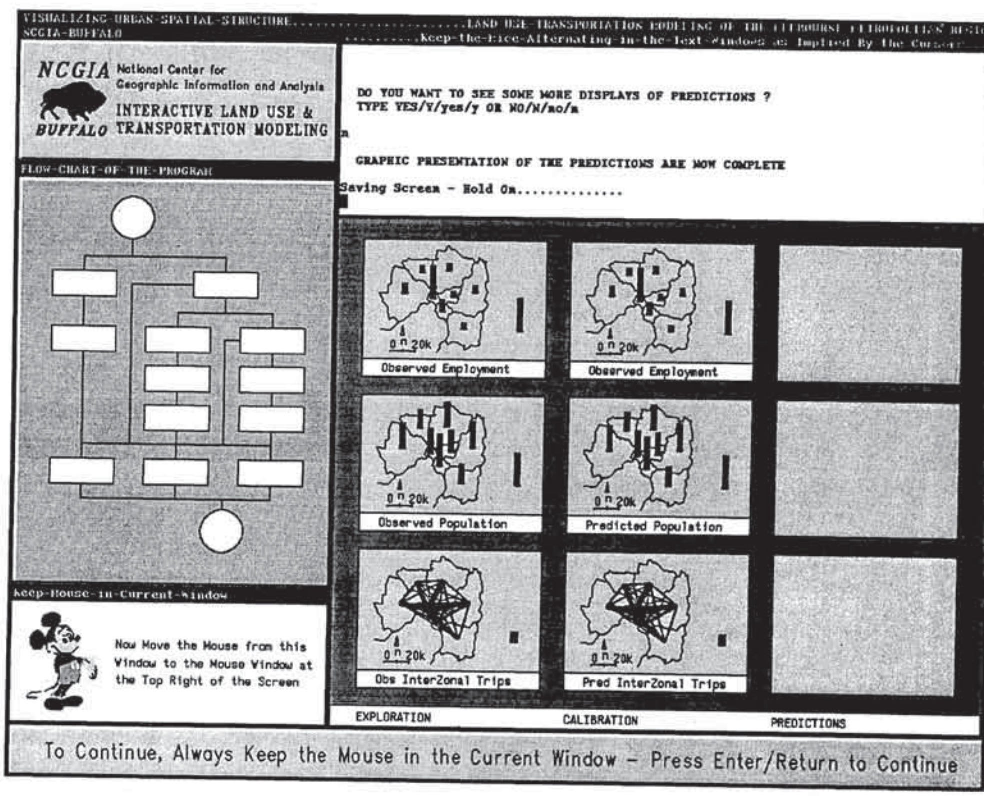


Figure 8. The completed exploration and calibration stages.

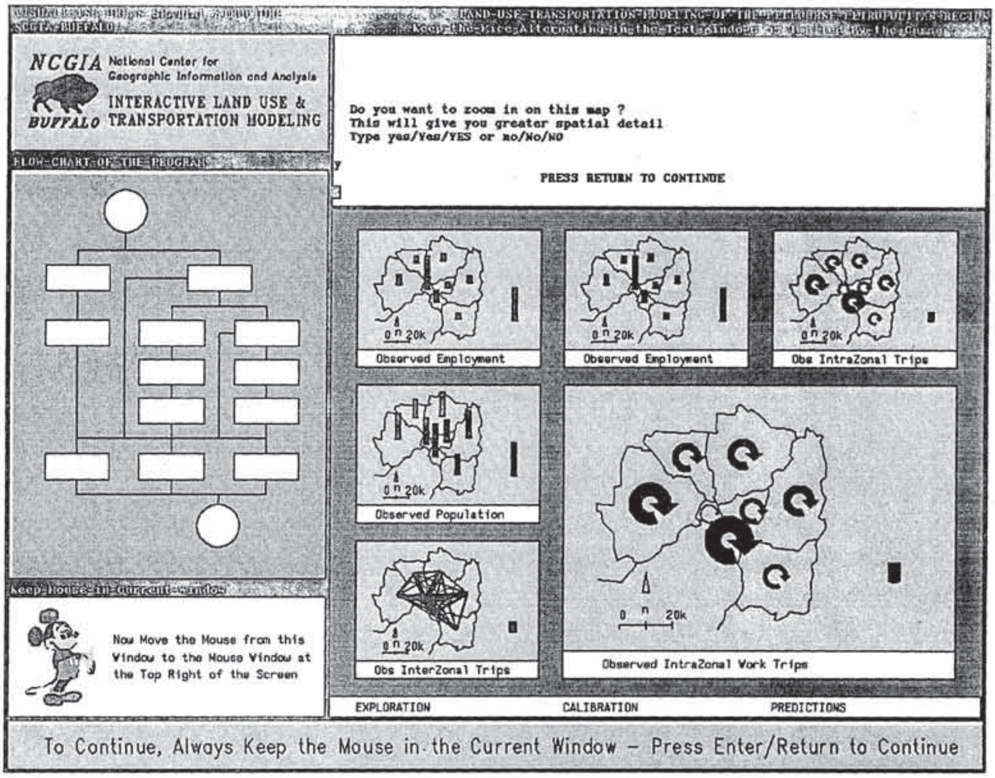


Figure 9. Assessing the predictions through intrazonal flows.

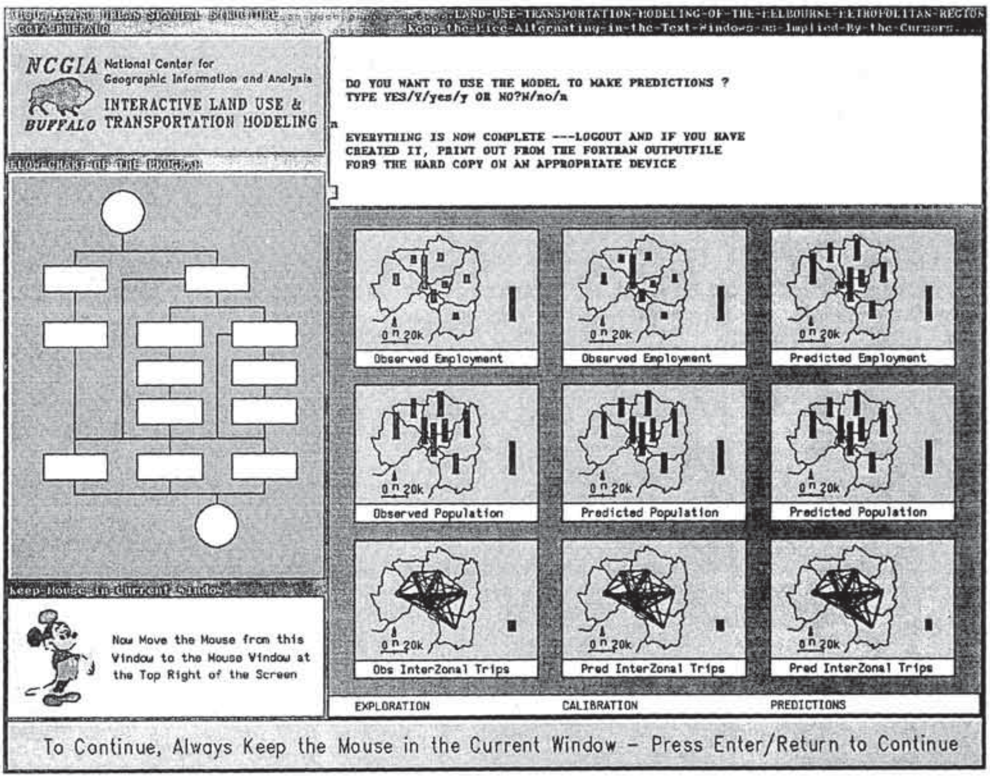


Figure 10. The completed run in terms of its visual trace.

stage ends then the third panel is filled out as shown in figure 10. The user can then continue to generate more predictions or can finish. A hard copy of the numeric, text, and statistical data can also be created by the user as the process proceeds and printed copy of the screen is available from screen dumps if a suitable printer is attached to the machine, either during the process or after it has been concluded. It is worth noting that, because the enlarged map window is created across the smaller panel windows, it is necessary in the calibration and prediction stages to dump the screen and then reload once the process moves on, so that the underlying graphics are preserved. Of course, it is possible to redraw the screen and with the small model developed here, this would take little time and even perhaps be faster than screen dumping and loading. But with the larger model envisaged, it is likely that such dumping and loading will be crucial to keeping up the speed of the application.

8 Conclusions

This project has only just begun and so far all that has been developed is a highly simplified prototype of a visual environment for urban modeling. Many of the features noted earlier which such an environment should contain have not yet been implemented and the model application to the Buffalo-Niagara region is in its early stages with the current application being the development of a pilot model for this urban region. In fact, this project covers many different objectives and there is always the danger that this plurality of goals might change the nature of the project as well as confuse its direction. The development of the visual environment described here is but one area of research where questions of design and the user interface are paramount.

Equally important is the development of the extended model itself which will incorporate many new developments based on new theories and models of urban structure which have been developed as the field has matured over the last decade. The incorporation of the optimization paradigm, for example, and the extended input-output model are areas where strong theory now exists but where applications are limited and the project will ultimately pursue these. There are problems posed by assembling data, particularly those arising from the need to stitch together databases from different countries, as well as the inevitable problems of making the temporal and spatial units consistent with one another. And there are also substantive issues involved in the application to Buffalo; for example, one of the critical issues in this urban region is the influence of Toronto and Southern Ontario on the local economy and the growing problems of many types of cross-border flow in the context of different pricing structures and costs of living as well as the existence of the Canada-US Free Trade Agreement.

All these diverse issues will not be incorporated in a single project and, at present, the visualization project is the one which holds the center stage, and even this is only beginning. There are several issues which are under active consideration, some of which are being slowly resolved at the time of writing. There is the issue of adding the fourth, prescriptive, process to the model and this is likely to be straightforward. There is a question of extending the modes-of-analysis windowing systems to encompass true aggregation, and we are already working with a small nineteen-zone spatial system based on the Buffalo-Niagara region at two levels of spatial aggregation. There is also the notion of extending the model to deal with two or more activity sectors and we are considering a pilot model of Buffalo incorporating the residential and retailing sectors in a Lowry-like framework.

There are many other issues relating to windowing which need to be resolved such as the construction of the model itself in modular fashion using Sunview windows to effect their linkage. And there is the need to develop better test screens and other forms of window which pop up and down in a somewhat more effective fashion than those we have at present. For example, we would like to develop the four processes of model building in such a way that we can iconify graphics as we use them, thus recalling them at any stage and making much better use of these as a visual trace of the model development process. All this is for the future but what is already clear is that the development of visual environments such as the one being pursued here is likely to become the rule rather than the exception in many, if not all, types of computer applications in the next decade.

Acknowledgements. This paper represents work that is related to research being carried out at the National Center for Geographic Information and Analysis, supported by a grant from the National Science Foundation (SES-88-10917). The Government has certain rights in this material. Support by NSF is gratefully acknowledged. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the NSF.

References

Anas A, 1982 *Residential Location Markets and Urban Transportation: Economic Theory, Econometrics and Policy Analysis with Discrete Choice Models* (Academic Press, New York)

Anas A, 1987 *Modeling in Urban and Regional Economics* (Harwood, New York)

Anselin L, 1988 *Spatial Econometrics: Methods and Models* (Kluwer, Dordrecht)

Batey P W J, Madden M, 1981, "Demographic-economic forecasting within an activity-commodity framework: some theoretical considerations and empirical results" *Environment and Planning A* **13** 1067-1083

Batty M, 1971, "Exploratory calibration of a retail location model using search by golden section" *Environment and Planning* **3** 411-432

Batty M, 1976 *Urban Modelling: Algorithms, Calibration, Predictions* (Cambridge University Press, Cambridge)

Batty M, 1978, "Urban models in the planning process", in *Geography and the Urban Environment: Volume I: Progress in Research and Applications* Eds D T Herbert, R J Johnston (John Wiley, Chichester, Sussex) pp 63-134

Batty M, 1983, "Linear urban models" *Paper of the Regional Science Association* **53** 5-25

Batty M, 1986, "Technical issues in urban model development: a review of linear and nonlinear model structures", in *Advances in Urban Systems Modelling* Eds B G Hutchinson, M Batty (North-Holland, Amsterdam) pp 133-162

Batty M, 1989, "Urban modelling and planning: reflections, retrodictions, and prescriptions", in *Remodelling Geography* Ed. B MacMillan (Basil Blackwell, Oxford) pp 147-169

Batty M, Yeh T, 1991, "The promise of expert systems for urban planning" *Computers, Environment and Urban Systems* **15** 101-108

Baxter R S, 1976 *Computer and Statistical Techniques for Planners* (Methuen, Andover, Hants)

Ben Akiva M, Lerman S R, 1985 *Discrete Choice Analysis: Applications to Travel Demand* (MIT Press, Cambridge, MA)

Bertuglia C S, Leonardi G, Wilson A G, 1990 *Urban Dynamics: Designing an Integrated Model* (Routledge, Chapman and Hall, Andover, Hants)

Birkin M, Clark G, Clarke M, Wilson A G, 1990, "Elements of a model-based geographic information system for the evaluation of urban policy", in *Geographic Information Systems: Developments and Applications* Ed. L Worrall (Belhaven Press, London) pp 132-162

Boyce D E, 1988, "Renaissance of large-scale models" *Papers of the Regional Science Association* **65** 1-10

Brail R K, 1990, "Integrating urban information systems and spatial models" *Environment and Planning B* **17** 417-427

Brand S, 1987 *The Media Lab: Inventing the Future at MIT* (Viking Penguin, New York)

Broadbent T A, 1973, "Activity analysis of spatial-allocation models" *Environment and Planning* **5** 673-691

Champernowne A F, Williams H C W L, Coelho J D, 1976, "Some comments on urban travel demand analysis, model calibration and the economic evaluation of transport plans" *Journal of Transport Economics and Policy* **10** 267-285

Couclelis H, 1991, "Geographically informed planning: requirements for planning-relevant GIS" *Papers in Regional Science* **70** 9-20

Creighton R L, 1970 *Urban Transportation Planning* (University of Illinois Press, Urbana, IL)
Cressie N, 1991 *Statistics for Spatial Data* (John Wiley, New York)
Dangermond J, 1990, "Geographic information system technology and development planning" *Regional Development Dialogue* **11** 1-14
de la Barra T, 1989 *Integrated Land Use Transport Modelling: Decision Chains and Hierarchies* (Cambridge University Press, Cambridge)
Densham P, 1991, "Spatial decision support systems", in *Geographical Information Systems: Principles and Applications* Eds D J Maguire, M F Goodchild, D W Rhind (Longman, Harlow, Essex) pp 403-412
Ding Y, Fotheringham A S, 1992, "The integration of spatial analysis and GIS" *Computers, Environment and Urban Systems* **16** 3-19
Echenique M, 1983, "The use of planning models in developing countries: some case studies", in *Urban and Regional Policy Analysis in Developing Countries* Eds L Chatterji, P Nijkamp (Gower, Aldershot, Hants) pp 115-158
Ferreira J, Evans D, 1992, "Sharing information in an imperfect world: understanding the interaction between technical and organizational issues", paper presented to the NCGIA Initiative 9: Institutions Sharing Geographic Information, San Diego, CA, 26-29 February; copy available from J Ferreira, Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge, MA
Forrester J W, 1969 *Urban Dynamics* (MIT Press, Cambridge, MA)
Fotheringham A S, O'Kelly M E, 1989 *Spatial Interaction Models: Formulations and Applications* (Kluwer, Dordrecht)
Fotheringham A S, Wong D, 1991, "The modifiable areal unit problem in multivariate statistical analysis" *Environment and Planning A* **20** 1025-1044
Haining R P, 1990 *Spatial Data Analysis in the Social and Environmental Sciences* (Cambridge University Press, Cambridge)
Harris B, 1967, "The city of the future: the problem of optimal design" *Papers of the Regional Science Association* **19** 185-195
Harris B, 1989, "Beyond geographic information systems: computers and the planning professional" *Journal of the American Planning Association* **55** 85-90
Harris B, Batty M, 1992, "Locational models, geographic information and planning support systems", TP 92-1, National Center for Geographic Information and Analysis, State University of New York, Buffalo, NY
Klosterman R, Brail R K, Bossard E G (Eds), 1992 *Spreadsheet Models for Urban and Regional Analysis* (Center for Urban Policy Research, Rutgers University, New Brunswick, NJ)
Lee D B, 1973, "Requiem for large-scale models" *Journal of the American Institute of Planners* **39** 163-178
Macgill S M, 1977, "The Lowry model as an input-output model and its extension to incorporate full intersectoral relations" *Regional Studies* **12** 337-354
Openshaw S, 1990, "Spatial analysis and geographical information systems: a review of progress and possibilities", in *Geographical Information Systems for Urban and Regional Planning* Eds H J Scholten, J C H Stillwell (Kluwer, Dordrecht) pp 153-163
Putman S H, 1983 *Integrated Urban Models* (Pion, London)
Putman S H, 1992 *Integrated Urban Models 2* (Pion, London)
Robinson V B, Coiner J C, 1986, "Characteristics and diffusion of a microcomputer geoprocessing system: the Urban Data Management Software (UDMS) package" *Computers, Environment and Urban Systems* **10** 165-173
Tukey J W, 1977 *Exploratory Data Analysis* (Addison Wesley, Reading, MA)
Walsh P K, Gibberd R W, 1980, "Development of an entropy model for residential location with maximum zonal population constraints" *Environment and Planning A* **12** 1253-1268
Webster F V, Bly P H, Paulley N J (Eds), 1988 *Urban Land-use and Transport Interaction: Policies and Models* (Avebury, Aldershot, Hants)
Wilson A G, Coelho J D, MacGill S M, Williams H C W L, 1981 *Optimization in Location and Transport Analysis* (John Wiley, Chichester, Sussex)

APPENDIX

The formal structure of the residential location model

The residential location model allocates the number of trips T_{ij} generated at employment zone i to residential zone j subject to two constraints on interaction which are specified as follows:

$$\sum_j T_{ij} = O_i \,, \tag{A1}$$

$$\sum_{i,j} T_{ij} c_{ij} = C \,, \tag{A2}$$

where O_i is the number of persons employed in zone i, c_{ij} is the travel cost between zone i and zone j, and C is the total cost of travel expended by all workers traveling between work and home in the system. The most probable model satisfying the origin and travel cost constraints in equations (A1) and (A2) can be derived by maximizing the conditional entropy S which is defined as

$$\underset{\{T_{ij}\}}{\text{maximize}}\; S = -\sum_{i,j} T_{ij} \ln\left(\frac{T_{ij}}{D_j}\right), \tag{A3}$$

where D_j is a measure of the attractiveness of the residential zone j which acts as the destination of the trips T_{ij}.

Equation (A3) is maximized subject to the constraints by forming a Lagrangian $\mathscr{L}$ defined as

$$\underset{\{T_{ij}\}}{\text{maximize}}\; \mathscr{L} = \underset{\{T_{ij}\}}{\text{maximize}} S + \gamma_i \left(\sum_j T_{ij} - O_i\right) + \beta \left(\sum_{i,j} T_{ij} c_{ij} - C\right), \tag{A4}$$

where γ_i and β are multipliers ensuring that the origin constraints (A1) and the travel cost constraint (A2) are satisfied. The model which is derived from this maximization is given as

$$T_{ij} = D_j \exp(\lambda_i + \beta c_{ij}) \,, \tag{A5}$$

where λ_i is the augmented multiplier, $\gamma_i - 1$, formed during the maximization. This model can be solved in several ways. At present, it is possible to solve for the parameters λ_i and β in two ways: first by explicitly calculating the factor λ_i by balancing equation (A1) and then iteratively converging on the value of β which ensures equation (A2) is satisfied using dichotomous search by golden section (Batty, 1971); and second by actually maximizing the entropy in equation (A3) subject to the two constraints using a penalty function method (Wilson et al, 1981).

The dual of this entropy maximization is a minimization problem given as

$$\underset{\{T_{ij}, \gamma_i, \beta\}}{\text{minimize}}\; D = \sum_{i,j} T_{ij} + \sum_i O_i \gamma_i + \beta C \,, \tag{A6}$$

which is subject to the following constraint equation

$$\ln\left(\frac{T_{ij}}{D_j}\right) - \gamma_i - \beta c_{ij} = 0 \,. \tag{A7}$$

As yet, the dual formulation has not been programmed as an option but it will be explored in the extended model when the environment is developed to embrace prescriptive usage of the model. The program also has an option of including inequality constraints on the location of trips at destinations. These constraints are formulated as

$$\sum_i T_{ij} \leqslant Z_j \,, \tag{A8}$$

where Z_j is the capacity of the residential zone j. If the model is run in this form, then slack variables and appropriately formed multipliers appear in the maximization and minimization problems. If these slacks take positive values, then this means that the destination constraints are operative in the solution.