GIS SUPPORT FOR MICRO-MACRO SPATIAL MODELING

Timothy L. Nyerges
Assistant Professor
Department of Geography
Smith Hall DP-10
University of Washington

ABSTRACT

Concepts, techniques and tools in geographic information systems (GIS) have come a long way primarily because of a) database management support for spatial data bases, b) spatial operators for point, line and area features - the current mainstay in spatial analysis as related to GIS, and c) graphics display. For years spatial modelers in human geography have been developing concepts and techniques for information analysis, however these developments have seldom found their way into a relationship with GIS. This paper indirectly explores reasons why this relationship has not developed faster to support model-based spatial analysis, given the length of time spatial modeling and GISs have been developing by examining the interrelationships between GIS and transportation modeling. Central to this problem has been the intended depth and narrow focus of spatial modeling applications versus the shallow breadth of GIS information management and analysis. GIS issues are presented in terms of system functionality. Spatial modeling issues are presented in terms of 'basic dimensions' in the context of spatial transportation modeling. The relationship between the two topics is explored by identifying GIS functions that can support various 'basic dimensions of modeling' in a transportation modeling context.

INTRODUCTION

Is it true that "(g)eographical information systems and model-based locational analysis: (are) ships (passing) in the night or the beginnings of a relationship?". (Birkin, Clarke, Clarke, and Wilson 1987 p 1). Birkin and others raise that question concerning the relationship between geographic information systems (GIS) and locational analysis in geography, but it could be considered for all socio-economic oriented modeling in geography including transportation modeling as well. They argue that "GISs and models in geography have largely grown through different cultures and traditions. Both schools have achieved obvious successes, but both have the same kinds of difficulty when it comes to practical application " (Birkin, Clarke, Clarke and Wilson 1987,1).

Birkin and others (1987) suggest that the history of GIS is tied rather closely to remote sensing (and by implication environmental sciences), and therefore GIS is inadequate when it comes to supporting social science studies. In addition, the authors claim that model building in academic geography has suffered from a lack of practical context, and therefore argue for a more "applicable human geography" (Clarke and Wilson 1987). The same case has recently been made in the *Annals* in an article titled "Geography Beyond the Ivory Tower" (Demko 1988). Calling for a better integration of theory and practice is by no means new, having been made over the years and in every academic field. But what is different now in geography is that GISs are being adopted rapidly by practitioners in all segments of society. Perhaps conceptual advances with GIS in support of social science studies can help further advances in model building in human geography.

The goal in both GIS and model building environments has been the production of "information" from raw data, with the definition of "information" being its "usefulness" in a decision-making process. How this is done is a matter of the traditions and methodologies of the different schools.

Nyerges and Dueker (1988) discuss the spatial analysis capabilities in GISs oriented to transportation applications in terms of spatial operators and model-based software for either socio-economic or environmental problems. Spatial operators are currently the mainstay of the GIS spatial analysis due to the fact that most modeling environments are applications specific. Model-based software modules have tended to be rather large in many instances and therefore not easily absorbed into GIS environments or the expertise to perform such interfacing has not been available. Only a few implementations of socio-economic models in a software environment with some GIS-like capabilities have been reported (Babin, Florian, James-Lefebure and Spiess 1982; de la Barra, Perez and Vera 1984; Birkin, Clarke, Clarke and Wilson 1987). These latter implementations point out the need for a better understanding of how GIS functionality can support socio-economic, model-based analysis.

The relationship between GIS and socio-economic, model-based analysis in geography needs to be developed further in terms other than just an occasional interfacing of software, if the two topics are to experience mutual benefits over a long term. The fundamental basis of integration/interfacing is in need of further exploration.

This paper indirectly explores reasons why the relationship between GIS and socio-economic modeling has not developed faster to support model-based spatial analysis, given the length of time spatial modeling and GISs have been developing. Central to this problem has been the intended depth and narrow focus of spatial modeling applications versus the shallow breadth of GIS information management and analysis. However, we can look upon this problem as an opportunity to better understand the foci of both GIS and spatial modeling if we contrast and compare the fundamental characteristics of each. That is, the functionality in a generic GIS and the 'basic dimensions of spatial modeling' in human geography. The context of this research is an exploration of GIS support for micro-macro spatial transportation modeling. In particular, travel demand forecasting models for state and urban transportation planning are being integrated/interfaced with a GIS environment.

GIS FUNCTIONALITY AND SPATIAL TRANSPORTATION MODELING

Transportation GIS Functionality

Several definitions for GIS have been proposed over the years, with perhaps as many definitions as there are view points on the subject. Cowen (1988) summarizes many of those definitions and identifies four basic approaches to defining a GIS. These approaches are: (a) a process-oriented approach in terms of input, storage, retrieval, analysis, and output, (b) an application which categorizes a GIS according to the type of data, (c) a toolbox approach incorporating a sophisticated set of computer-based procedures and algorithms for handling spatial data, and (d) a database approach as a refinement of the toolbox approach and focusing on the data retrieval performance of the system rather than the functionality. Cowen (1988, 1554) suggests that none of those approaches are suitable for a definition and concludes that "a GIS is best defined as a decision support system involving the integration of spatially referenced data in a problem solving environment." The focus is on decision making and problem solving. We can help clarify this GIS definition by turning to Bonczek, Holsapple and Whinston (1981, 3) who define a decision support system as "an information processing system that is embedded within a decision making system" whereby the decision making system might be human, machine or human-machine.

The definition of a GIS used here is most like that of the decision support orientation as favored in Cowen (1988). However, the modeling capability is left to functions that provide 'modeling support' rather than a specific 'modeling application'. The intent here is to identify functions in a GIS which can support modeling, and leave the system implementation as whether models are interfaced or integrated up to the system specific environment. A list of potential functions that might exist in a transportation oriented GIS appear in Table 1. All of these functions in some way support a modeling environment, but the interest here is in those that most directly support such an environment. The next section presents a list of 'basic dimensions of spatial modeling' to outline the issues that one considers when building a spatial-oriented socio-economic model, and in particular transportation models.

Table 1. Functionality in a Transportation GIS

Human interface

- Non-programming interface: an interface to support retrieval and analysis for both casual users such as menus and expert users such as a command language.
- Programmer's interface: a subroutine library interface to support programmer use for extending and creating applications.

Data capture and editing

- Digitizing: manual or automatic digitizing.
- Data validation: integrity constraints for data quality checks such as detection of 'overshoot' digitizing errors.
- Import/Export: ability to load in (import) or send out (export) bulk spatial and attribute digital data.
- Interactive editing: capability for a user to add/delete objects/data values one at a time with the use of retrieval criteria at the option of the user.
- Batch editing: capability for a user to add/delete objects/data values in bulk processing.
- Map edge match: matching the edges of maps by selecting a center line having a band width.

Spatial and thematic data management

- Map area storage/retrieval: continuous geographic domain for any area or group of areas can be stored and retrieved as a single database.
- Spatial data description: construction of point features, link/node topology with shape records, chain encoding of polygon boundaries.
- Locational reference: use of absolute referencing such as latitude/longitude, state plane or UTM coordinate reference system or relative referencing such as route-milepoint, address or other specialized system depending on problem orientation.
- Global topology: global network topology for any geographic domain and any set of data categories to be defined by system administrator at the request of users.
- Thematic data description: construction of attribute fields to qualitatively and quantitatively describe data categories.
- Data definition: software to manage descriptions and definitions of the data categories and the spatial and thematic data descriptors of these categories.
- Spatial selective retrieval: retrieval of data based on spatial criteria such as coordinate window, route-milepoint reference.
- Thematic selective retrieval: retrieval of data based on thematic criteria such as name of road, attribute of road.
- Browse facility: retrieval of any and all data categories.
- Access and Security: multi-user or single user access with read/write protection.
- Roll-back facility: supports restoration of database state in the event of system failure.
- Minimal data redundancy.
- Subschema capability: select parts of a corporate-wide database for special management.

- Database size: No limitation on the number of points, lines or areas per map, maps per data base, or coordinates per line or area should exist for logical storage of elements within the capacity of physical storage.

Data manipulation

- Structure conversion: conversion of vector to raster, quadtrees to vector.
- Object conversion: point, line, area, cell, or attribute conversion to point, line , area, cell, or attribute.
- Coordinate conversion: map registration, 'rubber sheet' transformations, translation, rotation, scaling, map projection change or image warping.
- Locational classification: grouping of data values to summarize the location of an object such as calculations of area centroids, proximal features, Thiessen polygons.
- Thematic classification: grouping of thematic data values into classes.
- Locational simplification: coordinate thinning of lines.
- Locational aggregation: grouping of spatial objects into a superordinate object.
- Class generalization: grouping data categories into the same class based on characteristics of those categories.
- Thematic aggregation: creation of a superordinate object based on thematic characteristics of two or more other objects.

Data analysis

- Spatial object measurement: individual object and interobject calculations for line length, area and volume, distance and direction.
- Statistical analysis: frequency analysis, measures of dispersions, measures of central tendency (mean, median, mode), correlation, regression.
- Spatial operators: point, line, area object on/in point, line, area object; gravity model primitives; network indices for beta, diameter and accessibility.
- Routing: identify routes based on spatial or thematic criteria.
- Model structuring: a model structuring environment that provides linkages between parts of models perhaps through a special language.

Data display

- Symbolization change: any graphic symbolization could be created at the option of the user.
- Softcopy graphics: viewing of maps, graphs on CRT monitor.
- Hardcopy graphics: maps, graphs on printer/plotter.
- Reports: reports on content of database, report formatting to support analysis, formatted summary tables.
- Display window: the area of the database currently being examined.
- Overview window: a window used for quick spatial orientation that shows the entire geographic domain.
- Pan: the ability to roam across a geographic domain bringing data to the CRT screen without having to change the display window.
- Zoom: changing the area of the display window to examine more or fewer features, resulting in a change of scale of

the display image. A change in accuracy usually does not occur.

Note: The list of functions is a combination of a list appearing in Table 1 of (Rhind and Green 1988) and a list appearing in a bid specification created with the assistance of the author (State of Alaska 1986). The list as prepared by Rhind and Green (1988) is a synthesis of several authors' works. Several of the classes as appear in (Rhind and Green 1988) have been reorganized and further elaborated based on the authors experience in writing system specifications.

Dimensions of Spatial Transportation Modeling

Models in human geography as related to transportation systems have been a topic of interest for many years resulting in too many models to mention in detail. Good summaries of the breadth of the field appear elsewhere such as (Chorley and Haggett 1967, Wilson and Bennett 1985). Even these extensive presentations do not completely describe the effort as suggested in the review of (Wilson and Bennett 1985) as provided by MacKinnon (1987).

Focus in this research involves transportation modeling which is part of an urban transportation planning process. Meyer and Miller (1984) describe the urban transportation planning process as consisting of the following decision-oriented steps:

1. Diagnosis and data management
2. Analysis and evaluation
3. Scheduling and budgeting
4. Monitoring

Since the mainstay of GIS is data management and graphics it is natural that part of current and future GIS research be involved with data management and other functionality suitable to support analysis for certain application areas, thereby systematically linking the above mentioned Steps 1 and 2. Transportation analysis and evaluation is one of the most complex and lengthy portions of the transportation planning process. A majority of the analysis process involves transportation modeling. According to Werner (1985) the spatial transportation modeling process is made up of the following steps:

1. **Estimate trip generation characteristics.** Inflows and outflows from each traffic analysis zone (TAZ) must be estimated at an appropriate disaggregated scale using trip generation models.

2. **Compute trip distributions.** Characterize the origins and destinations of the TAZs and balance these to conserve the total number of trips across the transportation area using trip distribution models such as growth factor, intervening opportunities or gravity models.

3. **Estimate the modal split.** Determine the proportions of the different modes of transportation that apply to each of the origins and destinations using modal choice models.

4. **Assign trips to routes.** All trips must be assigned to routes to load the network using network loading models.

5. **Evaluate alternative network structure and use.** Alternative network designs and the operation of these networks according to vehicle access are examined using models based on dynamic programming (shortest path) and linear programming (optimal use) algorithms.

A systematic approach to describing the idiosyncrasies of spatial transportation models can start with a set of basic dimensions of any human-geography related modeling task as identified by Wilson (1987, 414): entitation, scale, spatial representation, partial versus comprehensive, static analysis and dynamic analysis. These basic dimensions identify the issues and considerations that need to be addressed when developing (or in this case interfacing or integrating) models to a GIS. Consequently, the models in the above steps could be described according to the six dimensions; each description being used to identify GIS functionality to support those models.

The six dimensions that characterize the basic considerations for building spatial-oriented models in human geography are describe in more detail as follows:

1. **Entitation** - Enumerate the basic components of the system of interest, e.g. what are the important entity types (categories of information) of a transportation system (Wilson 1981) such as transportation routes (rivers and freeways), routes, vehicles (multimodal - cars, buses, trains etc.).

2. **Scale** - What level of resolution should be adopted, e.g. should all entity types be populated with data to a level such that all vehicles and all roadways are considered in the analysis?

3. **Spatial representation** - Determine how to treat space. Treat space in a continuous way, so that exact locations of activities can be determined as in the case with every origin and destination of a trip. Treat space in a discreet way such that zonal boundaries are more appropriate as the spatial aggregations or origins and destinations.

4. **Partial versus comprehensive** - Are the location of any marginal activity considered in the analysis, for example a firm or a household, given the rest of the system as an environment or are all activities considered simultaneously, thus reflecting competitive processes?

5. **Static analysis** - What theory and methods are used for the static analysis of structure and form.

6. **Dynamic analysis** - What theory and methods are used for the dynamic analysis of evolution and change.

Wilson (1987) offers several observations clarifying these dimensions. The first four dimensions are of a different character than the last two. The first four determine the way we look at the system of interest for analytical purposes; the last two relate to detailed specification of the theories and methods.

Differences in creating a model can arise because of dimensional choices for data representation even though the theoretical basis may be the same. For example, continuous space economic models and discrete space economic models look different because the constructs used to create the model are different, even though their theoretical underpinnings are still the same. Continuous space economic models are much more likely than discrete zone models to necessitate restrictive assumptions such as "all employment at the city center" inorder to make the mathematics practical. These differences result in the need for various functional requirements that match a particular modeling approach.

Wilson (1987) recognizes the importance for conditional forecasting that will remain useful for short-run planning, but emphasizes the need for a shift in perspective from long-run comprehensive, impact analyses which has not been very productive (Lee 1973) to analyses of stability and resilience of components in urban systems. Despite the differences in short-run partial versus long-run comprehensive approaches to modeling, Meyer and Miller (1984, 182) identify several issues that are common to both the long-run and short-run modeling tasks:

1. the dynamic nature of an urban area
2. the complexity of urban behavior
3. the need for good quality, detailed data

Providing support for the integration of micro-macro spatial transportation models that can address short-run and long-run issues at local and regional scales in a convenient and systematic way will test the capabilities of current and future GISs. In essence, a micro-macro approach is an attempt to build structured, comprehensive models while taking into consideration data availability constraints in the short-run, addressing these constraints and associated data requirements with GIS support.

Birkin and Clarke (1986) discuss micro-macro approaches to dynamic urban modeling. An example of a micro approach is a micro-simulation model of urban travel behavior. Such a model captures the decision making process of an urban traveler throughout the daily cycle of activities over the urban landscape. A macro approach uses trip data aggregated to origin and destination places represented as different traffic analysis zones (TAZs). The macro approach uses an accounting method for balancing all of the trips generated from origin zones to destination zones with the return trips from the destination zones back to the origin zones (Werner 1985).

GIS SUPPORT FOR TRANSPORTATION MODELING

No single GIS specification is likely to address all geographical applications, as no single GIS specification is likely to address all modeling applications, including transportation modeling. However, as a beginning to a systematic exploration of what GIS functions best support a modeling environment, the basic set of GIS functions outlined in Section 2.1 can be compared against the basic dimensions of modeling as described in the context of transportation modeling in Section 2.2. Each GIS function can be examined in turn as potentially providing some support to a transportation modeling environment, or alternatively each dimension can be considered as to how each function might support that dimension. The results are shown in Table 2.

Table 2. GIS Functionality for Transportation Modeling

	MODELING DIMENSIONS					
	1	2	3	4	5	6
GIS FUNCTIONALITY						
Human interface						
Non-programming interface	x	x	x	x		
Programmer's interface	x	x	x	x		
Spatial and thematic data management						
Map area storage/retrieval				x		
Spatial data description			x			
Locational reference			x			
Global topology				x		
Thematic data description	x					
Data definition	x					
Spatial selective retrieval			x			
Thematic selective retrieval	x					
Browse facility				x		
Access and Security				x		
Roll-back facility						
Minimal data redundancy				x		
Subschema capability				x		
Database size				x		
Data manipulations						
Structure conversion	x		x			
Object conversion		x				
Coordinate conversion			x			
Locational classification		x	x			
Thematic classification	x	x				
Locational simplification		x	x			
Locational aggregation		x	x			
Class generalization	x					
Thematic aggregation	x					
Data analysis						
Spatial object measurement			x			
Statistical analysis				x		
Spatial operators			x	x		
Routing				x		
Model structuring	x		x	x		

Some functionality is more directly related to the modeling interface/integration than is other functionality. The human interface, data management, data manipulation and data analysis functionality categories are directly related to a modeling environment. The data capture and data display functionality categories indirectly support integration/interfacing and are not considered in this evaluation (although this is the subject of related research).

For brevity, all models mentioned in Section 2.2 are taken as a general category, therefore only a single column exists as it applies to all models. A more detailed analysis would include a separate column for each model on each dimension. In addition, the dimensions for static analysis and dynamic analysis are not considered in this evaluation, since very little (if any) GIS functionality can truly support dynamic analysis to an extent of other than static time slices. Currently "time" in most data bases is represented as a thematic attribute. Such an approach is in need of further conceptual development (Langran and Chrisman 1988).

The following generalizations can be made with regard to human interface, data management, manipulation and analysis functionality for support of transportation modeling. Since entitation is a topic that involves the basic substance of what phenomena are being modeled, the GIS functionality that involves thematic attribute management - either naming them or deriving their descriptions from other data - is fundamental to developing the basic set of entities to be used in the modeling process. The spatial data management of locational data is fundamental to the spatial representation dimension, but from a static analysis point of view. That is, data management is concerned with the storing and retrieving of static descriptions. Current GISs would have difficulty with the dynamic representation of entity classes and/or their spatial representations.

Data manipulation functionality involves the conversion of one data object into another data object to support a particular scale of analysis. The scale of analysis dictates an appropriate data resolution for the data classes (determined through entitation). Data resolution indicates how general or detailed the data values are for the sampled entities. Different levels of detail can be supported through forward and reverse generalization, classification, aggregation and simplification processes of the data objects. This data manipulation supports changes in scale for a micro-macro approach to modeling.

The micro-macro approach involves more than a matter of spatial scale, but also is concerned with what can be called thematic scale and temporal scale. Map scale (as a term that is defined more precisely than spatial scale) is a term that refers to the relationship between distance on the ground and distance on a map. Hence spatial scale refers to how much distance, or what area is included, focusing on spatial position. The term scale includes measurements for both resolution and accuracy. That is, the units of resolution for knowing how much distance or area is involved together with whether the measurement can be replicated with confidence is important. Thematic scale involves the degree of specificity for any given attribute and whether the measurement is accurate. The same would apply to time, as in the case of "when something occurred". The temporal characteristics of geographical information is a topic of current research (Langran and Chrisman 1988) in a GIS context.

Data analysis functionality involves basic spatial operators that are application specific, basic statistical analyses and spatial object measurement. The analysis functionality would primarily support a partial versus comprehensive approach to building models. Spatial object measurements could be made for the raw data to be submitted for statistical analyses. Statistical analyses for central tendencies and regression could support parameter estimation for the primitives of models. The spatial operators might include gravity model operators that are the primitives used to construct more sophisticated spatial interaction models. Spatial operators might also include network indices, such as beta diameter and accessibility indices, for describing the complexity of a network, hence being used to decompose the network into simpler terms for quicker analysis at a general level, then more detailed analysis at a local level.

SUMMARY and CONCLUSIONS

This paper started out by asking the question: "After years of development for both spatial modeling and GIS, are these developments passing as ships in the night or is there a relationship beginning to emerge?" If we assume the latter to be occurring; it will only pass a superficial level by taking a more systematic cross-fertilization of the two areas of expertise.

GIS solutions for many geographical problems are now being investigated. One of these is in transportation applications. Transportation modeling is beginning to come out of an experimental realm and is being put to use in several municipalities across the United States.

Several issues about GIS functionality to support spatial transportation modeling are still unresolved. The evaluation presented here is preliminary, and would most likely differ from one modeling problem to another. However, an overall indication of the nature of GIS and modeling functionality can be assessed using this approach.

Further refinement is needed of the approach used in identifying GIS functionality that can support transportation modeling. Rather than indicating "direct" or "indirect" support or a "yes" or "no" level of support, the level of support could be indicated along a scale and prioritized for certain applications. This can be done through a more thorough examination of particular transportation models. This topic is currently under investigation through an examination of the use of TIGER/Line files as a base network for transportation modeling in the State of Washington.

REFERENCES

Babin, A., M. Florian, L. James-Lefebure and H. Spiess 1982. EMME/2: Interactive graphic method for road and transit planning, *Transportation Research Record*, 866, 1-9.

Birkin, M. and M. Clarke 1986. Comprehensive dynamic urban models: integrating macro- and micro- approaches, Working Paper 439, School of Geography, University of Leeds.

Birkin, M., G. P. Clarke, M. Clarke, and A. G. Wilson 1987. Geographical information systems and model-based locational analysis: ships in the night or the beginnings of a relationship?, Working Paper 498, School of Geography, University of Leeds.

Bonczek, R. H., C. W. Holsapple and A. B. Whinston 1981. *Foundations of Decision Support Systems*, New York : Academic Press.

Chorley, R. J. and P. Haggett 1967. *Models in Geography*, London : Methuen.

Clarke, M. and A. G. Wilson 1987. Towards an applicable human geography: some developments and observations, *Environment and Planning A*, 19:1525-1541.

Cowen, D. J. 1988. GIS versus CAD versus DBMS: What are the Differences?, *Photogrammetric Engineering and Remote Sensing*, LIV(11):1551-1555.

de la Barra, T. B. Perez and N. Vera 1984. TRANUS-J: putting large models into small computers, *Environment and Planning B*, 11(1):87-102.

Demko, G. J. 1988. Geography beyond the ivory tower, *Annals* of the Association of American Geographers, 78(4):575-579.

Langran, G. and N. R. Chrisman 1988. A Framework for Temporal Geographic Information, *Cartographica*, 25(4): in press.

Lee, D. A. 1973. Requiem for large-scale models, *Journal of the American Institute of Planners*, 39(3):163-178.

MacKinnon, R. D. 1987. Review of *Mathematical methods in human geography and planning*, by A. G. Wilson and R. J. Bennett 1985.

Meyer, M. D. and E. J. Miller 1984. *Urban Transportation Planning*, New York : McGraw-Hill.

Nyerges, T. L. and K. J. Dueker 1988. Geographic information systems in transportation, Technical Report, U. S. Dept. of Transportation, Federal Highway Administration Planning Division, also included in Proceedings of the Computer-assisted Cartography in Transportation Conference.

Rhind, D. W. and P. A. Green 1988. Design of a geographical information system for a heterogeneous scientific community, *International Journal of Geographical Information Systems*, 2(2):171-189..

State of Alaska 1986. Bid Specification for a Digital Mapping System for the State of Alaska Department of Transportation and Public Facilities. Department of Administration, Bid No. 11451, Juneau, Alaska.

Werner, C. 1985. *Spatial Transportation Modeling*. Beverly Hills, CA: SAGE Publications.

Wilson, A. G. 1981. *Geography and the Environment: Systems Analytical Methods*. Chichester: John Wiley & Sons.

Wilson, A. G. 1987. Transport, location and spatial systems: planning with spatial interaction models, Chapter 7 in *Urban Systems: Contemporary Approaches to Modeling* (eds.) C. S. Bertuglia, G. Leonardi, S. Occelli, G. A. Rabino, R. Tadei, and A. G. Wilson, London: Croom Helm.

Wilson, A. G. and R. J. Bennett 1985. *Mathematical Methods in Human Geography and Planning*. Chichester, Sussex : John Wiley & Sons.