KUNGL. TEKNISKA HÖGSKOLAN
Royal Institute of Technology

DEPARTMENT OF
INFRASTRUCTURE AND PLANNING

TRITA-IP FR 01-100
ISSN 1104-683X
ISRN KTH/IP/FR--01/100--SE

REPORT

GIS in Transport Modelling

Svante Berglund

Division of Transport and Location Analysis

GIS in Transport Modelling

by

Svante Berglund

November 2001

Royal Institute of Technology
Department of Infrastructure
SE–100 44 Stockholm, Sweden

Set in LaTeX using EmTeX

Akademisk avhandling som med tillstånd av Kungliga Tekniska Högskolan i Stockholm framlägges till offentlig granskning för avläggande av teknologie doktorsexamen fredagen den 30 november 2001 kl 13.00 i sal Q 1, Kungliga Tekniska Högskolan, Osquldas väg 4, Stockholm.

Högskoletryckeriet, Stockholm 2001

GIS in Transport Modelling, Berglund S.
Department of Infrastructure
Royal Institute of Technology, Stockholm, Sweden

Abstract
This dissertation consists of five papers and an introduction. In *Barriers in travel models*, the focus is on modelling mobility in border regions, both borders within nations and between nations. We use travel surveys and networks from Sweden and Norway to model the barrier effect caused by the borders in the region between two counties on each side of the national border.

In *Identifying local spatial association in flow data* we develop a statistic for spatial association for flow data by generalising the statistic by Getis and Ord (1992) and Ord and Getis (1995), G_i and G_i^*. This local measure of spatial association, G_{ij}, is associated with each origin-destination pair and can thus be used for explorative analysis of e.g. matrices of travel flow. In *Detecting barrier effects* we use a modified G_{ij}-statistic, $G(\Gamma)$, in an attempt to identify barrier effects in flow data. The $G(\Gamma)$ approach for studying barrier effects is compared with the traditional approach of using dummy variables to model the barrier as in *Barriers in travel models.* We conclude that the $G(\Gamma)$ is more flexible than the usual dummy variable approach in an explorative spatial data analysis.

One central evaluation criterion of the transport system is accessibility. In *Path based accessibility* we develop an aggregate accessibility measure that takes one mandatory activity pattern into account i.e. travel to work trips. This measure can be used to evaluate accessibility obtained from multi-purpose trips.

In *Experiences from GIS based travel demand modelling* we summarise our experiences from different types of GIS–model integration. We present a model system that consists of tools for zone design, model estimation, scenario calculation and accessibility analysis. When using GIS as a modelling environment, we have not encountered any limitations with regard to model implementation and design of the user interface. Furthermore we found the GIS environment excellent for data preparation, accessibility analysis and visualisation of model output.

Keywords: transport modelling, GIS, accessibility, local spatial association, flow data, barriers, travel demand modelling

Preface and acknowledgement

All you need to write a dissertation are good supervisors, financial support, inspiring colleagues and an understanding family. First I would like to thank my supervisors Lars Lundqvist and Lars-Göran Mattsson for all their comments, their support and patience. During the work with Papers 2 and 3 I had the pleasure to work with Anders Karlström which was very inspiring, fruitful and fun. Financial support was provided by the Swedish Institute for Regional Research (SIR) and The Swedish Communications Research Board (KFB). Other important contributors were The Swedish National Road Administration and Helena Braun at SIKA who supported us with data.

During the process of writing this dissertation I have received many constructive comments during informal discussions and formal seminars. I would like to thank Dr. Jonas Eliasson, professor Lars Westin and specially Dr. Siv Schéele for comments on earlier versions of this dissertation. I also would like to thank Siamak Baradaran for comments on GIS related issues, Tom Petersen for comments on Paper 1 and Katja Berdica for help on volume delay functions during the preparation of Paper 5.

A family is not necessary in order to write a thesis, but if you have one it must be patient and understanding. Thank you Carin. Finally I would like to thank my daughter Malva for coming at the right moment (a few months before finishing this dissertation) and made me realise that what I was doing was not so important after all.

Contents

List of papers

1 Berglund, S. and Lundqvist, L. (1998) "Barriers in travel models", paper presented at the 37^{th} Annual Meeting of WRSA, in Monterey.

2 Berglund, S. and Karlström, A. (1999) "Identifying local spatial association in flow data", *Journal of Geographical Systems,* 1:219-236.

3 Berglund, S. and Karlström A. (2001) "Detecting barrier effects".

4 Berglund, S., (forthcoming) "Path based accessibility", *Journal of Transportation and Statistics.*

5 Berglund, S. (2001) "Experiences from GIS based travel demand modelling".

GIS in Transport Modelling
an Introduction

Svante Berglund

Introduction

The common theme of this dissertation is the use of Geographical Information Systems (GIS) in transport modelling; not in transport analysis or transport planning but just in modelling. Planning and analysis are wide concepts that cover a variety of disciplines and techniques, while modelling is a much more narrow concept. The focus is on ways to enhance modelling with the use of GIS. Some of these enhancements just represent moves into a GIS of tasks that have previously been handled with other tools, while some tools are new and would never have been thought of outside a GIS. In order to identify possible contributions and limitations of GIS, this technology has been used in several applications and in different parts of the modelling process. The subject is further substantiated in this introduction and in five articles. The introduction consists of a literature review and a detailed description to the original contributions of this thesis.

The background to this dissertation is theoretical and applied work during the last few years on GIS and transport models. This work covers database management in a GIS environment, descriptive statistics, parameter estimation, development of tests for spatial dependence and development of programmes for implementation of results.

The division between the material in this introduction and the material in the articles has been guided by the ambition to keep theoretical parts and parts describing principles of the subject separate. In a dissertation like this one containing much practical work such as programming in GIS, it is impossible to avoid a description of this effort for several reasons 1) it represents a large share of the research volume, 2) the results in the articles depend on this foundation and finally 3) some readers could be interested in this work on its own merits. There is one additional reason to report on practical efforts. A large body of the literature concerning GIS and transport modelling is of a conceptual nature and it is often unclear whether the content of the article has been implemented, will be implemented or if it is possible at all to implement it. The ambition with the partly quite technical description contained in the introduction is to make it clear what has been done and how it works. This is an important contribution although not a theoretical one.

The literature on the use of GIS (and mapping facilities) in transport analysis and transport modelling has grown tremendously during the last few years. Roughly speaking there seem to be two parallel paths of development where GIS plays an important role. One path starts from the spatial perspective, where GIS has been used extensively. Within this community the main criticism of GIS is the absence of up-to-date modelling tools that are capable of addressing questions concerning mobility in space. The development efforts along this path are consequently concerned with the issue of adding modelling tools to GIS. The other development path starts from the transport modelling perspective, where a need to display modelling results has grown[1]. The integration work along this second path is often concerned with conversion of networks from transport planning packages to GIS. There are principal differences between the two development paths concerning data structures, model integration and scope of the modelling process.

In addition to the developments in GIS and transport modelling, methods to treat space and spatial relations are established within spatial statistics and spatial econometrics. Taking advantage of the development in spatial statistics, crucial questions within transport modelling can be addressed such as model quality and the spatial relationships in background data and model output. Attempts to use techniques from spatial statistics have been made in the modelling process e.g. in the creation of zones and tests of model output.

The contents of this introduction cover, in conjunction with descriptions of tools developed for the articles, serve to position the work within the GIS – modelling frame. Basically, the first half of this introduction is a general text on GIS in transport modelling while the latter part is focused on the tools used in the papers. The second half mainly draws on pieces from separate technical descriptions of the different programmes used to perform the necessary calculations for each paper.

All five papers in this thesis are concerned with mobility: modelling mobility, studying errors of models of mobility, finding out what can be gained from mobility and finally finding a suitable platform for the modelling of mobility. In Paper 1, the focus is on modelling mobility in border regions, both borders within nations and between nations. We use travel surveys and networks from Sweden and Norway to model the barrier effect caused by the borders in the region between two counties on each side of the national border. The work with this first

[1] You do not have to be a GIS purist to have the opinion that GIS is not solely about visualisation but this is a frequent misunderstanding among people coming in from other areas.

paper *'Barriers in travel models'* required a large amount of work with databases in a GIS environment, databases which we have recycled in the remaining papers. The work with Paper 1 also required the development of a GIS-based tool for estimation of logit models of combined mode and destination choice which has been used in the remaining papers. Like all research, Paper 1 gave rise to more questions than answers and what we asked ourselves was "How would the model have been affected if we had not estimated the barrier effect of the border?" Our hypothesis was that the residuals from the model would have been systematic in a spatial sense. These ideas resulted in the Paper 2 and Paper 3 in this thesis *'Identifying local spatial association in flow data'* and *'Detecting barrier effects'*.

In the second paper we develop a statistic for spatial association for flow data by generalising the statistic by Getis and Ord (1992) and Ord and Getis (1995), G_i and G_i^*. This local measure of spatial association, G_{ij}, is associated with each origin-destination pair and can thus be used for explorative analysis of e.g. matrices of travel flow. In the third paper we use a modified G_{ij}-statistic, $G(\Gamma)$, in an attempt to identify barrier effects in flow data. The $G(\Gamma)$ approach for studying barrier effects is compared with the traditional approach of using dummy variables to model the barrier as in Paper 1.

One central evaluation criterion of the transport system is accessibility. The use of GIS in transport planning has renewed the interest in developing accessibility measures adapted to this new technology. Traditionally, accessibility measurements have been based on zones, while much of the renewed interest is towards individual measures of accessibility based on the space-time framework in geography. Zone-based measures have been criticised for being unrealistic and for neglecting the fact that most people face a mandatory daily travel pattern, which is not taken into account by zone-based accessibility measures. Individual measures on the other hand require detailed information on space, the individual and the means to overcome space. This is a serious drawback from an operational point of view. In Paper 4 *'Path based accessibility'* we try to combine these two traditions and develop a traditional accessibility measure that takes one mandatory activity pattern into account i.e. travel to work trips.

In our first stumbling attempts to implement our research efforts into useful software, we wrote our programmes in the FORTRAN programming language. We fed our programmes with data from GIS by importing files and visualised our results by exporting data back to the

GIS. This type of data shuffling is obviously not efficient and secure and we needed to find alternatives. Our solution to the problem was to use a macro programming language that is built into a GIS (TransCAD). This programming environment worked satisfactory with regard to data management and data transfer. In addition, we also obtained the possibility to provide our algorithms with a user-friendly interface. We used this programming environment to implement the results from the previous papers in a user-friendly GIS based software. We also wrote a traditional 4-step travel demand model in the same programming environment. However, the advantages of working inside a GIS come with a price, which in this case was computational efficiency. Some routines in transport modelling are very demanding from a computational point of view. So when time was scarce, we had to take one step back and rely on "traditional" programming languages. This time we wrote our programmes to take advantage of the database within the GIS and thus minimise the problems with data transfer. In this way, we gained experience from different types of GIS – model integration which is of interest for its own sake. We summarise our experiences from travel demand modelling in GIS in Paper 5, where our programmes are described and evaluated in a GIS context.

Why use GIS in transport modelling?

This work rests on one technical platform – GIS. Referring to the discussion in the introduction, GIS and transport modelling is approached from the spatial perspective with the ambition of adding modelling tools to a GIS.

Before going into details of the specific system used in this study, it is appropriate to specify the meaning of the term GIS. GIS has turned out to be a concept that covers a variety of tools related to computer based visualisation, storage and retrieval of data (Bailey and Gatrell, 1995). Definition of GIS has been done numerous times before[2] without achieving a generally accepted agreement. The most ambitious definitions cover everything from hardware, software, databases and algorithms to the organisation of people that operate the system, see e.g. ESRI (1990). A single "GIS user" with a "GIS software" would

[2] See Heikkila (1998) for a listing of a few definitions.

according to this definition not necessarily be regarded as a GIS. At the other extreme you may find definitions like: GIS is what is in the box (Heikkila, 1998). There seems to be an agreement that at least (spatial) database management facilities, routines for (spatial data) analyses and visualisation facilities should be included in a GIS. The problem at hand then may determine the requirements of each of the three facilities just mentioned. This "elastic" definition is quite good for the researcher who thinks of GIS as a platform to add functionality to a specific problem. The contents and scope of different GIS vary a lot and the term GIS is used in different ways. A transport orientated GIS (GIS-T) contains many facilities that are not present in standard GIS packages and that may be difficult to understand without a transport modelling background.

In transportation analysis, some of the three cornerstones of a GIS need further development to reach the desired level. Besides the issues of handling transport data and flow data, the analytical side of most GIS is far from satisfactory. Most GIS applications are based on making queries on spatial databases for visualisation. Sometimes this is mixed up with spatial analytical tools, which are something different[3]. Bailey and Gatrell (1995, p. 7) define spatial data analyses as

> "... where observational data are available on some process operating in space and methods are sought to describe or explain the behaviour of this process and its possible relationship to other spatial phenomena. The object of analysis is to increase our basic understanding of the process, assess the evidence in favour of various hypotheses concerning it, or possibly to predict values in areas where observations have not been made."

A tool for this type of analysis requires (besides spatial queries and visualisation) serious modelling components, use of inferential statistics and some kind of forecasting method that takes space into account. This process applies of course to transport.

So why use GIS in transport modelling? Among the factors behind the use of GIS in transport modelling there are developments that made it possible to work with GIS as the main platform and developments that made it desirable. The basis of every transport modelling system is the transport network. This makes it necessary

[3] With a slacker definition of analytical tool, this is analysis.

to use some kind of network database to compute travel times and travel costs between origin - destination pairs (OD-pairs). Such a functionality has been available in different GIS for some time and with transport-orientated GIS it is possible to carry out such operations for large OD-matrices. Another important factor that made it possible to use GIS was the increasing availability of geographical data. Transportation analysis in a GIS is quite demanding as regards the quality of data. In a normal transportation network, the important attributes are link cost and connectivity of the network. These kinds of data do not necessarily need a geographical reference. GIS data on the other hand have detailed geographical references but usually not the information needed to perform transportation analyses. Normally, the representation of roads in a GIS is limited to a picture of the road containing attribute information and not the data needed to perform transport analysis. Bringing the best from these two worlds together is quite demanding. Availability of data has increased in recent years but there is still room for improvement. In order to arrive at some ideal data, fusion of data from different sources is necessary.

After concluding that it is possible to use GIS in transport modelling we must discuss if it is desirable. One possible drawback is that the information normally stored in a GIS contains "excess information" to a transport model. A link in a network stored in a GIS will most likely consist of an array (with more than two pairs) of coordinates defining the real world location of the link. Since a transport model just needs a network with a from- node and a to- node, this is unnecessary information which slows the system. If we do not intend to take advantage of the "excess information", using a GIS could imply unnecessary work.

Assuming that we have this excess information, how does this information affect the potential for modelling? In Figure 1, some standard geographical information including a network is shown. What does it tell us? Apart from the fact that we know where the transport infrastructure is situated, we know its spatial relation to all other objects in space. e.g. what kind of land the road passes through and what kind of service facilities can be found along the road. In a GIS such spatial relations are not just used for visual information, they can be utilised in database queries where we are able to put questions based on distance, overlap etc. Using space as the reference, we can integrate almost any kind of data without any other key in the database except space. GIS has proven to be a good vehicle for integration

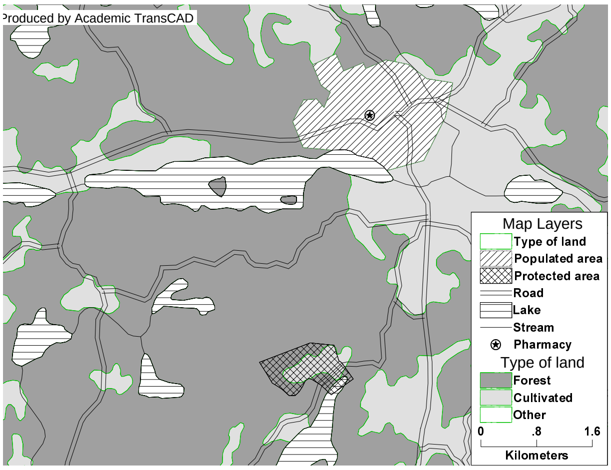


Figure 1: Example map of geographic information.

of data from seemingly unrelated themes. GIS provides the link between different technical systems through the spatial dimension. This function is usually referred to as "space as a data integrator". The transportation system is not a separate system. There is an increased awareness of the interaction between different systems such as social, ecological and economic systems (Thill, 2000). Bringing the spatial dimension into transport models is the main contribution of GIS.

1 Data for transport applications

Geographic databases[4] as such consist of spatial and non-spatial data with complex data structures used for complex analysis. In the introduction, I stated that transport data are special in a number of respects. Some characteristics of transport data are:

- Transport data are spatial
- Network data structures
- Data often represent flows

[4] Worboys (1995) is an excellent book on geographical databases (with some applications to transportation) which is a key reference for this section.

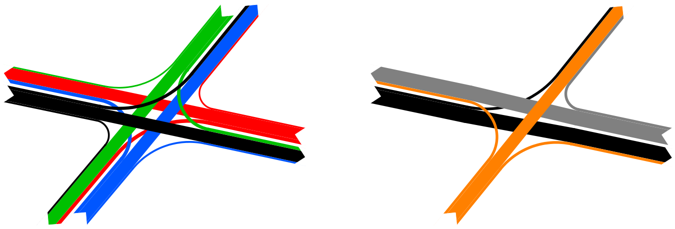
- Large amount of data
- Require array and matrix data structures
- May consist of composite geographical objects

Transportation data suffer and benefit from the same problems and advantages as other spatial data. Notable problems are the choice of spatial resolution (see discussion on page 20) and quality of spatial data. At least with Swedish data, quality is more of a problem the higher the spatial resolution is. This is mainly due to the fact that the routines for data collection on a fine scale have not matured yet. The advantage of regarding data as spatial is that space in itself provides an interpretation of data. In non-spatial databases, data are identified by their position in a file or by the relational data model. For spatial databases, space itself provides a (not always unique) key between different data objects. Space always provides some kind of information about data.

An important geographical object of transport data is the network[5], which in a GIS interpretation will be arcs (links) and nodes. In addition to the usual nodes, some objects (usually point objects) act as points of interchange between demand and supply in the transport system. The transport network is not just a representation of objects in space, it also represents legal aspects of the transport system as well as economic aspects (tariff structures etc.). Legal aspects are rules that are applied to an entire network as well as rules that apply for specific objects, e.g. turning restrictions. These aspects are not always represented in a conventional GIS (Thill, 2000). Being able to represent legal aspects is a problem with regard to data, the software and a combination of both. This is clearly a problem that deserves more attention. An example is shown in Figure 2(a)-2(b) where Figure 2(a) is a standard network from a mapping agency and Figure 2(b) is the correct representation.

The geographical objects mentioned above are typical vector-based objects. For computational speed it is an advantage to use vector-based systems. There are however examples where combinations of raster and vector based systems are used. These combined systems are used mainly for two different purposes 1) to disaggregate data from zones to raster cells (see e.g. Spiekerman and Wegener, 1995), 2) to model impacts on the environment that are continuous in space.

[5] See Nielsen *et al.* (1999) for a comprehensive text on traffic networks in GIS.



(a) (b)

Figure 2: Representation of the same intersection from two different data sources.

Data for transport applications often represent flows of people and goods between points or zones. Geographically this means that each data object is associated with two spatial objects instead of one which is standard in a GIS. The corresponding data type is a matrix or array of values instead of a scalar value for each geographical object. This is one reason why transport analysis is a special case of GIS analysis that is quite demanding. Most GIS do not handle matrices as an internal database format in an efficient way, which is a complication for transport analysis. However, the system we have used so far, TransCAD (TC), contains a database format for matrices.

Transportation applications are normally based on a set of models where some take advantage of common data but also where separate data are used for each model step. This requires large databases. In most GIS applications, data are used for display and simple manipulations. Transport models on the other hand are computationally very intensive, which requires efficient storage and fast access to data.

Above we mentioned flows which are data objects associated with at least two geographical objects: an origin and a destination. Other more complex data structures are routes that may consist of multiple geographical layers such as a line layer representing roads, nodes for intersections and a layer representing bus stops. Several routes may utilise the same infrastructure and consequently there are a lot of ”one to many” relationships within transport databases. These kinds of dependencies between objects are demanding with regard to database structure. Most GIS do not meet such requirements.

A GIS-based database for transport modelling is expected to support several modelling tools in the form of several programmes. For this dissertation, we have written about 15 different programmes for analysis, data manipulation etc. Most programmes are written in internal programming languages and can thus take advantage of internal data transfer routines (internal to the GIS). Some other programmes are written in standard programming languages and the data transfer thus becomes more fragile.

When approaching the problem of supporting several types of models with data and traditional GIS analysis, we have some requirements:

- All data are spatially referenced
- No duplication of data
- Performance
- Integrity

In order to motivate the GIS environment we want all data to be used in a mapping context. This also allows us to analyse our input data in a spatial sense. We do not want to maintain one database for specialised algorithms to operate on and one database for the GIS. Besides being a waste of space, such duplication of data implies a risk of database inconsistency and a risk of loss of integrity. An example is when two separate networks are maintained: one in a transportation simulation software and one in a GIS for visualisation of results. This is quite a common solution and could be advantageous from a second criterion - algorithmic performance. Fast access to data is critical to transport modelling since these types of models are computationally intensive. Some computational routines used in this work are written in the FORTRAN programming language and it has not been possible to immediately read the GIS database into the algorithm. In these instances, a temporary database has been created that is deleted after completion of the calculation.

The common technical platform in our work is a transport-orientated GIS, TransCAD [6]. The main reason for our choice is that this system handles data structures that are used in transportation analysis such as matrices and networks[7]. In addition, TC also has fundamental

[6] See www.caliper.com for a description.
[7] Network in the sense a transport modeller would use the term, i.e. a condensed set of information that supports computational speed.

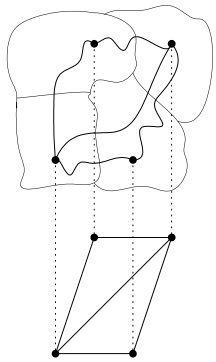


Figure 3: Representation of space in a GIS and a transport modelling package.

functionality for generating data for transportation analysis.

2 GIS and transport modelling

In this section we will go into the relationship between GIS and transport models in some detail. What can a GIS do for transport models and what can a modelling perspective do for a GIS? There is no doubt that there is an ambition with GIS to support analysis of spatial processes and enhance spatial modelling. However, somehow the results of these ambitions are quite simple tools for quarrying and visualisation.

GIS are essentially data-driven while transport models are driven by theory. Roughly speaking, GIS data are detailed while data for transport modelling just contain as much information as is needed by the algorithm. This can be illustrated as in Figure 3. The upper part represents the GIS way of looking at a zone- and network system. In the GIS world, the roads are shown with their correct location and shape. The lower part of Figure 3 is what is necessary to the transport model. A link in the transport model (and related software) is made up of a from- and a to- node, while a link in a GIS may be made up of hundreds of coordinates. This fundamental difference falls back to the second item above – the modelling perspective. A theory does not need more data than actually used. Most GIS are not designed with any specific theory in mind or are not designed for a single purpose but have a strong foundation in the spatial perspective. The lack of theoretical guidance that too often characterises GIS work often results in excessive data collection. With a theoretical foundation from the modelling perspective, much unnecessary data collection could be avoided. In a transport model the spatial perspective is replaced by the connectivity of the network. This can be regarded as the fundamental theoretical and conceptual difference between the two systems.

Transport modelling from a GIS perspective

Data in a GIS often represent land use at certain previous time points[8]. Even the simplest treatment of time, the path from one state of a system to another state is badly represented in GIS. The treatment of time in transport models has also been subject to criticism but transport models can handle simple transitions from one state of a system to another. Transport models are designed to find out what happens as a result of changes in the transport infrastructure or of changes in land use. In GIS, most "what if" questions will remain unanswered. Land use as discussed here is usually something as simple as the number of work places or the number of residents.

GIS is the ideal tool to describe, present and make sketches of current and future land use. It is simple to store, retrieve and manipulate geographical objects to make a map of the planners intentions. This can be seen as keeping track of land use at different moments in time, at the present and in the future. In Figure 4, the GIS perspective is illustrated. At time t the system is fed with geographical data: zones, networks, environmental data, service facilities and planning restriction etc. These objects are associated with data, e.g. population, employment, income and housing type and so forth for each type of object. Carefully set up, a geographical database is an excellent planning tool, since zones can be modified to serve different planning purposes and we can extract information based on combinations across space and across different variables. Now when we have zone-based data, transport-related indicators can be computed such as potential values and simple accessibility measures using buffer zones and other types of queries. Simple analytical tools are available in GIS. GIS can also be

[8] Representation of time even in a very simple way is difficult in a GIS and the treatment of the time dimension in GIS is currently one important field of research.

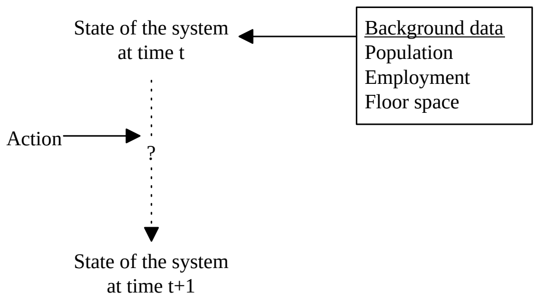


Figure 4: The GIS perspective on transport and land use.

used as tools for simple exploratory data analysis. By thematic mapping which can be enhanced by multiple views and graphs associated with maps, insights can be gained about complex spatial phenomena. A good understanding can be obtained from simple analytical tools which in some cases may be enough.

Similar slices in time can be made for numerous periods but a fundamental problem is illustrated in Figure 4. A GIS in itself does not provide any connection between the different time slices and we lack casual and theoretical relationships between the different states of the system. Batty and Xie (1994) state that:

> "GIS is essentially a storage and display medium for spatial data, and if processes are to be modelled in whatever geographical domain, such modelling must be achieved outside the GIS."

The article by Batty and Xie was written a few years ago but in this respect not much has changed. Since GIS is marketed as a planning tool (among other things), this inability to handle processes is problematic. Planning is about changes. The current state, which the data describe, is just a starting point.

The data-driven nature of GIS assumes that the programme is fed with data, much like pure mapping. GIS is not a system that generates new data about changes based on theories. The data-driven paradigm suggests a one way direction from external models and other data sources to GIS.

GIS from a transport modelling perspective

This section looks at GIS from a transport modelling perspective. Starting with the different parts of the traditional four step model, we identify GIS functionality that can be useful in a transport modelling perspective. Each part of the model as well as the whole system depends on estimation (calibration), scenario calculation and evaluation. The use of GIS can be discussed in relation to these three tasks.

In Table 1, the four steps in the traditional four step model are listed in the first column along with a data preparation step and an evaluation step. This is an evaluation framework used by Nyerges (1995). For each step, necessary input data are associated with the geographical objects and these data can be subject to manipulation.

Preparation of data – zones and network

The discussion starts with the least prestigious part of the modelling process – preparation of data. Setting up a GIS database is time consuming. A large share of the average project resources is consumed before the analysis has even started. The benefit in using GIS is that the data will obtain a quality that makes them useful for various purposes outside analysis of the transport sector.

In an early stage of the modelling process it is necessary to decide which geographical resolution to work with. It has been suggested by e.g. McCormack and Nyerges (1997) that GIS could be a useful tool to define proper regions for transport analysis. This is probably one of the most widespread utilisations of GIS in transport modelling. According to Miller (1999) the spatial resolution has generally not received much attention among practitioners. This is remarkable since it has been shown that zone design affects parameter estimates and goodness-of-fit measures and the corresponding results of scenario calculations (see Miller, 1999).

The problem of construction of proper geographical units to analysis and modelling is not unique for transport modelling. It has been on the research agenda in geography for many years. The observation that results from analyses based on geographical data could differ depending on the underlying set of zones was first observed by Gehlke and Biehl (1934). This problem is known as the modifiable area unit problem (MAUP), a term introduced by Openshaw and Taylor (1979). The MAUP can be divided into two components, the scale effect and the zoning effect. The scale effect originates from the observation that

Table 1: Illustration of the different steps in the four step model and associated data in relation to geographical objects in a GIS.

Model step	Data	Geographical object	Geographical manipulation and computation
Preparation of data	Land use, indicators of zone characteristics	Zones, network	Creation of zones, centroid positioning, connection of zones to network
Generation	Socio-economy, accessibility	Zones, centroids	Exploratory analysis of model output
Distribution, mode choice	Travel time, travel cost	Network, zones/centroids, nodes	Exploratory analysis of model output
Network assignment	Volume delay function, parameters	Network	Exploratory analysis of model output
Evaluation of alternatives	Accessibility, externalities	Everything with a location	Computing accessibility, pop. at risk

different statistical results can be obtained from the same data set depending on the level of spatial aggregation (from smaller to bigger regions). The zoning effect denotes the differences in results related to the spatial subdivision at a given scale.

It has been argued that the finer the zonal resolution, the better. This is false and ignores the fact that a zoning system is a generalisation operator with a spatial analysis function (Openshaw, 1996). Openshaw (1996) argues that zone design is an analytical tool that can be used to e.g. capture patterns in data, describe space and be used for testing of spatial hypotheses. However, it must be noted that aggregation adds noise and that a finer resolution increases the flexibility of the data. There are no simple guidelines for zonal creation but the zonal system must say something about the nature of the data and the model (see Openshaw, 1996).

The obvious disadvantage of using small zones is that they are likely to be highly spatially autocorrelated. Small zones may also produce unreliable results due to few observations. These first two drawbacks apply to all kinds of spatial modelling work, including transport modelling[9]. On the other hand using large zones will increase the intrazonal variation and, specifically in transport models, the number of intrazonal trips will increase and lead to underestimation of traffic volumes on the networks. You *et al.* (1998) use a measure of spatial autocorrelation (Moran's I) to design traffic analysis zones (TAZ). The whole zone design process in You *et al.* (1998) is carried out in a GIS (ArcInfo) using macros. This is probably one of the best examples of traffic analysis zone design in GIS. There are other examples as well based on e.g. clustering techniques that do not explicitly address the core of the MAUP, the spatial autocorrelation.

The typical working mode in zone design with GIS is to carry out the zone aggregation within the GIS and then export the zone centroids to a transportation modelling package. This strategy is e.g. used in Choi and Kim (1997) and Ding *et al.* (1994), where ArcInfo and TRANPLAN are used.

An example of zoning

For different parts of the modelling work reported here, it was necessary to change the zoning system. Instead of using the quite am-

[9] The role of spatial autocorrelation in standard travel demand modelling is not satisfactory studied.

bitious zoning algorithm addressing spatial autocorrelation by You *et al.* (1998) a simple rule-based framework was chosen, which was considered to be useful in this particular context. It is important in our context to take advantage of the GIS environment by combining different spatial data sources, and to require that the zoning system represents the realities of the transport system. An advantage of using simple rules to create zones is that the result is easily understood, a disadvantage compared to e.g. You *et al.* (1998) is that the core of the MAUP is not addressed. However, some features of zoning for traffic analysis may not be subject to compromises related to possible autocorrelation, so a system based on simple rules might be appropriate.

As a starting point for zoning we need a description of the relationships between the zones under consideration. The relationship between zones can be described by a contiguity matrix of dimension $n \times n$ where n is the number of zones. Each element of this matrix can be given a value representing zone i's relation to zone j. Usual ways of representing relationships between zones are 0/1 (neighbour or not), common border length or distance between midpoints of the zones. These values can be translated to rules for merging zones, i.e. zones must be adjacent to be merged, or priorities for best merger, i.e. longest common border length. The principle of merging zones with a long common border is usually referred to as compactness. Contiguity matrices can also be based on spatial attributes other than internal relationships between objects of one class. One example of such a object in the case of zone mergers is the presence of barriers between a pair of zones.

In our case three simple rules were adopted and implemented in a short programme. *Rule one: Continuity.* To allow a pair of zones to be merged they must share a common border. This is a common rule in zoning practice. Rule one is simply implemented by using a standard binary adjacency matrix.

Rule two: No water between zones. The greater Stockholm area is water-based. Besides thousands of islands, the whole region is split into two halves by Lake Mälaren and in addition there are hundreds of other water areas to obstruct mobility. To merge two zones separated by water would not be a good representation of the possibilities of mobility within the zone. The Stockholm example of water could easily be translated to some other type of barrier important to transportation or zoning for some other purpose. Rule two was implemented by creating a matrix of the $n \times n$ zones representing adjacent zones,

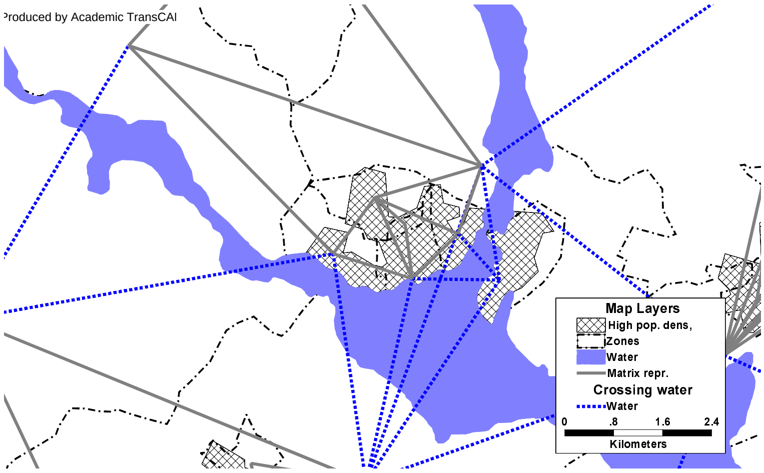


Figure 5: Illustration of data behind zoning.

except in cases when water areas (or rivers of specific width) larger than a specified value are defined as an obstacle. This is illustrated in Figure 5, where the lines represent neighbouring zones and thus potential zones of merger. Some lines (dotted) go across water and are thus not acceptable candidates of merger. The overlay class of objects (water) is used to compute the matrix representing rule two (0 if zone i is separated from neighbouring zone j by water 1 for all other pairs of zones).

Rule three: Homogeneous from a functional point of view. This rule reflects a preference to merge zones that belong to the same cluster of dense population and not to merge zones of high and low density. To implement this principle, the original zones were overlayed with polygons of high population density and an identifier was added to the zones as a variable of similarity (same value if both zones belong to the same dense area and no value if both zones are independent of different clusters of dense population).

Finally if all rules are fulfilled and there are two or more candidates of merger, the two zones with the longest common border are merged. Beside these rules, some candidates or independent zones ("seeds") are needed to start an aggregation. These zones were selected according to a size criterion.

Does it work? That depends on whether the rules are found to be acceptable. This example illustrates the use of generic GIS functionality to prepare input data for analysis.

Trip Generation

Some modellers like to see the whole modelling process as an art where creativity and imagination are as important as formal modelling skill. This process involves trying out and identifying spatial phenomena. In this process, looking at data is one of the tools available to establish an understanding of the underlying process.

The formal first step of the four step model is trip generation. Data for this model step are zone-based. This simply means that a scalar value is associated with each zone which is a data structure that all GIS can handle. In order to check the quality of the data visualisation is an important tool.

Trip generation can be expressed as the total number of trips made from a zone explained by socio-economic characteristics of the zone. Alternatively trip generation can be defined at the individual/household level and expressed as the likelihood of making a trip. In the first (aggregate) case, standard regression techniques apply and they will be subject to the spatial effects that spatial regression models suffer from (spatial dependence and spatial heterogeneity). When dealing with such problems, GIS will be necessary in various stages of the model estimation process. If the model is estimated at the household level, typically with a logit-model, the tool box from spatial econometrics will not be available. Despite this, it is often useful to incorporate spatially lagged variables for explanation [10]. Accessibility is a variable that is commonly used in travel demand modelling which can be seen as a spatially lagged variable [11].

Trip distribution and mode choice

Trip distribution and mode choice is discussed in the same section since the two steps are modelled simultaneously in the applications later on in this introduction.

Data for this model step are usually a mix of scalar values associated with zones (e.g. parking costs) and matrix data associated with

[10] This can be done using a GIS see e.g. Berglund and Karlström 1999b.

[11] If we let $\exp(-\beta d_{ij}) = w_{ij}$ we have the standard lagged variable WX.

origin- destination pairs (travel- time, cost). Matrix data are notoriously difficult to handle for GIS. Most GIS lack this kind of functionality, which makes relational databases the alternative for storing data. This is of course not very efficient and even small matrices require considerable storage space. However, we use one of the available GIS that can handle both matrices and relational databases (TransCAD). Shortest path routines are standard tools in practically all GIS. Preparation and storage of data can thus be done entirely within a GIS. The geographical objects associated with the background data are different networks (car, public etc.) and zones. Often the travel time from a link is given by an equation, a volume delay function, which does not cause any problems in a GIS.

The estimation of the combined trip distribution and mode choice model is usually regarded as the core of travel demand modelling. This requires separate estimation programmes that are fed with data from external sources. As in the case of trip generation, trip distribution provides output in the form of estimated data – flows. Flows are matrix data associated with OD-pairs. As noted above, this may cause problems in a GIS. There are some examples in the literature where GIS are used in model estimation and calibration. One example is visual display of model outputs, which has been an aid in model calibration (MDOT, 1998). In Berglund (1998) a GIS was used to provide data for an estimation programme (see separate section on GIS and model estimation below). A close relationship between model estimation and GIS could be beneficial for analysing model output (i.e. not only for data display). The use of spatial statistics for analysis of transport data has not yet been satisfactorily explored. The use of GIS is essential for incorporation of tools from spatial statistics into transport modelling by providing input data such as spatial weight matrices. Berglund and Karlström (1999b) used a generalisation of the well known G_i- statistic allowing for flow data (G_{ij}) to analyse residuals from a combined mode choice and destination model. The G_{ij}- statistic has later been programmed in a GIS for convenient and interactive analysis of flow data (Berglund and Karlström, 1999b). In Berglund and Karlström (1999a) significant G_{ij}- values were selected and displayed as in Figure 6.

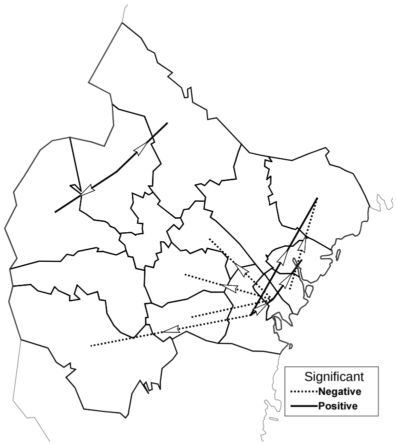


Figure 6: Illustration of identification of badly predicted values. Arrows representing flows with high (low) $G_{ij}(W^d)$- statistic, indicating high (low) residual flow from zone i to a neighbourhood of zone j. Source: Berglund and Karlström (1999a).

Network assignment

Data resulting from assignment are scalar: the number of vehicles on a link per unit time. The assignment itself is normally done independently of GIS even if a software like TransCAD contains assignment routines. When data are brought back to a GIS, a variety of analysis possibilities are open, including descriptive statistics and inferential spatial statistics (see e.g. McCormack and Nyerges, 1997). An analysis of network and link attributes using spatial statistics has been carried out by e.g. Black and Thomas (1998).

Evaluation of model results and alternatives

Evaluation may be one of the strongest contributions of GIS to transport modelling. Model results in raw form are quite demanding to interpret. The built-in visualisation tools in standard transportation packages do not reach the level of a standard GIS. The advantages of using GIS in presentation and interpretation of results is stressed by many authors, e.g. McCormack and Nyerges (1997), USDOT (1998). Even if visualisation is important, it is of limited interest as a research object.

There are many ways of supporting the evaluation of alternatives by using GIS. Common examples are to feed GIS-based emission models with link flows. Other examples are accessibility calculations within GIS, see e.g. Berglund (1999), Kwan (1998), Geertman and van Eck (1995) and Paper 5 in this dissertation. The large amount of literature on GIS and accessibility shows the central position of GIS-based evaluation tools.

3 GIS and analytical tools

A common complaint about GIS is the lack of built-in analytical facilities. In systems containing built-in models, these are often over-simplistic and worse than the modelling procedures used outside GIS. According to Fotheringham (1995) these tendencies for using GIS in conjunction with bad models represent a step backwards for spatial modelling. Modelling tools for transport analysis are no exceptions. Even in a specialised GIS such as TransCAD, the available tools can be regarded as below the aspiration level of the advanced transport modeller. Most GIS are made to serve general purposes and it is unrealistic to expect to find a perfect tool for specialised needs. Being aware of the danger of getting trapped in the use of bad models when trying to integrate GIS and transport models, we choose not to rely completely on tools that are built in to a GIS. Most GIS are open to communication with other programmes and there are several possibilities for adding functionality to GIS and for adding GIS-functionality to programmes written by the user. Since there are alternative ways of integrating models and GIS, we will briefly mention the main principles in order to position our own work in the continuum between total separation of models and GIS and full integration.

A simple alternative is to export data from the GIS in a format that can be used by other programmes. We call this type of integration loose coupling. Export of data can be automated using the GIS interface and the external programme can be called from the GIS (this has been done in the estimation programme described in this introduction). Examples of external programmes relevant for transport applications are estimation and scenario calculation routines. The processing of data outside the GIS can by these means be made invisible to the user and still controlled from an interface that is similar to

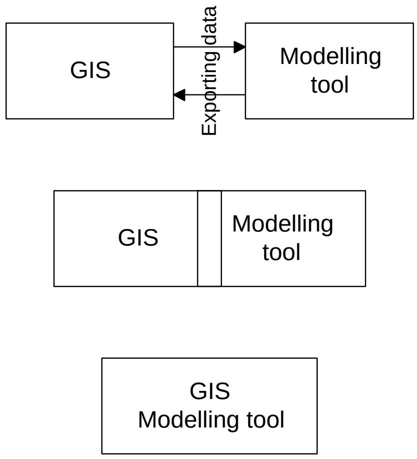


Figure 7: Different strategies for integrating GIS and modelling tools. From top of figure: loose coupling, close coupling and full integration.

the native GIS environment.

Adding models can be achieved from the inside of a GIS by writing programmes in macro languages supplied with the GIS. We call this type of integration close coupling (see Figure 7). Such macro languages are designed to perform tasks that can be written in reasonably short programmes. One example of this strategy is the computation of accessibility as described in this introduction. Macro languages are, however, not suited for operations that require large amounts of computation since they are considerably slower than standard programming languages (such as C or FORTRAN). To maintain computational efficiency, an alternative approach is to call GIS routines from an external programme. It is hard to obtain flexible GIS functionality in this way but if mapping alone is enough, it is an option. Both alternatives for GIS – model integration described so far can be done from the GIS interface. From the user's point of view the difference could be negligible. We could from the same interface control several programmes integrated in different ways. The possibilities of using GIS as a base for model management have been pointed out by e.g. Miller (1997).

The last alternative in Figure 7, i.e. full integration of modelling tools and GIS, would imply that the whole system should be within the same executable unit. That is possible only if one has access to the source code of the systems, which is not the case for the ordinary researcher.

One of the alternative strategies – loose coupling between GIS and the analytical procedures – relies on exporting and importing data back and forth between different software or using common files. Examples of different types of loose coupling can be found, e.g. between ArcInfo and different statistical software or transportation modelling packages (see e.g. Choi and Kim, 1997) The major argument against coupling of two independent software systems like GIS and statistical packages is that both may be very large, complex and costly.

In most cases it is hard to achieve a smooth integration without the possibility of modifying some parts of the source code in either the GIS or the analysis routine. Writing an interface between two complete pieces of software is a temporary solution to the integration problem. With possibilities of editing the source code in any of the two programmes (GIS or analysis module) smooth and secure data transfer may be achieved. This is the strategy we will develop.

4 Accessibility

In Paper 4 (Path based accessibility, PBA) of this dissertation an accessibility measure is developed. PBA represents a mix of two traditions in accessibility analysis: the constraint approach where daily activity patterns are central and traditional aggregate accessibility. Since individual accessibility measures are not so frequently used we take a look at the types of measures and the theory behind them. We also provide a small application. In short individual papers there is no room for describing the application platform. This is done for the case of accessibility at the end of this section.

Accessibility is the individual's possibility to take advantage of resources with a fixed location in space that requires presence. It is obvious that access to resources is restricted by the possibility to overcome distance. Mobility requires resources in terms of time, money, fixed capital, environment etc. Access also faces restrictions with regard to the above- mentioned resources and some additional ones e.g. opening hours or appointments at a fixed time. Resources can also be subject to restricted access regardless of presence. Not everyone may be allowed to or be able to enter the resource. Possibilities to overcome distance and other obstacles to mobility differ between individuals depending on e.g. where they live, work and their mobility resources.

The transport system provides accessibility as its main positive effect. All side effects of transport such as noise, air pollution, energy consumption and accidents are undesired. These side effects are often modelled in GIS and presented in GIS with e.g. colourful maps. Accessibility is usually computed with less advanced tools and will perhaps in the end be presented as maps. How the positive effects of transport are measured and presented is thus critical.

In this section, three easy-to-use programmes are presented that compute accessibility along two different theoretical lines. The first programme computes standard aggregate accessibility measures (AM). Aggregate AMs can be computed simply as cumulative opportunity measures or spatially discounted opportunity measures based on gravity principles. Both types of AM are computed in the programme. Aggregate measures represent the standard way of studying accessibility but recently there has been a growing interest in individual AMs. This is partly due to the developments in GIS. It has become simple and convenient to write programmes and visualise individual behaviour in the spatial environment GIS provides. The theory behind individual AMs is however mature and originates from the work by Hägerstrand (1970) and Lenntorp (1976). In time - space geography the key determinants of accessibility are the restrictions in time that set the limits of mobility. The strength of individual AMs is the description of individual mobility and the opportunities within the activity area. The drawback of these individual AMs is the difficulty in providing overall measures of the accessibility of an area or of a group of people. The second programme presented in this report computes an AM for descriptive purposes inspired by the time - space geography.

A third AM that is a mix of the two previously mentioned traditions is also provided in a programme. This "path based accessibility" is based on Paper 4.

The traditional methods of studying accessibility originate from Hansen (1959). Other important contributions have been made by Weibull (1976, 1980). Recently there has been a renewed interest in activity-based accessibility measures. This renewed interest is partly-technology driven by the developments in GIS. One representative of these GIS-based accessibility studies is the work by Kwan (1998).

Other important contributions have been made by Miller (1999) where the axiomatic approach by Weibull (1976) has been used to formulate space-time measures of accessibility.

Accessibility measures

Accessibility measures are based on two components, a source of supply and a friction of distance. These two components can, depending on the functional form, be given different weights and importance. Accessibility can broadly speaking be regarded as the utility discounted by a function of distance. Depending on what is stressed, the sources of supply or the transportation system, different aspects of the AM are more or less developed. Some authors are neutral with regard to supply and distance friction, such as Koenig (1978):

> "The concept of accessibility usually associates an appreciation of the availability of satisfactory potential destinations with respect to a given need."

Others inspired by the time - space school of geography stress the constraints faced by people, such as Weibull (1980).

> "Accessibility concerns physical and temporal constraints on behaviour and thus is an aspect of the freedom of action of individuals."

The latter definition is somewhat wider in that it opens up for other types of frictions besides those caused by the transportation system. The focus on restrictions is mainly associated with the time - space geography framework and is normally not incorporated in traditional AMs. Time - space measures are more orientated towards description than towards measuring the overall utility of a system.

Another important question related to AMs is whether to study accessibility in the overall transportation system or in each part of the system. Often the aim is to improve the system or parts of a specific system so it is important to be able to identify the contribution of an improvement in a specific system. This calls for separated transport systems and the use of network distances in the AMs.

The full list of AMs covering different aspects of accessibility would require a large amount of data. Normally it is not necessary to use variables for studying accessibility other than those used in transportation models. A description of accessibility could cover the following factors:

- Attributes of individuals (age, gender, income)
- Attributes of space (travel time, travel cost)
- Attributes of mobility resources (available modes, access to transport networks)
- Attributes of the resources (capacity, opening hours, congestion, quality of the resource)

The complete description of all the different attributes can in most cases not be achieved. Often the lack of data on individuals is the main restriction to a complete accessibility analysis. Instead, measures of the aggregate population are used in most cases. Nevertheless for descriptive purposes, individual accessibility measures can be of great interest. Such measures can be computed by using small samples or by using hypothetical individuals.

Individual accessibility

Individual AMs can be used to find out what resources a person can reach from a location given a set of restrictions. Aggregate AMs are normally used to measure the accessibility from a certain location to a given type of resource without any type of restriction. Individual accessibility differs between places and individuals depending on relative location, mobility resources and commitments in time and space. Everyone may experience his own accessibility which may differ from the accessibility of other individuals. Access is unequal and certainly not equally distributed in space or across the population.

Hägerstrand (1970) identifies three types of constraints: capability constraints, coupling constraints and authority constraints (for a discussion of these concepts see Hägerstrand (1970) and Lenntorp (1976)). Data for capability constraints are mainly associated with individuals, while coupling constraints data are associated with sources of supply. The supply source may be e.g. another person or a store. Examples of coupling constraints are opening hours or a person-to-person appointment. Authority constraints (can be temporary) mainly represent spatial limitations of access. The data of authority constraints are typically associated with the area, which could be a supply source but which has also got coupling constraints associated with it. Individual AMs can deal with these types of constraints one at a time or simultaneously.

The difficulty of modelling constraints increases with the number of constraints. Even taking a single constraint into account is not simple in practice if space, travelling time and supply sources are represented in a reasonably realistic way.

The simplest individual AM computes the possibility of being at a location k at a certain point in time t given one origin- (i) and one destination- (j) constraint with fixed departure and arrival times. If we add information about travel times t_{ij} we can define the Potential Path Space (PPS) defined in Equation (1), see Kwan (1998). The PPS as stated above takes individual time constraints and the travel time into account. The potential path space can be computed as

$$PPS = [(k,t)|t_i + t_{ik} \leq t \leq t_j - t_{kj}] \qquad (1)$$

where t_i = the latest ending time of activity at i, t_j = the earliest starting time of activity at j, t_{ik} = travel time from i to k, and t_{kj} = travel time from k to j. With this information we can e.g. illustrate the space time prism and its projection onto the map: the Potential Path Area (PPA). PPA gives us the maximum time which it is possible to spend at each location k. It is obvious that for an area k to be in the PPA, t must be greater than 0 in the PPS. In this framework it is assumed that the PPS is fixed given the departure and arrival constraints. It is not possible to get away from any constraints by e.g. paying an additional fee, i.e. all valid opportunities or supply sources must be within the PPA. In the short run such an absolute restriction in time may be dubious but hardly in the long run.

From the information in the PPS a number of accessibility measures can be calculated, e.g. 1) the number (or sum) of opportunities within the opportunity set, 2) the maximum time that it is possible to spend at each location, 3) the sum of the opportunities weighted by some function of distance to the opportunity. In the time-space setting it is also possible to calculate the extra travel time that an activity requires (our fourth AM) given the two initial activities at i and j. This aspect is an important aspect that is not taken into account by other types of accessibility measures. The extra travel time caused by going to k is given by Equation (2)

$$t_{k|ij} = (t_{ik} + t_{kj}) - t_{ij}. \qquad (2)$$

i is allowed to be equal to j which means not making a trip. If $i \neq j$ and k is along the road from i to j the extra travel time equals 0.

Finally we have to add the time consumed by the activity (t_k^l) itself at each location k. The total time consumed by an activity will then equal $t_{k|ij} + t_k^l$. If t_k^l is impossible (or difficult) to obtain we could use some stop penalty.

Alternative 1 calculating the sum of the opportunities does not take human behaviour into account apart from the time constraints. Time constraints are not necessarily associated with the individual, rather most constraints are consequences of the surroundings. They do not say anything about individual preferences and the utility of individuals, i.e. the utility of a location is equal across space as long as it is within the PPA. Alternative 1 has a close relationship to cumulative opportunity measures, the difference here is that the sum is restricted to the PPA given a required presence at fixed locations.

Application – Accessibility in GIS

In this section we illustrate the different accessibility concepts by describing the use of the two programmes. Since the input data differ considerably between the space-time inspired AMs and the traditional AMs they are programmed separately.

The programme presented in this introduction are written in the internal programming language of TransCAD (GISDK[12]) and use the same interface, data and internal compiler and are thus closely coupled. Close coupling provides an elegant solution and possibilities for frequent interaction between the analytical routine and map display.

The programme for individual accessibility and path-based accessibility are described below. Since aggregate accessibility is so well established we omit it.

Individual accessibility

In this programme the number of constraints is limited to two locations in space and two constraints in time. With a minor programming effort it is possible to generalise the programme to allow for a more complex set of constraints. The individual AMs computed by the programme are the following:

- Time that it is possible to spend in each location
- Sum of opportunities within the PPA

[12]Geographic Information System Developer's Kit

- The log-sum restricted to the PPA

- The extra time that is required to go to k given i, j

- And of course the PPA (PPS) itself

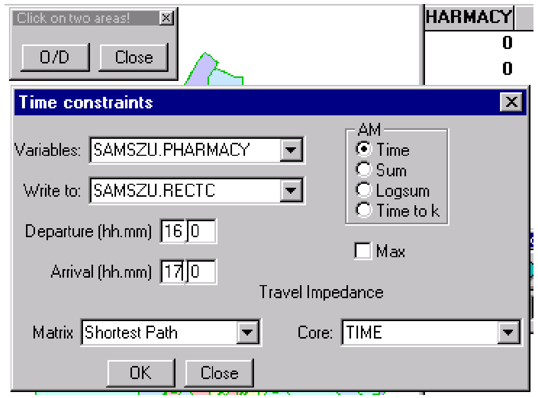
All alternatives can be computed as the sum (aggregated over opportunities) or as the max (best opportunity) within the PPA which leaves us with eight different individual AMs. The use of the programme proceeds as follows:

1. Invoke the routine and a dialogue box is shown on the screen that prompts the user to click on two zones with the mouse (different areas or the same two). These two zones are used to specify for example the constraints in space, i.e. i and j in Equation 1. The dialogue box is shown in the upper left corner of Figure 8(a).

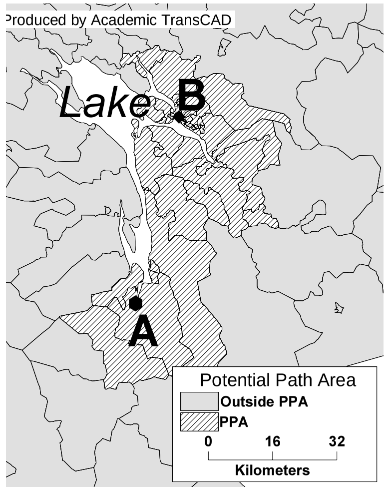
2. A new box appears that prompts the user to type in the time constraints associated with i and j above. We also need to specify the supply source, a variable where the output is written, the type of AM, the use of the maximising type of AM and a travel time matrix, Figure 8(a). The summation type of AM is the default choice. A difference compared to other AMs is that zones outside the PPA will contribute with an accessibility score of zero.

3. The programme is written to allow the user to work in an interactive fashion. The dialogue boxes will remain on the screen until they are closed. To alter the time constraints, the user just has to type in new departure and arrival times and the map will be updated with the new PPA.

Individual AMs do not provide an overall picture of the accessibility in the system but provide illustrations of the limits of individual mobility. AMs of this type cannot be used as stand-alone evaluation tools of changes in infrastructure but are useful for describing potential impacts on individual mobility of travel time reductions. It is for instance possible to study the PPA before and after changes in the transport system or new sets of opportunities generated by alternative scenarios.



(a) Screen picture



(b) PPA

Figure 8: Illustration of programme for individual accessibility.

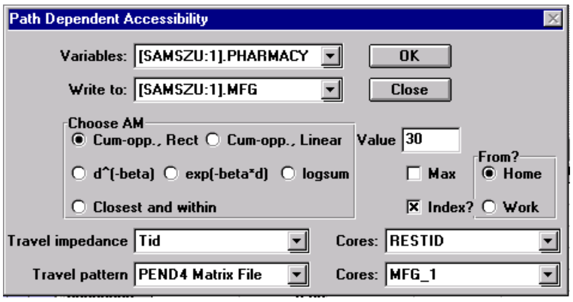


Figure 9: Picture of dialogue box for path based accessibility.

Path based accessibility

This programme was written to implement the ideas presented in Paper 4. Like the previous programme it is written in a GIS programming language (GISDK). This programme also exists in a loosely coupled version written in FORTRAN. The comparison of runtime between the two versions shows a considerable difference in computational efficiency and thus supports the conclusions in Paper 5[13]. In this programme different types of impedance functions can be used (see Figure 9 and what differs compared to a traditional AM is that this measure is weighted by a travel pattern. This travel pattern is represented by a matrix that is selected at the bottom of the dialogue box in Figure 9). As an option the user can choose whether to compute accessibility with regard to work or residence.

The parameter value for the parameterised AMs refers to the travel cost sensitivity with regard to the secondary trip, given a travel pattern. Estimations based on data from the Stockholm region show a rather high cost sensitivity, which indicates that the secondary trips are often short. The use of the programme is intended to be simple.

[13]A problem when using-GIS based programming languages is that these languages are considerably slower than standard programming languages and thus less suitable for computationally intensive routines.

5 GIS and model estimation

All articles in this dissertation require estimation of logit models. The model structure (for the mode and destination choice) used throughout this work is of structured or nested logit type with two choice levels. The model can be formulated with either the destination choice ("normal structure") or the mode choice ("reverse structure") on the upper level. A third alternative is to keep the λ parameter fixed at 1 and estimate a simultaneous model. The parameters are estimated using maximum likelihood simultaneously for all parameters. Regardless of structure, the model is formulated as a simultaneous choice of residential zone/employment zone and mode choice conditional on employment zone/residential zone.

To obtain the parameter estimates of the logit model we have relied on a software written "in house", *ester*. *ester* is a subroutine to an optimisation package "MINOS" and was originally written by Lars-Göran Mattsson[14] to estimate different transport and location models for the Stockholm region. One example is the the IMREL model system (Integrated Model of Residential and Employment Location), Anderstig and Mattsson (1991). The programme has been enhanced in several steps to allow for more parameters to be estimated and to be run in a Windows environment (originally the programme was used in a UNIX system).

This programme is written for one purpose: to provide estimates of the parameters of a combined destination and mode choice model of doubly constrained logit type. For other modelling purposes there are probably better alternatives available. The advantages are: the programme can handle many alternatives (particularly good for the destination choice) and can estimate many constants for the zones (doubly constrained!). A practical advantage of the current version is that it can be used from inside a transport-orientated GIS using the internal data structures and the parameters can be estimated without knowledge of any additional software (other than the GIS). Given the Combined Mode and Destination choice model (CMD-model) the programme is also flexible with regard to the nesting structure of the model and with regard to modelling location of jobs and households respectively.

ester has gone through several development stages. For Papers 1

[14]A separate documentation for ester without the GIS interface is available from Mattsson (in Swedish).

to 3 *ester* was used as free-standing software reading TC input files. During the work with the two last papers, *ester* was controlled from an interface within TC. One of the most important changes include a generalisation to allow for different numbers of parameters, different structures etc. without programming and compiling. Other important changes in the current version are that *ester* reads TransCAD files as input data: OD-matrices, travel times, starting values for the optimisation solver and settings that control the estimation procedure. The original version of *ester* was designed to read input data from EMME/2. During these changes errors may have been introduced. The author of this dissertation is solely responsible for any error in the current version of the programme.

The last version of the programmes appears as an integral part of the GIS but consists of three different parts, 1) an optimisation programme – MINOS[15], 2) a subroutine to MINOS – *ester* where the ML-problem is defined and 3) an interface – "TC-est" between TC and MINOS+*ester*. MINOS and *ester* are both written in FORTRAN and are compiled and linked to one single executable unit while the interface is written in a macro language (GISDK) internal to the GIS. The interface provides input data to *ester* through files all written in TC's file format. The output from *ester* is parameter estimates that can be read by TC and the scenario calculation routines (see below). Part of the parameter data in the doubly constrained model is of a geographic nature (the balancing parameter associated with each zone) that can e.g. be mapped. These parameters are then used within the GIS to compute scenarios and calculate accessibility measures. This solution is chosen in order to provide a smooth integration between the GIS and the estimation programme. The user will only be confronted with the interface which looks like a standard TC dialogue box. Everything behind will be invisible to the user. TC-est appears as an integral part of TC and does not visually differ much from any other routine. The programme structure and the information flows are illustrated in Figure 10. The estimation programme is still a separate executable unit and thus loosely coupled to the GIS. One difference between this coupling and several other loosely coupled programmes is the fact that we have access to the source code both of the interface that calls the programme and the estimation programme itself. This enhances the integration but it is nevertheless – loosely coupled.

[15]Murtagh and Saunders, 1995

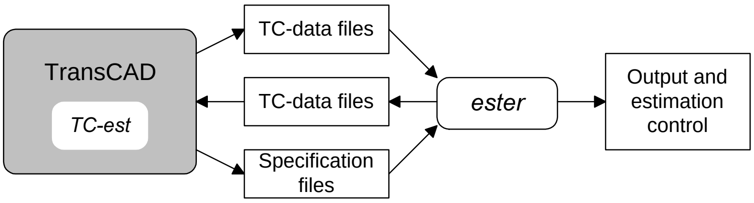


Figure 10: Data flows in TC-est.

6 Analysis of flow data

Papers 2 and three are concerned with the analysis of local spatial non-stationarities in flow data. Both papers draw heavily on the G_i-statistic developed by Getis and Ord (1992) and Ord and Getis (1995). Local measures of spatial association would hardly exist without GIS and the implementation platform is thus given. Despite this we made our first implementation in a free standing programme and relied on manual import and export of data between GIS and our programme. In Berglund and Karlström (1999a) we realized our mistake and conclude by writing:

> ”In the perspective of the growing interest in explorative spatial data analysis (ESDA) methods [see, e.g. Anselin and Bao (1997), Unwin (1996), Ding and Fotheringham (1992)], it is a natural next step to implement the G_{ij} statistic more fully integrated with a GIS. In the same way as the G_i statistic has been implemented as macros in GIS software such as ArcInfo [Ding and Fotheringham (1992)] and ArcView [Scott and Lloyd (1997)], we think it would be worthwhile to implement the G_{ij} statistical measures of local spatial association in transportation related GIS, such as TransCAD.”

To immediately remedy what we failed to do we wrote a programme closely coupled to a GIS which is available along with a technical report (Berglund and Karlström, 1999b). We present the ideas behind this programme briefly in this section.

The concept of the weight matrix is central in spatial statistics. The weight matrix defines a neighbourhood of different objects in space in a way which is operational from a computational point of view. Ideally weight matrices should be derived from theory. However, this is not simple and one often has to turn to what is practical and convenient. One of the advantages of being inside a GIS and in particular in a transport-orientated GIS is that it is simple to construct weight matrices of different types. In Paper 2 we define a metric in flow space which is based on a single spatial weight matrix and in that sense analysis of flow data does not differ from analysis of scalar data.

In order to make some general tools for spatial analysis available in our transport modelling environment, we included a few simple routines such as: variable standardisation, row standardisation of spatial weight matrices, converting a continuous spatial weight matrix to binary, computing the first order spatial lag and computing the G_i- and G_i^*-statistics. Manipulation of weight matrices is applicable to analysis of both flow and scalar data while the other tools only apply to scalar data. Although the analysis of flow data is fundamental to transport analysis, analysis of scalar data is also relevant in a transport context.

To compute the G_{ij} as in Berglund and Karlström (1999a) we need a spatial weight matrix and flow data associated with geographical objects (compatible with the selected weight matrix). All data are defined from windows menus as for all programmes written within the framework of this dissertation (see Figure 11). Flow data are stored in a standard TC matrix. The resulting matrix of G_{ij}-statistics will also be in matrix form. However, this matrix will not be displayed directly. Instead, an array from the matrix will be written for display to a column in the relevant attribute table. The array is defined by a mouse click on the map. Any thematic map of the highlighted field will be updated every time a new area is selected. Once the matrix of G_{ij} is computed, the result can be displayed interactively by using the OD-button (see Figure 11) and clicking on areas on the map. This is an example of interactive analysis of flow data. As noticed several times in this introduction and in Paper 5 the computational performance of GIS programming languages is not very strong. The computation time may be considerable if the data set is large.

7 Scenario calculation

In Paper 5 an additional (and central) part of the model system is introduced – the scenario calculation module. This module contains what is often referred to as a 4-step transport model. Our model is

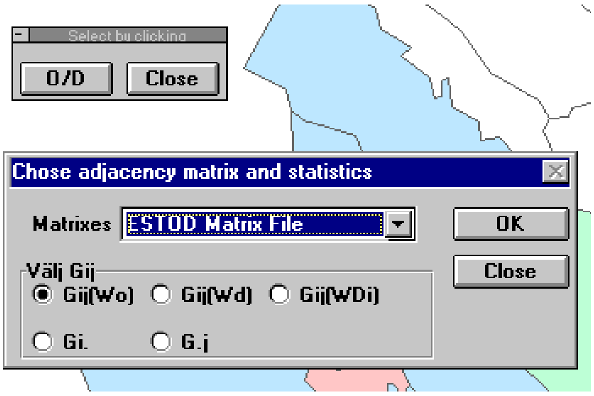
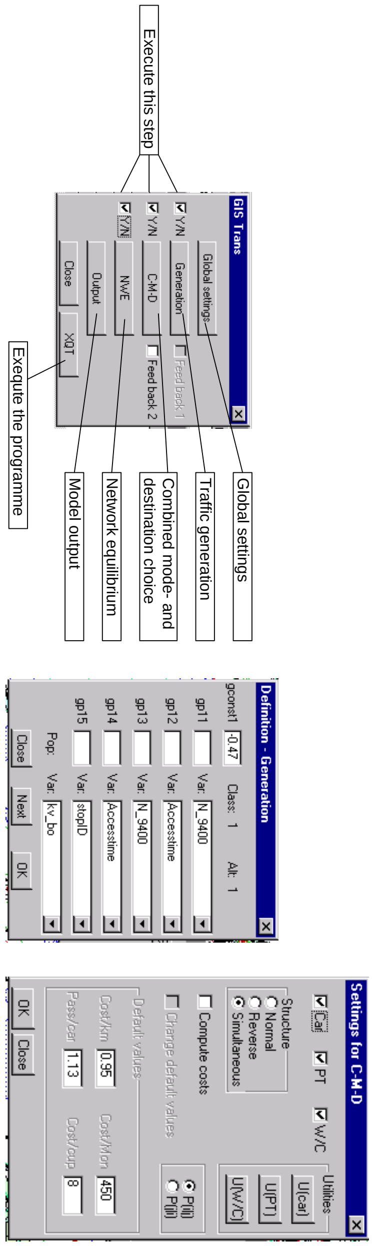


Figure 11: Dialogue box where the computation of the G_{ij}- statistics is controlled.

of traditional type including traffic generation, combined mode and destination choice and network assignment from TC. The programme takes care of definition of input data and feedback loops between the demand side and the supply side, see Figure 12(a). The ambition with the programme is to provide a flexible tool for scenario calculation and consequently all parameters and input data can be changed. The number of socio-economic groups can be altered under global settings (G settings in Figure 12(a)).

Zone-based input data can be stored in any type of TC data file and OD data are stored in TC matrices. The initial traffic generation step is entirely zone-based and the parameters are typed in a dialogue box, see Figure 12(b). The programme saves the parameter values from the previous session together with the corresponding values of the variables and these values will be used as default for the next session. Since the programme for estimation of the parameters of the combined mode and destination (CMD) choice model is produced "in house" the parameters are written to a parameter database that can be read immediately by the scenario programme. The CMD can be altered in four ways 1) number of modes (currently less than three) 2) nesting structure 3) computation of residential or job-based travel demand model 4) definition of travel cost variable (see Figure 12(c)). Currently the assignment algorithm cannot be substituted without modification of the source code.



(a) Initial dialogue box.

(b) Generation.

(c) Mode and destination choice.

Figure 12: Interface for scenario calculation.

The output will be standard TC files (matrices and tables). As default, output data are written to separate matrices with one layer for each socio-economic group.

8 Concluding remarks on the use of GIS in transport modelling

The main use of GIS in transport planning is data collection, management and display of model inputs and outputs. There is however also an emerging (research orientated) literature where the analytical benefits of GIS are taken advantage of. Examples of the latter are systematic design of traffic analysis zones compared to ad-hoc solutions and analysis of model output. The analysis of model output can be enhanced by inspiration from spatial statistics.

The use of GIS in transport modelling requires high quality of the data which results in a time-consuming start up process if no readymade database is available. If the problem at hand is to be solved once and the database will be of limited use afterwards, it may be appropriate to consider another modelling environment due to the initial cost.

The principles for model integration outlined in this introduction differ in a few important respects from many other similar approaches. Most important is the use of one commercial software and the decision to let all additional programmes be subject to modification to allow smooth integration. The programme environment used so far can be further developed to greater flexibility. One of the advantages of working with GIS is the open database provided by such systems which makes analyses and presentation of results simple and user friendly.

Based on the experiences from putting together the tools presented in this introduction, we can now summarise a few findings. In Table 2, a subjective grading of GIS with regard to some essential tasks in the modelling process is provided. Performance with regard to (wrt) user refers to efficiency in interaction between model and analyst. Computational performance refers to the efficiency of the software when it has received its instructions.

General geographic data management routines in GIS have the capability of handling any type of data that may be of interest for modelling purposes. Since GIS is essentially an information management system, it is not surprising that performance with regard to any

Table 2: Grading of GIS functionality in transport modelling.

Model task	Performance wrt user	Computational performance
Manage data	Very good	Good
Generate input data	Good	Good
Quality control of input data	Very good	Good/Poor
Hypothesis generation	Very good	Good/Poor
Model estimation	Poor	Poor
Model evaluation	Good/Poor	Poor
Generate forecasts/scenarios	Good	Poor
Presentation of results	Very good	Good

type of data handling is good. Visualisation of data is an important step in the modelling process, errors are detected and hypotheses generated. The drawback of GIS in this situation is observed when going beyond the simple query and visualisation tools in GIS. Then there is a need for own programming. Since a programming environment is provided by most software providers, this is not a problem for the advanced user. However, for the regular user the lack of built-in tools is clearly a shortcoming.

The core of the modelling process, the model estimation, often takes place outside GIS-based systems. This makes the modelling process far from being an interactive process between modelling, visualisation and evaluation. One of the expected benefits from working with GIS is the possibility of carrying out the modelling process in an interactive fashion. With the systems available today, this is not possible. Model evaluation suffers from the same problems as quality control of data, i.e. the tool box consists of simple tools and if you want to go further you have to do your own programming. If you do your own programming you will run into problems with computational performance if the size of your problem is large.

Generation of forecasts and scenarios involves organisation of lots of data both as input and out put. In this part we can rely on one of the strong sides of GIS. Data management and the interface programming facilities available in most GIS make this convenient with regard to the user. However transport modelling is computationally intensive and running times of several days are not uncommon. Computational efficiency is thus important. In this respect, GIS-based systems have a lot to prove. Our present application involving a four step transport model with feedback loops (see Paper 5) illustrates the performance gap between GIS and standard programming languages.

Presentation of results constitutes the analyst's means of communicating with the decision makers and the public. Even though this is not formal modelling, this part should not be underestimated and cannot be excluded from the modelling process.

References

Anderstig, C. and Mattsson, L-G. (1991) An integrated model of residential and employment location in a metropolitan region. *Papers in Regional Science 70:167-184*.

Bailey, T.C. and Gatrell, A.C. (1995) *Interactive Spatial Data Analysis,* Longman, Harlow.

Batty, M. and Xie, Y. (1994) Modelling inside GIS: Part 2. Selecting and calibrating urban models using ARC-INFO, *International Journal of Geographical Information Systems*, Vol. 8, NO. 5, 451-470.

Berglund, S. (1998) *Analysing Transportation Patterns in Cross-Border Contexts*.

Berglund, S. (1999) TECHNICAL REPORT: Computing Accessibility in GIS – Program Documentation.

Berglund, S. and Karlström, A. (1999a) Identifying local spatial association in flow data, *Journal of Geographical Systems*.

Berglund, S. and Karlström, A. (1999b) TECHNICAL REPORT:Computing G_i and G_{ij} – Program Documentation.

Berglund, S. and Lundqvist, L. (1998) *Barriers in transport models*, Paper presented at the 37^{th} Annual Meeting of WRSA, in Monterrey.

Black, W.R. and Thomas, I. (1998) Accidents on belgium's motorways: a network autocorrelation analysis, *Journal of Transport Geography*, Vol. 6, No1. pp. 23-31.

Choi, K. and Kim, T.J. (1997) A hybrid travel demand model with GIS and expert systems, *Computers, Environment and Urban Systems*, Vol. 20, No. 4/5, pp. 247-259.

Ding, C., Hanley, P. and Kim, J.T. (1994) ARCTRAN: ARC/INFO - TRANPLAN interface for transport planning, *Proceedings of the Fourteenth Annual ESRI User Conference*.

ESRI (1990) *Understanding GIS the ARC/INFO Method*, Environmental Systems Research Institute, Redlands CA.

Fotheringham, A.S. (1995) GIS-Based spatial modelling: A step forwards or a step backwards? Paper presented at GISDATA Specialist Meeting, Friiberghs Herrgård, Sweden, 14-18 June, 1995.

Geertman, S. and van Eck J.R. (1995) GIS and models for accessibility potential: an application in planning, *International Journal of Geographical Information Systems*, Vol. 9, No. 1, pp. 67-80.

Gehlke, C.E. and Biehl, K. (1934) certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association*, 29 (Supplement): 169-70.

Getis, A. and Ord, J.K. (1992) The analysis of spatial association by use of distance statistics, *Geographical Analysis*, 24(3), 189–206.

Hansen, W.G. (1959) How accessibility shapes land use, *Journal of the American Institute of Planners*, 25:73-76.

Heikkila, E.J. (1998) GIS IS DEAD; LONG LIVE GIS!, *Journal of the American Planning Association*, Summer 98, Vol. 64 Issue 3, pp 350-361.

Hägerstrand, T. (1970) What about people in regional science. *Papers of the Regional Science Association*, Vol 24, pp. 7-21.

Koenig, G. (1978) Accessibility and individual behaviour: accessibility indicators as a determinant of trip rate and urban development, Paper presented at the PTRC Summer meeting.

Kwan, M-P. (1998) Space-time and integral measures of individual accessibility: A comparative analysis using a point-based framework. *Geographical Analysis*, Vol. 30, No.30, pp.190-216.

Lenntorp, B. (1976) *Paths in Space-Time Environments – A Time-Geographic Study of Movement Possibilities of Individuals.* Lund, Lund studies in geography.

Lussier, R. and Wu, J.H. (1999) Development of a Data Exchange Protocol between EMME/2 and ARC/INFO, *http://WWW.inro.ca/publications/arcinfo_emme2.html*

Madden, J.F. (1981) Why women work closer to home, *Urban Studies*, 18, pp. 181-194.

Mattsson L-G. (1993) *Estimation of a Combined Mode and Destination Choice Model: User's Manual for the Estimation Tool Ester.* Working paper, Division of Regional Planning, KTH.

McCormack, E. and Nyerges T. (1997) What transportation modelling needs from a GIS: A conceptual framework, *Transportation Planning and Technology*, Vol. 21, pp. 5-23.

MDOT, (1998) *Transportation Case Studies in GIS*, Case Study 4: Maine Department of Transportation State-wide Travel Demand Model.

Miller, H.J. (1997) *Towards Consistent Travel Demand Estimation in Transportation Planning: A Guide to the Theory and Practice of Equilibrium Travel Demand Modeling.* http://www.bts.gov/tmip/papers/feedback/miller/toc.htm

Miller, H.J. (1999) Potential contributions of spatial analysis to geographic information systems for transportation (GIS-T), *Geographical Analysis, 31, 373-399.*

Murtagh, B.A. and Saunders, M.A. (1995) *MINOS 5.4 User's Guide* (Stanford University, Stanford 1995).

Nielsen, E.R., Nielssen, O.A. and Israelsen, T. (1999) BRIDGES to GIS, Deliverable D5 and D6 to the European comission.

Nyerges, T.L. (1995) Geographical Information Systems support for urban/regional transportation analysis, In *The Geography of Urban Transportation* (ed. S. Hanson). The Guilford Press, New York pp. 240-265.

Openshaw, S. (1996) Developing GIS-relevant zone-based spatial analysis methods, in *Spatial Analysis: Modelling in a GIS Environment*, P. Longley and M. Batty eds., Geoinformation International.

Openshaw, S. and Taylor, P.J. (1979) "A million or so correlation coefficients: Three experiments on the modefiable areal unit problem." In N. Wrigley (ed.) *Statistical Applications in the Spatial Science* London: Pion, pp. 127-144.

Ord, J.K. and Getis, A. (1995) Local spatial autocorrelation statistics: Distributional issues and an application, *Geographical Analysis*, 27(4), 286–305.

Spiekermann, K. and Wegener, M. (1995) Freedom from the tyranny of zone: Towards New GIS-based spatial models, Paper presented at the GISDATA Specialist Meeting, Friiberghs Herrgård, Sweden, 14-18 June 1995.

Thill, J-C. (2000) Geographic information systems for transportation in perspective, *Transportation Research Part C* 8 pp. 3-12.

Thill, J-C. and Horowitz, J.L. (1997) Modelling non-work destination choice sets defined by travel-time constraints, In *Recent Developments in Spatial Analysis*, Eds. Fisher and Getis.

USDOT (1998) *Transportation Case Studies in GIS*, Case Study 2: Portland Metro, Oregon - GIS Database for Urban TransportationPlanning.

Weibull, J.W. (1976) An axiomatic approach to the measurement of accessibility, *Regional Science and Urban Economics* Vol. 6 No. 4.

Weibull, J.W. (1980) On the numerical measurement of accessibility, *Environment and Planning A*, Vol 12, pp. 53-67.

Worboys, M.F. (1995) *GIS A Computing Perspective*, Taylor and Francis, London.

You, J., Nedovic-Budic, Z. and Kim, T.J. (1997) A GIS-based traffic analysis zone design: Implementation and evaluation. *Transportation Planning and Technology*. Vol. 21, pp.69-91.