Feeding the Animal. Can a GIS Satisfy the Hunger?

Scott Wilkinson
Engineer
Queensland Department of Transport

Abstract:

Transport planning models are data hungry

As urban conurbations throughout Australia grow, the coverage of our urban transport models has been creeping inexorably outwards to keep up with the urban sprawl Increasing complexity of questions asked of transport planners has also seen an increase in the testing capabilities built into our models.

In Queensland, traditional road based models have been enhanced to encompass public transport services in the first real effort in many years towards integrated transport planning While the United Kingdom trend has been to move towards more coarse zoning systems to address the questions raised by integrated transport issues. Queensland Transport has chosen to enhance existing methodologies by modelling mode choice behaviour in micro-detail, thereby requiring finer model zoning systems and networks.

The appetites (for input data) of the new breed of transport models have taken a leap beyond being hungry into a condition of gluttony

Faced with feeding these models, Queensland Transport have migrated into the world of Geographical Information Systems (GIS) to cope with the task of storing and displaying the huge amounts of data This paper outlines the initial observations relating to the coding of road, rail bus, and ferry infrastructure and service data into a Gis, the plans for converting and extracting the data for use within enhanced node based transport models. and future directions for the display and interrogation of model outputs via the GIS medium

---

Contact Author:

Scott Wilkinson
Queensland Department of Transport
GPO Box 1412
BRISBANE Q 4001

Telephone: (07) 834 2420
Fax: (07) 834 2201
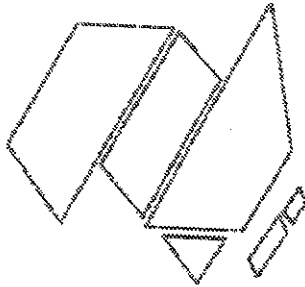
1.0   INTRODUCTION

Be careful. It BITES! This may be the warning that Queensland Transport's transport planners may soon find plastered over the procedural manuals for the Department's new breed of travel evaluation and forecasting models

Queensland Transport has reached a turning point in its development of travel models and has taken the leap, or perhaps, fall into the world of Geographical Information Systems (GIS) to satisfy the hunger of these models for input data.

Background

Since the creation of the current Queensland Department of Transport in 1990 (from the former Departments of Main Roads, Harbours and Marine, and Transport), a growing emphasis on integrated transport planning has emerged. Public awareness of the effects of increasing transport usage on the urban environment and the benefits of humble bus and train is increasing The result is a push for the urban and transport planners to reconsider the important interrelationship between land use and transport.

Changes in the focus of planning have been most noticeable in the South-East Region of the state where the expected population growth has potential to burst the urban fabric at the seams. Techniques for planning for this growth were developed for the assessment of Brisbane's transport needs and the lessons learnt have since been generally applied throughout the state.

The relative importance of the roles of private and public transport has changed since Brisbane's first study into travel demand in the mid 1960's by Wilbur Smith and Associates (1965) Since then, escalating public awareness of the environmental and sociological impacts of transport, with the reduction of personal mobility caused by road congestion may be producing a reduction of private car use. Over the same period (since 1965) our assessment methodologies associated with transport infrastructure provision have also changed considerably In parallel, computer modelling of the demand and distribution of travel has boomed due to the ever increasing power of the personal computer, allowing investigation of the personal travel route and mode choice in ever increasing detail

Changes in the focus of transport planning has resulted in the current development of a set of more detailed modelling techniques in the Integrated Travel Forecasting and Evaluation Method (ITFEM). The major development within ITFEM will be the treatment of the choice between private and public modes of transport While this endeavour has been attempted before, its treatment in ITFEM is aimed at making

detailed simulation of the factors influencing the public's choice between use of public transport versus private transport.

A prototype version of ITFEM is being developed on a single study area though it will have application to all urban areas in the state.

How does a GIS fit in?

The answer is simple. The models being created are developing appetites for data.

While the modelling of public transport use is not new, detailed modelling of the choice between private and public transport modes in detail has for the last decade been avoided (in Queensland). One primary consideration when assessing whether public transport is a realistic option for any trip, is the access to and from the public transport service at either end of that trip. Modelling of this access problem is the most difficult part of creating a realistic integrated transport evaluation method. Within ITFEM the computer representations of the road and public transport system has therefore been aimed at the most detailed level within practical limits, to incorporate some measures of the many public transport access options.

Computer models were developed for the first transportation study in 1965 (Wilbur Smith and Associates (1965)) and have been progressively enhanced until the completion of the Brisbane Traffic Study (Brisbane City Council (1989)). These models were, however, largely based on representations of the road transport system. As part of the new breed of model, a full representation of the total transport system will be created, including infrastructure and service details for urban road, rail, bus, and ferry transport.

The scale of the data collection, collation, and storage process, because of the level of detail now required, has increased many times. Previous data storage methods were a combination of documentation in technical papers and filing cabinets, and bewildering arrays of computer files on many storage media. In short - a nightmare! It was recognised early in the development of the ITFEM methodology that a radical change in the collation and storage method for input data was needed. Considering that many data sets upon which the computer models rely are spatial, the decision to move towards a GIS was clear cut.

The ITFEM project is developing on many parallel paths, one of which being the encoding of the data associated with Brisbane's transport infrastructure and services. While the GIS will have many future benefits, this paper will focus on the change in procedures already experienced in feeding the transport model animal with input infrastructure and service data.

2.0   DATA FOR A TRANSPORT MODEL

Data required to describe the transport infrastructure can be divided into three main sections, the zoning system, road infrastructure, and public transport infrastructure and services   A discussion on each of these follows

Transport Zoning System Data

The study area used for the ITFEM prototype covers the Australian Bureau of Statistics (ABS) Brisbane Statistical Division.   Analysis of the travel from one area to another is achieved by dividing the total area into smaller portions, termed transport zones, each in turn assessed as to the amount and distribution of trips entering and leaving.   While these zones do not form a recognisable part of the transport infrastructure, their modelled access to the transport network is important.   The size and location of these transport zones influence the level of detail required in the modelled representation of the road and public transport networks.

The demographic data associated with the estimation of the amount of travel is aggregated into groups of households, primarily to maintain the confidentiality of personal data   Much of the data collected by the ABS is grouped into various levels of geographical areas to prevent association of any particular data item with a specific household or location.   In keeping with the aim of creating a model capable of reflecting mode choice, the zoning system was based on the smallest aggregation practical, the ABS Census Collection District (CCD)   The bulk of the data collected by ABS and Queensland Transport will be aggregated to this level of aggregation for use in ITFEM

Other data sets that have been collected in the past have been aggregated using several other zoning systems.   One of these is the information collected in the Census by ABS on the mode of travel from home to work.   Aggregation of this data was not done to CCD level   Queensland Transport was responsible for the development of the travel zoning system used to aggregate the destination (work location) data.   Most travel zones are groups of or single CCDs, but in some areas (notably the Central Business District) the CCDs were split into a finer than CCD zoning system   Splitting CCDs was kept to a minimum to avoid the difficult task of apportioning ABS data into smaller areas.

To maintain compatibility with the bulk of the ABS data (including the Journey to Work data), the boundaries from both zoning systems have been amalgamated to form the ITFEM zoning system   The size of the zones are therefore equal or smaller than the CCDs   Some special land uses have also been specifically catered for in the zoning system on an individual basis (eg shopping centres, industrial estates, freight terminals, hospitals, large tourist attractions, etcetera).

The total number of transport zones in the prototype (Brisbane) study area is likely to approximately 2500 indicating, as a rough measure, the intended improvement on the sensitivity of the model, when compared to the previous total of 242 internal zones used in the most recent comparable study by the Brisbane City Council (1987).

Road Infrastructure Data

Statistics collected for coding into a road infrastructure network can be grouped into (a) the physical attributes of road segments that relate to their capacity to move traffic, (b) the surrounding interaction of land uses with the traffic on the road segments, and (c) the physical attributes of intersections that affect their capacity to transfer traffic from one road segment to another.

Treatment of traffic movement on the road network within ITFEM will be at a strategic level. The treatment of intersections will not be a surrogate for a more detailed traffic analysis. ITFEM will be a planning tool, not a design tool.

Collection of the road inventory data, was done by the staff in the local authorities and Queensland Transport districts and passed on to the ITFEM team on hard copy maps. The process used to transfer this information to the GIS will be discussed later and a comparison made with the traditional coding procedures.

Public Transport Infrastructure and Services Data

A separate project being undertaken concurrently by Queensland Transport is aimed at creating a GIS based public transport inquiry system. The data collected for the Integrated Public Transport Information Service (IPTIS) will be used directly to create the public transport infrastructure and service database required for ITFEM. The information stored in the IPTIS system to be extracted for ITFEM includes the route, fare structure, connectivity, and timetable data for all of the rail, bus, and ferry services operating in the study area.

The conversion of this data for use in ITFEM will not be discussed in as much detail as the road infrastructure process. The conversion of this data is, however, of some importance to the future direction of the use of a GIS in Queensland Transport.

3.0   PREVIOUS CODING METHODS

Discussion of the benefits of using a GIS to code and store transport model related data must logically begin with a description of the methods used prior to the GIS approach. This comparison will only be made for the procedure used to code a road based

transport network although the points made would also apply to the coding of public transport raw data

Coding Procedure

The coding procedure can be broken down into a series of major tasks. The following description of the conventional coding tasks has been deliberately simplified for this discussion:

(a)   Preparation of base maps is often a crucial task, as the scale of the map(s) must allow accurate measurement of distances and yet not be an unwieldy size.

(b)   Identification of the roads to be modelled and an appropriate zoning system is done, often the product of compromises between the accuracy desired of the model and the realities of computer run times, data collection costs, and anticipated calibration effort.

(c)   The road network and zoning system are converted into a series of road "links" jointed by uniquely numbered intersection "nodes" and each zone is represented by a single "centroid" and connected to nodes by one or more links.

(d)   A series of link records (identified by the node numbers at each end - Anode and Bnode) can now be entered into a computer file that describes the connectivity in the network. Each link's length is measured, often by scaling from the base map, and entered into the link record file along with the behavioural characteristics calculated from the real data (collected from various authorities as described earlier).

(e)   An optional addition to the data set at this point is a digitised set of node and zone centroid coordinates.

(f)   Building the network is the name of the process given to the checking of the connectivity of the network and the checking of the link information as described by the link record file. Coding of a new network is generally carried out initially by creating ascii link records and followed by checking using graphical viewing/editing programs.

(g)   Many checks are now done on the coded link parameters, starting with checking that the data coded for each link is correctly entered, and eventually checking that the shortest paths through the network as selected by the model are sensible (based on free flow link travel speeds/times entered into the coded data).

Advantages of the Traditional Approach

Considering that this paper is being written to comment on an improved method of coding transport infrastructure network data, a discussion of the advantages of the superseded coding method may seem out of place. The reality is that many of the procedures perfected over the last twenty years in Queensland are still of merit.

One definite side effect of coding a network using the above methods is that the transport planner becomes intimately familiar with the network that he or she has coded. It should be pointed out that the coding process is only a small part of the total transport planning process. Following the coding of a transport network, the ominous task of calibration of the travel model begins. The process of convincing the modelling software to mathematically replicate the travel movements, observed in various surveys, often becomes the subject of nightmares. At this point, if the transport planner is not familiar with the network, both the real network and the coded representation, misunderstanding of discrepancies can often occur. Strictly speaking, the calibration process should largely involve fine tuning of trip estimation and distribution procedures. Changes to the coded network should be minimal, beyond some checks of the network model's route choice behaviour under various levels of traffic loading.

Given that the accuracy of the coding process can affect the eventual calibration of a transport model, the majority of the thinking and checking involved in coding is often conducted personally by experienced transport planners, thereby promoting a more careful approach to the whole coding process.

Disadvantages of the Traditional Approach

Familiarity with the data in a network model is good, boredom is not. Coding is a tedious task and coding of a road network model for an area like the greater Brisbane area is enough for the average transport planner. The added detail required by ITFEM and the fine level of detail (ie CCD zoning system and matching road network) would induce insanity using the traditional approach to coding. The monotonous process of associating data to a system of numerically numbered nodes and links would be a recipe for error. No amount of checking would catch every error in a network of that size.

The procedure described above is also a very slow process. Each individual link has to be manually assessed to categorise it into a particular type of road segment, its length measured, and converted to a computer record and key punched. Measurement of the link lengths, usually by scaling from the base maps, can account for an appreciable proportion of the total coding time.

A serious problem, not directly associated with the coding procedure, is the potential loss of the original data and decision process used to create the original coded link records. Often the calibration of a network model will herald the start of a period of continued testing and modification of the network model, sometimes for many years. Over this period, typically the only data maintained in something like its original form are the computer files required by the transport modelling software. Unfortunately, the original data upon which these records are based tend to remain in their original hard copy form. While the hard copy maps stored in plan drawers may never be lost, the maps will rarely be used to check that modifications to the network in later analyses adhere to the same criteria originally used

4.0   THE CREATION OF A GIS DATABASE

Most of the improvements to coding procedures implemented through the GIS tool relate to the process used to code the road network infrastructure. Coding and preparation of the data relating to the public transport network and transport zones was done via different processes due to the availability of good electronic versions of source data The coding of the road network, public transport infrastructure and services, and the transport zone boundaries will each be discussed individually below.

Road Network Coding with a GIS Package

Revising any procedure consists of duplicating the portions of the process that work well and replacing the portions that do not The design of the GIS process for road network infrastructure coding was based heavily on this aim, in particular by:

(a)   fostering the familiarity with the road network (minus the boredom) by providing recognisable geographical features on the GIS screen,

(b)   removing some of the laborious tasks involved,

(c)   creating GIS applications based on the experience of the transport planner to allow the use of less experienced staff for the bulk of the data input,

(d)   representing the data on the GIS screen using an annotation legend similar or identical to that used on the hard copy data maps supplied by the various road authorities,

(e)   using the GIS to produce hard copy output at the base map scale for ease of data checking, and

(f)   storing the "real" infrastructure data in the GIS before modification (by a computer based process) into link records as required by the transport modelling software

Many of the steps involved in the coding process from the traditional approach were still carried out for the ITFEM data. The differences lie in the changes within those steps afforded by the use of the GIS tool.

A GIS is not of much use without data. While the aim of the coding procedure was to create a group of GIS data sets, some source information was required as guiding input. The source information collected for use in the road network coding process has been listed in Table I.

A series of base maps were prepared as the initial task in the coding process. A GIS was never and will never be a substitute for an accurate set of maps. A series of 1:25000 scale topographic maps, updated where out of date, proved most appropriate

The controlling feature determining the detail to be coded into the ITFEM road network was the transport zoning system. The decision to model public transport required that the zone size be equivalent to ABS CCDs. As a guide to the impact of this decision, plots were produced by the GIS on transparent film at the base map scale. Comparison of these overlays with the base maps allowed the quick identification of the road segments to be included in the network model. A hand drawn set of link maps were then prepared, copied and sent to the local authorities for collation of relevant data. Annotation legends were prepared to standardise the responses.

While awaiting the data, the skeleton of the road network was created within the GIS. The roads of interest were selected from an existing GIS map of road centre-lines. The numbering of nodes, digitising of node coordinates, naming of links with the "A-node B-node" tag, and measuring of link lengths were all automatically done by the GIS programs developed. Checks were made by creating further overlays and comparing the GIS links map with the hand drawn links map. This comparison of link overlays proved highly successful in identifying missing and extra links

Upon return of the data from the local authorities, each of the data items was associated with the GIS road segments and stored in an external relational database via a live connection to the GIS. The programs developed for this task allowed the GIS operator to select links by direction along a series of links, or by area, and change/enter the data. The GIS screen was then updated with the data, represented graphically in the form on the returned hard copy maps for instant confirmation and checking. Further overlays were also produced as an independent check

Table 1   Input GIS Data Sets

Data Set (source)	Application within GIS Coding Process
1991 Census Collection Districts (Australian Bureau of Statistics)	The 1991 CCDs were the starting point for the development of the transport zoning system.
Street Centre-line Maps (ERSIS)	This data set was used directly to give the alignment of the majority of road links entered.
Digital Cadastral Data Base (Qld Department of Lands)	The property boundaries were used to give the location and size of special traffic generators. Road reserve boundaries were also useful for digitising road segments not shown on the street centre line data (eg freeway ramps and new roads)
Integrated Public Transport Information Service (Qld Department of Transport)	The IPTIS database contained both GIS maps and relational database tables. This information was used directly in a reformatted form to produce the public transport infrastructure and service data
Miscellaneous Queensland Transport Data (Qld Department of Transport)	The development of GIS applications is currently progressing in many branches of the Department. The GIS maps created for cartographic production of the Department's district maps were very useful as background view data throughout the coding process and included coastlines, rivers, railway lines, declared roads, and suburb names.

Once the data from the local authorities was entered into the GIS, the data was converted into a form suitable for export to a transport modelling package. This process was achieved within the relational database and was based on the traditional procedures used to convert "real" road statistics to link records. Some further data items were required to facilitate this conversion process as a substitute for the hands on ability to

visually assess the data. Creation of the correct files for the transport modelling programs was then simply a matter of exporting the correct fields from the database tables.

Public Transport Infrastructure and Services

Coding of the public transport data was achieved by one of the rare occurrences of good fortune/planning within the Department. A separate project under development within Queensland Transport, independent of the development of the prototype GIS, was the Integrated Public Transport Information Service (IPTIS). The project was undertaken to update and expand an existing bus information hot-line to encompass all public transport information within a GIS based system. Details including the service type and operator, the fares, timetables, routes and stops were being collected and transferred into a GIS for all of the public transport services in south-east Queensland region.

Conversion of the IPTIS database for use in ITFEM was achieved through an automated computer process. The route maps were converted to infrastructure links and added to those of the road network links and any new nodes created as required. Timetable information was modified to give an average frequency (for representative periods of the day) and fare and stop information was reformatted for export to the transport modelling software.

Creation of programs to encode raw public transport information manually has not yet been developed. Programs to enter this data into the GIS, for areas that are not included in the IPTIS information, when developed will similar to those already developed for the road network data.

Transport Zoning System

Yet again, existing databases were used as aids for the development of the GIS transport zone maps. The zoning system was developed initially from the GIS maps of the CCD boundaries supplied by the Australian Bureau of Statistics, however substantial modifications were made to the CCD boundaries for two purposes.

Data was collected in the 1991 Census of Population and Housing relating to the Journey to Work (origin, destination, and mode). The zoning system used for the aggregation of the data in south-east Queensland was developed by Queensland Transport. Generally the boundaries of the Journey to Work zones are similar to those of aggregated CCDs. Some exceptions are subdivided further for more meaningful Journey to Work data. These boundaries were also adopted for the ITFEM transport zones to maintain compatibility with the data. The locations of the split boundaries

were digitised directly into the GIS zoning map using property boundaries as a guide.

Other additions to the GIS transport zone maps were for the purpose of identifying special traffic generators. The GIS property boundary maps were used directly to locate the boundaries' of these special zones.

Connection of the zonal areas within the transport modelling software is achieved via special links called "centroid connectors" that connect the centres of the transport zones to one or more network nodes. These were added using an extension of the procedure used to enter the road network links. The centre of each zone was identified and a path "driven" along the street centre line maps to the access point(s) for the zone. This had the advantage of allocating a realistic average length to the access path to locations within the zone.

5.0   OBSERVATIONS ON THE CREATION OF THE GIS DATABASE

Let's face it. Using a GIS is fun! At least it seemed like fun in comparison to the monotonous traditional coding procedures, and the GIS is more efficient and accurate for the task of coding. The charm and accuracy of a GIS are both attributable to its ability to produce pretty pictures on a live screen. The operator operating the GIS is instantly shown the fruits of his labours and so can check the data visually upon input. This leads to a more productive and accurate data entry process. Some more specific observations on the use of a GIS for coding data are reviewed below.

Source Information

Reference has already been made to the use of source information to create the GIS maps. This existing information was used in various stages of the coding procedure for different reasons. Some was converted and used directly in the ITFEM maps and some was only used as background layers to facilitate location of the input infrastructure statistics. Table 1 below shows the useful information from these bodies, the source, and its application within the coding process.

The use of these data sets was not without some heartache. Using third party GIS data has shown that some care must be exercised when converting from foreign GIS software. Particular care must be taken when importing area features to ensure that the correct names are associated with those areas (eg Census Collection Districts).

Of course, there were some hefty charges for the supply of the data. This is not surprising, considering the effort involved in its collection and encoding. The cost, when compared to creating the data sets as part of the development of the transport GIS,

was not too excessive The importance of obtaining the correct information was therefore obvious, and so demanded care and judgement

The "A-node B-node" Paradigm

Coding transport planning network models can form bad habits. So ingrained were the procedures used prior to the GIS, that the structure of the GIS was based around the numbering of nodes and the "A-node B-node" link naming convention. This seemed logical. Or was it?

The only reason that the node numbering convention was used in the traditional coding procedures was to tell the transport modelling software which links connected, the locations of those intersections, and to associate the information with the correct links. It seems a little superfluous to be adhering to these numbering systems when the GIS can do these checks by itself.

All is not lost The ultimate aim of the GIS procedure was to create a data entry and storage system for transport data. The nodal system of network description was still required for export to the external transport modelling software. Storing the data in the GIS in the node based system was not a total loss, in fact, it worked quite well The drawback was that a great deal of effort was expended in creating a GIS system that created nodes and named links in the "A-node B-node" convention. The GIS programs would have run much faster without the overheads of creating the correct link and node names

Each point, line, and area feature in the GIS maps is assigned a unique identification number upon creation, independent of the name of the tag given to it by the GIS operator or programs. The tag of the links could be then be used to store a much more practical reference Perhaps, something as human as the street name could be used!

Efficiency

The development of the GIS infrastructure coding programs was somewhat of a fast tracking exercise The programs were written as progress was made, and the tasks became more fully defined This has both good and bad effects on the development. The programs developed were given a "birth by fire", resulting in a more robust system Of course, the penalty was that some of the development could have benefited by a little more experimenting Despite this rushed development, the result still remains; the process worked and a huge transport infrastructure and service information GIS data set has been successfully created.

The efficiency of the GIS for coding transport data is difficult to assess. On occasions the coding staff were required to change tasks to allow program writing to catch up. A noticeable trend was that each program, once written, allowing the staff to enter that particular data item at a remarkable rate.

The efficiency of the total process of storing all the information in the GIS and the automated conversion to the transport modelling link records cannot be fully assessed at this point. The benefits of having this data permanently stored in the GIS will be an outcome that time will soon illuminate. A true assessment of the efficiency of the GIS for data entry will soon be possible when it is used for the next major transport infrastructure coding project (i.e. without the delays imposed by the GIS program development and testing)

6.0 OPPORTUNITIES BEYOND INFRASTRUCTURE CODING

The use of the GIS is not intended to cease with the development of the infrastructure coding facility. In fact, this can be viewed as the beginning of a wonderful friendship. Launching into the world of GIS for transport planning will have many benefits for Queensland Transport. Coupled with the transport modelling software, the GIS will provide many new analysis possibilities never before available to planning agencies in Queensland.

Transport planning procedures are designed largely to assess the effects that land use changes have on the requirements for transport infrastructure provision. The reciprocal effect that transport has on shaping the land use patterns in urban areas has largely been left in the "too hard basket". The abilities of the GIS software to assess spatial distributions of land use data may assist in analysing this important subject.

As the global use of GIS grows, the compilation of data sets specifically for use in ITFEM studies may be a far more simple process. Involvement of government bodies in the creation and maintenance of GIS databases at all levels is already gaining momentum. The ability to simply transform, or directly use another GIS database has already demonstrated by this project in the use of the IPTIS data. This practice must eventually result in improvement in the accuracy of data used for transport assessments. Coding of a transport infrastructure in the future may simply be a matter of extracting/converting relevant data via electronic connections to other agencies' databases. The assessment results would also be available for use by the wider forum via the same process.

7.0 CONCLUSIONS

Queensland Transport has launched into the wonderful world of GIS. The growing appetite of the new breed of transport models for input data has been sated for the moment by a modification of a GIS to allow data to be coded directly into a Geographical Information System based database.

Initial assessments of the possibilities made available through the development of a GIS based infrastructure coding procedure have been favourable. Improvements over the traditional procedures were made in the areas of accuracy and efficiency.

Existing data sources were used to reduce the initial workload associated with converting other agencies' data. The potential for future transport data transfer was demonstrated through the use of the Integrated Public Transport Information Service database. The possible direction for the future data transfer procedures was exploited successfully.

Further development of the potential of the GIS for coding procedures and analysis will be investigated in the future as part of the future planning function of Queensland Transport.

References

Brisbane City Council (November 1989) *Brisbane Traffic Study*.

Wilbur Smith and Associates (November 1965) *Brisbane Transportation Study*, 2 Volumes and Summary.