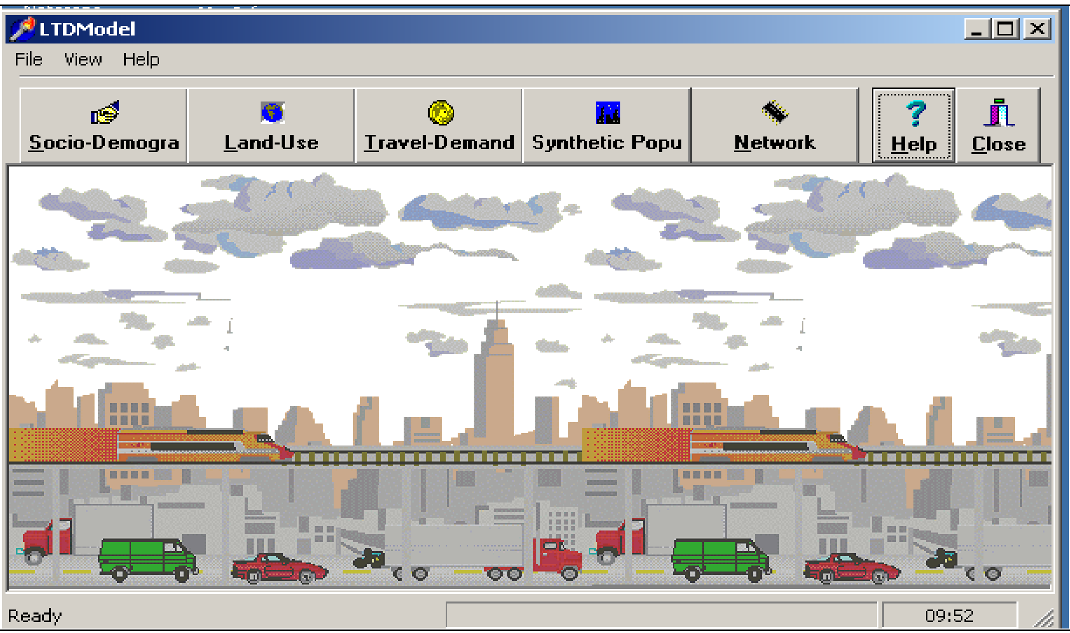


A microsimulation model of travel behaviour for use in urban transport corridor analysis

Min Xu, Transport Systems Centre
University of South Australia
Michael A P Taylor, Transport Systems Centre
University of South Australia
Steve Hamnett, Transport Systems Centre
University of South Australia

Conference paper
Session IV

Moving through nets:
The physical and social dimensions of travel
10th International Conference on Travel Behaviour Research
Lucerne, 10-15. August 2003

A microsimulation model of travel behaviour for use in urban transport corridor analysis

Min Xu
Research Scholar
Transport Systems Centre
University of South Australia
Adelaide SA 5001
Australia
Tel: +61-8-8302-2212
Fax: +61-8-8302-1880
E-mail: min.xu@postgrads.unisa.edu.au

Michael A P Taylor
Professor of Transport Planning and Director
Transport Systems Centre
University of South Australia
Adelaide SA 5001
Australia
Tel: +61-8-8302-1861
Fax: +61-8-8302-1880
E-mail: map.taylor@unisa.edu.au

Steve Hamnett
Professor of Urban and Regional Planning
Transport Systems Centre
University of South Australia
Adelaide SA 5001
Australia
Tel: +61-8-8302-2297
Fax: +61-8-8302-1880
E-mail: steve.hamnett@unisa.edu.au

Abstract

In recent years, many planning initiatives have focused on the relationship between land use and travel behaviour, with the aim of reducing car use and encouraging forms of more 'sustainable transport', together with the application of Travel Demand Management (TDM) principles. Travel behaviour is often highly complex and, similarly, TDM measures and the relationship between land use and travel behaviour are also multifaceted, often requiring 'micro-level' analysis and decision-making for successful implementation. Activity-based travel demand microsimulation models are based on the notion that travel is a derived demand from the need to pursue activities that are scattered in time and space. They focus on the analysis of the individual's travel behaviour resulting from his/her participation in a sequence of activities, and possibly influenced by interactions with other people (such as fellow members of a household). Microsimulation models are disaggregate and dynamic in nature, offering effective and practical tools for the representation of complex travel behaviour, and are thought to be the most suitable models for such purposes.

This paper describes the development of a microsimulation model of household travel, responsive to alternative transport networks, land use patterns and densities, and socio-economic characteristics of individuals. The model or is being tailored for use in multimodal urban transport corridors, with the study area focusing on the north-west metropolitan corridor of Adelaide in South Australia. The primary objective of the model is to be able to provide estimates of the likely usage of different travel modes, with an emphasis on including features which can determine the likely usage of public transport and non-motorised modes, in particular where the incorporation of alternative land use patterns supportive of such modes is envisaged.

The model system or is being developed in the object-oriented software design paradigm. Four sub-models – socio-demographic, land use, activity-travel simulator, and network – are formulated into the model system. All of the sub-models are loosely connected with each other and can be used independently. The socio-demographic model forecasts the evolution of individuals and households over time from an assumed base year. The land use model simulates future land use development based on current land use and possible changes in land use policy. The network model estimates travel movements (including public transport) in corridor networks. The activity-travel simulator, which is the key component of the model system, generates activity/travel patterns and forecasts travel demand, as well as scheduling activity and travel demonstrated by individuals and households in time and space.

The activity-travel simulator is based on the concept of decomposition of a daily activity/travel pattern into segments to which certain aspects of observed activity/travel behaviour are matched. The activity/travel pattern can be deconstructed into various segments, in terms of activity type, activity duration, activity location, and mode choice and transition. Initially, the activity-travel simulator assigns a 24-hour activity/travel pattern for each member of a household by using a set of representative activity/travel patterns[1]. Then, from the start time point, the simulator simulates an activity type, its duration and location based on the observed activity distribution aligned with the activity/travel pattern and time step. A new activity and its characteristics are simulated at the finish time of the preceding activity based on the activity/travel pattern. This process continues until the complete 24-hour activity/travel pattern is determined for each individual. The representative activity/travel patterns are derived from the observed activity/travel behaviour of the individual and household, and are categorised into different groups according to the socio-demographic characteristics of the household. Monte Carlo simulation is employed to generate the daily travel patterns. A nested-logit model is used to estimate the mode choice within the simulator.

In addition to the four sub-models above, a synthetic population generator is also one of the major components of the microsimulation model system, and this paper outlines the process developed for this study. The individual traveller is the basic entity for simulating travel behaviour in microsimulation travel demand models. Whilst it would be desirable to work with an entire population of individuals (or households) within the microsimulation model, a complete state estimation is basically impossible, because some data is systematically missing due to

[1] These patterns have been obtained from analysis of the 1999 Metropolitan Adelaide Household Travel Survey (99MAHTS), which provides detailed data on individuals' travel and activity participation.

privacy issues. Therefore it is necessary to use alternative techniques, such as the development of a synthetic population.

Most of the microsimulation models developed to date, whether in a short-term or in a medium and long-term forecasting application, have required a representative sample of population. However, an updating procedure of the representative sample must be included in a medium or long-term application. One of the approaches to deal with this is to use a survey sample as representative of a whole population, such as used in the development of the Activity-Mobility Simulator (AMOS) (RDC Inc., 1995). An alternative approach is to use census data together with a micro-data sample to generate a synthetic population as representative of a study area (for example, Beckman *et al*, 1996). The approach adopted for this study is the synthetic household method. The basic concept involves using a GIS model of the corridor, which includes dwelling types and locations. The synthetic population (households and individual household members) is generated at the census collector district (CCD) level using the Australian Bureau of Statistics' (ABS) census data and one per cent household sample data. Detailed land use data is available, including the cadastral database for the study region, which identifies the location and attributes of all dwellings and other building and site developments in the region. The land use data is then incorporated into the model by assigning the synthetic population to the available dwellings in the CCD to match the known demographic characteristics of the population. Monte Carlo simulation is employed to generate the synthetic households. For effective use of the model, several different synthetic populations are sampled, to allow validation of the model with observed behaviour. Further synthetic populations, for alternative planning scenarios for the corridor, may also be generated, to test their likely travel outcomes.

The 1999 Metropolitan Adelaide Household Travel Survey database (99MAHTS) is used for activity/travel simulator estimation and validation. The 99MAHTS database includes all the information of activities (in-home and out-of-home) and trips that all individuals had taken in two consecutive days for all households in the survey sample. It also includes respondents' socio-demographic information, such as age, sex, occupation, employment status and personal income.

The paper concludes by showing some initial results from the model and indicating its future development.

Keywords

Travel demand modelling, microsimulation, synthetic population, activity scheduler, International Conference on Travel Behaviour Research, IATBR

Preferred citation

Xu, Min, Michael AP Taylor and Steve Hamnett (2003) A microsimulation model of travel behaviour for use in urban transport corridor analysis, paper presented at the 10th International Conference on Travel Behaviour Research, Lucerne, August 2003.

1. Introduction

In recent years, many planning initiatives have focused on the relationship between land use and travel behaviour, with the aim of reducing car use and encouraging forms of more 'sustainable transport', together with the application of Travel Demand Management (TDM) principles. Travel behaviour is often highly complex and, similarly, TDM measures and the relationship between land use and travel behaviour are also multifaceted, often requiring 'micro-level' analysis and decision-making for successful implementation. Activity-based travel demand microsimulation models are based on the notion that travel is a derived demand from the need to pursue activities that are scattered in time and space. They focus on the analysis of the individual's travel behaviour resulting from his/her participation in a sequence of activities, and possibly influenced by interactions with other people (such as fellow members of a household). Microsimulation models are disaggregate and dynamic in nature, offering effective and practical tools for the representation of complex travel behaviour, and are thought to be the most suitable models for such purposes. To the present, however, there have been few reports of fully implemented microsimulation models.

This paper describes the development of a microsimulation model of household travel, responsive to alternative transport networks, land use patterns and densities, and socio-economic characteristics of individuals. The model is being tailored for use in multimodal urban transport corridors, with the study area focusing on the north-west metropolitan corridor of Adelaide in South Australia. The primary objective of the model is to be able to provide estimates of the likely usage of different travel modes, with an emphasis on including features which can determine the likelihood of people using public transport and non-motorised modes, in particular where the incorporation of alternative land use patterns supportive of such modes is envisaged.

The model system is being developed in the object-oriented software design paradigm. Four sub-models – socio-demographic, synthetic population generator, activity-travel simulator, and network – are formulated into the model system. All of the sub-models are loosely connected with each other and can be used independently. The activity-travel simulator, which is the key component of the model system, generates activity/travel patterns and forecasts travel demand, as well as scheduling activity and travel demonstrated by individuals and households in time and space. The activity-travel simulator is based on the concept of decomposition of a daily activity/travel pattern into segments to which certain aspects of observed activity/travel behaviour are matched. The activity/travel pattern can be deconstructed into various segments, in terms of activity type, activity duration, activity location, and mode choice and transition. Initially, the activity-travel simulator assigns a 24-hour activity/travel pattern for each member of a household by using a set of representative activity/travel patterns[2]. Then, from the start time point, the simulator simulates an activity type, its duration and location based on the observed activity distribution aligned with the activity/travel pattern and time step. A new activity and its characteristics are simulated at the finish time of the preceding activity based on the activity/travel pattern. This process continues until the complete 24-hour activity/travel pattern is determined for each individual. The representative activity/travel patterns are derived from the observed activity/travel behaviour of the individual and household, and are categorised into different groups according to the socio-demographic characteristics of the household. Monte Carlo simulation is employed to generate the daily travel patterns. In addition to the activity-travel simulator, another major component of the microsimulation model system is the synthetic population generator. The basic platform for the approach is a GIS model of the corridor, which includes dwelling types and locations. The generated synthetic households have their associated dwelling and similar socio-demographic characteristics in a study area. Again, Monte Carlo simulation is used for this approach.

The paper is structured as follows: in the next section, a brief description of the project background and review of activity-based travel demand models is given, with a specific focus on micro-simulation approaches. Section 3 provides an introduction to the data sets used in the model system development, including the Australian Bureau of Statistics (ABS)1996 census data, the 1996 Census Household Sample File data[3], the 1999 Metropolitan Adelaide Household Travel Survey (99MAHTS) data and the geographical zone data. The overall model design is

[2] These patterns have been obtained from analysis of the 1999 Adelaide Household Travel Survey (99MAHTS), which provides detailed data on individuals' travel and activity participation.

[3] This data file is a one per cent sample of complete individual household records from the 1996 census for the Adelaide Statistical Division.

described in Section 4, and the data processing methods carried out for the various data sets are then presented (Section 5). Section 6 describes the methods employed in obtaining a synthetic population, together with a validation of the results. The activity-travel simulator is then introduced and some initial results are demonstrated (Section 7). The final section provides a discussion and summary, including an outline of future work.

2. Background

Urban transport corridors provide a convenient means of describing multi-modal or multiple-route trip movements in an urban area. Corridor analysis is of increasing importance in transport systems modelling, especially in the planning and evaluation of multi-modal urban transport systems. The interest in corridor studies is based on three main issues:

- the current emphasis on incremental improvements to existing transport infrastructure, which reinforces a need for tools for sensitivity analysis of small changes to operational parameters;
- travel demand management (TDM) concepts aimed at reducing reliance on private car transport by providing alternative transport facilities and services operating in parallel along a designated urban corridor and possibly involving alternative allocations of available road space to the competing modes (e.g. Ramsay, 1994; Stephanedes and Kwon, 1988); and
- a need for efficiency in the operation of models and similar decision support tools, minimising overall data needs whilst ensuring that sufficient detail is available (Yagar, 1988). In general this need may be realised in hierarchical systems of models (e.g. Taylor, 1999), but the geographic form of the corridor may accentuate this distinction between the 'network' of immediate interest and that of the surrounding region (Roy *et al.*, 1996).

A corridor approach is a useful artefact for large scale studies where there are strong directional travel movements (along the corridor's principal axis). Detailed studies require the use of techniques that can split trip movements according to the behavioural characteristics of the travellers and operational and locational characteristics of the transport modes, services and routes within the corridor. This requires more detailed decision support tools than those normally applied in strategic network analysis. At the same time, most corridor studies involve incremental analysis and thus demand a maximum efficiency of use of any tool, especially in terms of computing time and effort. Effective analysis tools thus need to use the simplifying spatial characteristics of the corridor, whilst retaining the necessary detail for route and mode choice analysis.

Sustainability concepts applied to integrated land use and transport planning require that attention be given to the use of modes such as walking and cycling for trips in a local area, and that these modes be considered as potential 'feeder services' to public transport for longer distance travel (e.g. Cervero and Radisch, 1996). This study focuses on the ability to increase public transport usage in a particular corridor, the north-west corridor of Adelaide (a city of about one million people and the capital of the state of South Australia), through a combination of land use planning (for instance, by increasing residential densities and concentrating trip generators) and transport planning (by improving service capacity and operational performance along the corridor). The north-west corridor is approximately 22 kilometres in length from the city centre to the Outer Harbour. It is not untypical of many suburban Australian contexts at present – slow-growing, low residential densities by most international standards, and with low levels of public transport usage. In this context, a key question to be investigated is what changes to land use and transport policies are likely to have the most effect on reducing car travel and encouraging a shift to more sustainable modes of travel.

3. Literature review

The advantages of an activity-based modelling approach for analysing complex travel behaviour and forecasting travel demand have been stated by many researchers (e.g. Goulias, 1997; Kitamura, 1997; Pas, 1997). In general, there are two main approaches - namely the econometric model system and the microsimulation model system - in the application of activity-based models. Extensive literature on econometric modelling (e.g. Ben-Akiva *et al*, 1996; Bowman *et al,* 1998; Golob, 1998; Mannering *et al,* 1994; Hensher and Johnson, 1981; Louviere, Hensher and Swait, 2000) has appeared for some time, although the intense data needs for the models have restricted their widespread use. These models are normally based on the random utility theory to analyse and predict individual choice behaviour. Microsimulation models have also been postulated for some time, but are only now becoming feasible with the advent of readily available high capacity computers (Kitamura *et al*, 2000; Nagel and Axhausen, 2001).

Miller (1997) lists the reasons for and merits of adopting a microsimulation approach for travel demand forecasting. Most of these merits have proven themselves in models that have been developed or are currently under development. For example, the Sequenced Activity Mobility Simulator (SAMS), developed by the RDC, Inc team, was envisaged as the next generation of the urban transportation demand forecasting mode system (Kitamura *et al.*, 1996). The core of the system, the Activity Mobility Simulator (AMOS), is designed to predict traveller behaviour through a microsimulation of transportation activities and decisions. It generates a daily travel itinerary for each individual, including departure times, travel modes, destinations, chained trips, routes and travel partners, as well as some socio-demographic characteristic of the individual, and provides a response to a TDM measure. The calibrated AMOS is also the first operational tool and was implemented to study a sub-sample of commuters in the Washington, DC metropolitan area in order to predict the impacts of selected TDM measures.

Miller and Salvini (2001) discussed key issues that have been encountered and dealt with in the development of the Integrated Land Use, Transportation, Environment (ILUTE) modelling system. This modelling system is being developed by researchers at several universities in Canada as a state-of-the-art modelling system, designed to be a fully integrated microsimulation modelling system of urban transportation-land use interactions and including the environmental impacts of these interactions. Miller and Salvini discuss the key issues, including the representative population, the representative economic activities, the treatment of space and time etc., and, in particular, the issue of how the individual sub-models fit together within the overall modelling system. These are the most essential aspects that apply in the development of the majority of the activity-based travel demand microsimulation models, and such discussions and ideas not only indirectly indicate the advantages of such a model system but, at the same time, highlight the difficulties encountered in the development process. Such information is necessary, especially as the microsimulation of travel demand is still a relatively new field of modelling, and enables researchers in the same area to gain enlightenment from some of these very helpful ideas.

Another microsimulation model of travel behaviour – MOVER - was developed by Dornemann (2001). MOVER is a simulation tool with a core micro-geographical system which permits a very detailed analysis of mobility and travel behaviour. It focuses on each individual's life cycle, based on a list of attributes that reflect persons' life cycles, such as birth rate, education system, status of employment, income, household structure, etc. The system has the potential for individual status or the population structure in any future point of time to be simulated. Consequently, the effects of future modification of socio-economic structures and their results on travel behaviour and travel demand can be forecasted. Furthermore, the system includes the possibility to analyse the interrelation between mobility, travel behaviour, mode choice. The paper concludes with an outline of the proposed further development of the model, which includes some new ideas, such as the system having the ability to update data for commercial vehicle simulation and also to model the impact of e-commerce behaviour.

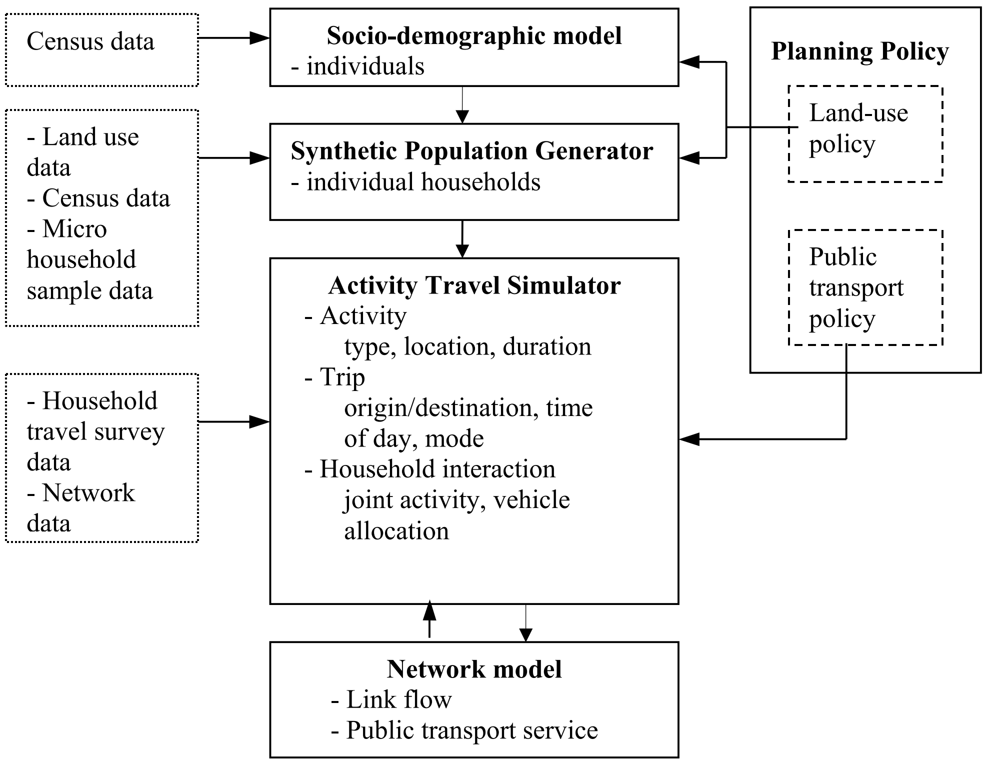
With the recognition of the merits of using microsimulation to model individual and household travel behaviour, together with the ability to develop practical tools for large scale urban transport modelling, a new area of microsimulation of travel survey data has emerged in favour of reducing the large cost of household travel surveys which are the main data source of travel demand forecasts. Stopher *et al* (2002) introduced a method of simulating household travel survey data with a case study in Adelaide, South Australia – see also Stopher and Pointer (2003). The procedure involved constructing distributions of relevant variables that could be used in travel demand models, such as numbers of trips by purpose, mode of travel, time of day of travel and trip length, along with a sample of local residents, which was drawn from disaggregate census data together with detailed information of their socio-economic characteristics. Travel data were then simulated from a national travel survey based on these socio-economic characteristics and using Monte Carlo simulation. In their study, the 1996 Census Household Sample File data for the Adelaide region was used to create a sample of households, and the US Nationwide Travel Survey data was used to simulate the travel data for Adelaide, which was then compared with the 1999 Metropolitan Adelaide Household Travel Survey (99MAHTS) data. The results suggest that this method is an acceptable method for simulating household travel survey data in particular as the results could also be improved by conducting a small local survey and using the results to update the distribution data.

4. Model description

Proper consideration of both short and longer term traveller behaviour requires integrated models that reflect a number of choices by individuals and households, including household location, vehicle ownership, destination, mode and route choice. An integrated land use and activity-based microsimulaton travel demand forecasting model system is a desirable model system to describe these variables. Figure 1 shows the general structure of the model system that is being developed as part of our research study. Four sub-models are included in the model system: Socio-demographic model, Synthetic Population model, Activity-travel simulator and Network model. In our earlier work (Taylor *et al.*, 2002) a separate land use sub-model was included in the model system. This current version of the model uses the land use data as an exogenous input only. The future incorporation of land use data is available in the form of alternative scenarios for the study area, provided by the state urban planning agency Planning SA. All these sub-models are loosely connected with each other and can be used independently. The functions of the sub-models are as follows:

- the socio-demographic model forecasts the evolution of individuals and households over time from an assumed base year;
- the synthetic population generator creates a synthetic population as an input representative population for the activity-travel simulator to predict travel demand in a specific study area;
- the activity-travel simulator, which is the key component of the model system, generates activity/travel patterns and forecasts travel demand, as well as scheduling activity and travel demonstrated by individuals and households in time and space; and
- the network model estimates travel movements (including public transport) in corridor networks.

Figure 1 General structure of the model system



The fundamental aspect of the model system is that it is an object-oriented paradigm, developed within the model system, which provides a user-friendly interface to a modeller. The Geographic Information System (GIS) is incorporated into the model system and is used as a platform to manipulate all spatial elements without zonal aggregation. It allows for the joint layering of a study area land uses, transportation, vacant land, household locations and household activity patterns. Taylor *et al* (2002) discuss the primary aim of the model development and the factors related to modal choice that require consideration as follows:

> '*the primary requirement of the model is to investigate factors relating to choice of travel mode in the study corridor, with emphasis on the usage of both public transport and non-motorised modes, and to suggest land use and transit service factors that could increase public transport usage (or reduce private motor vehicle usage). These factors could also include encouragement of households with a pro-*

> *pensity to use public transport to take up residence in the corridor. Thus the model is intended to be served as a comprehensive public transport policy analysis tool and also is designed to be sensitive to changes in demographics and land use development patterns in the corridor, and its inputs include alternative land use scenarios representing these alternative situations.*'

This model system may also be used as a 'laboratory' for modelling applications, testing and evaluation, with the north-west corridor being an area having similarities to corridors in other Australian cities which may require modelling of the relationships between land use, activities and transport.

To date, the model has been designed to forecast the participation of households in activities/travel in a typical weekday day in a study area, and it is planned to be developed to forecast travel demand in several continuing days. The basic process of the model system is to forecast travel demand within a study area using a base year population. The steps required are as follows:

- Step 1: to generate a synthetic population for a study area, the census data, census household sample data and current land use data are input into the synthetic population generator;
- Step 2: the synthetic population data, together with the household travel survey data and road network data are input into the activity travel simulator to forecast schedule activities and travel for each member of a household within the whole synthetic population and to obtain a forecasted travel demand; and
- Step 3: the travel demand data is then incorporated into the network model to estimate the link flow and public transport service requirements.

Any changes in planning policies could be tested by the model by obtaining information on changes in travel behaviour. A change in land use patterns can be included directly into the model, as described in step 1, while travel behaviour affected by any changes in public transport policy would be part of the process outlined in step 2. To measure any possible public transport policy change, the activity travel simulator would modify the basic activities and travel patterns to derive a new travel forecast set.

5. Data description and process

This section describes the data requirements and processes required for implementation of the model into the study area. The project data includes the ABS 1996 census data, a complete data set incorporating aggregated socio-demographic information that is required for generating the synthetic population, as well as providing a base year population to forecast future population based on death, birth and migration rates. Also included is the 1996 Census Household Sample File data, a disaggregated data set with the socio-demographic information of households, comprising a one percent sample of private dwellings, with their associated family and person records, and a one percent sample of persons from all non-private dwellings, together with the associated records. The use of this household sample data is necessary to formulate the sample data required to create the synthetic population for the study. In addition, the study uses the 1999 Metropolitan Adelaide Travel Survey (99MAHTS) data, which is the latest available travel survey data, containing detailed information on activities (in-home and out-of-home) and trips that all individuals in all households in the survey sample had taken on two consecutive days. It also includes socio-demographic information about the respondents, such as age, sex, occupation, employment status and personal income. Lastly, the project data includes the study area land use data, traffic analysis zone (TAZ) data and road network data. Most of these data have to be processed to establish databases for input into the models which are tailored for use in the study area.

5.1 The 1996 census data

The census data is used to obtain the observed characteristics of the population in the study area, and to generate the synthetic population based on the aggregated characteristics of the population. While the approach of acquiring the synthetic population used for this study can generally be applied to any household without taking account of the household types, travel behaviour and activity patterns may differ by different characterises of households, (for example, activities participated in by a lone person household would differ from those by households with children). Therefore, it is important to divide households into different classes to examine the travel behaviour and travel patterns in activity-based travel demand models. To do this, three types of households - family households, non-family households and group households - are segregated and then synthetic households generated respectively. Family households comprise the households with two or more related persons, and are further classified into adult couple family households and single parent family households. Households with persons living alone or non-related persons living together are categorised as non-family households. Persons living in dwellings such as college dormitories and hospitals are classed as group households. In this paper, only the process used to generate the adult couple family with children households is presented.

According to the census data, adult couple family with children households are categorised into seven classes, as follows:

Adult couple family with children 0-14 and independent children
Adult couple family with children 0-14 and dependent students 15-24
Adult couple family with children 0-14 and dependent students 15-24 and independent children
Adult couple family with children 0-14 only
Adult couple family with dependent student 15-24 and independent children
Adult couple family with dependent student 15-24 only
Adult couple family with independent children only

The socio-demographic variables of households, in terms of number of cars, number of workers, number of persons, and (possibly) age of the head of the household are used to construct synthetic populations within each family household type.

The tables that relate to these variables were extracted from the census data include:

Family type and relationship in household by age
Family type by weekly family income
Household type by family type by number of car
Household and family type by number of person
Dwelling structure by household and family type

The number of total households, total households by number of persons (3 persons household to 6+ persons household) by their family income, and total households by number of cars (0 car to 4+ cars, and car not stated) from each census collection district (CCD) were obtained. Summary statistics obtained from the census data for the study area are illustrated in Table 1 and 2. Table 3 shows an example of the total number of households by couple family with children 0-14 for part of the study area CCD[4]

Table 1 Total number of persons per household by adult couple family with children

	Number of persons				
	3	4	5	6+	Total
Total	6284	7245	2846	1056	17431

Table 2 Total number of cars per household by adult couple family with children

	Number of cars						
	0	1	2	3	4+	Not stated	Total
Total	406	4350	8529	2762	984	400	17431

[4] As there are 337 CCDs in the study area, the table shows a sample of CCDs, for illustrative purposes only.

Table 3 Example output: Total number of households by adult couple family with children 0-14 only by income

	Income								
CCD	Neg/Nil	$1-$299	$300-$499	$500-$699	$700-$999	$1000-$1499	>=$1500	NA[1]	Total
4100101	0	0	0	0	0	0	0	0	0
4100201	3	3	10	13	12	10	3	3	57
4100202	3	0	3	5	7	3	0	0	21
4100203	0	3	6	7	13	4	0	0	33
4100204	0	5	7	4	7	4	0	0	27
4100205	0	0	15	7	6	4	0	3	35
4100206	0	3	10	14	9	6	0	3	45
4100207	0	0	5	6	5	0	0	0	16

[1]Not applicable

5.2 The 1996 census household sample file data

The census household sample data is used to obtain detailed characteristics of households as basic sample households (with their members) to create a synthetic population for the study area. Considering that the sample size of the study area in the one percent sample file is small, a whole of Adelaide sample with 10,690 records was extracted from the file. The first step of the data process involves separating the data into the three different household types according to the census data. As described in Section 5.1, the categories are family households, non-family households and group households. In each family household type, the households were also processed further into adult couple family with children households[5], couple family only households and single parent family households. To correspond the same couple family household categories with the census data, the couple family households types were re-categorised into seven classes[6]. Each class was then again classified into eight small groups according to the family income, as follows:

Negative/Nil income
$1-$299
$300-$499
$500-$699
$700-$999
$1000-$1499
>=$1500 and
Not Applicable income

The variables included in each household are the number of cars, number of workers, number of persons, and age of the head of the household. Each sample household also has a family type and is associated with an ID number.

The demographic characteristics of the census household sample data are a little different from those observed in the study area data from the census, as the household sample data was extracted from the whole of the Adelaide metropolitan area. In order to reduce the possible bias and for the data to correspond with the observed characteristics, an adjustment was applied by putting weights on the household sample data according to the characteristics of the study area population as determined from the census data.

[5] In this paper, only the process of the adult couple family household data and used to create the synthetic adult couple family with children households are presented.

[6] Refer to the seven adult couple family with children household classes listed in Section 5.1 above.

5.3 The 1999 Metropolitan Adelaide Travel Survey (99MAHTS)

The 99MAHTS data comprises 5886 households, with 14,004 person records containing all person travel activities and information, and the associated personal information. After an initial data check, households with complete person records missing (one or more than one family member) were eliminated. Using inference methods, some missing or incorrect data in the person records was successfully repaired. A total number of 5415 households remained in the data set after further missing data repairing procedures were carried out using imputation methods to obtain approximately 20 percent of the income data which was missing[7] In the 99MAHTS data set, 858 of the households are part of the study area. As the activity travel simulator will initially only be used for simulating the activity/travel of households occurring in a typical weekday day, the activity/travel records in weekends and public holidays were removed into another file. According to the household types categorised in the census data and census household sample file data, the households were grouped into small classes, and representative activities/travel patterns were then generated by the number of persons and number of cars in each small group.

5.4 The traffic analysis zone data, census collection district zone data and the digital cadastral data base

Zones are necessary and important elements in a microsimulation model's spatial data management system (Miller and Salvini, 2000). One main reason is that input data, for example, individual households, firms, activity places, etc are all easily accessible in the zone level. The second reason is that the output results can be more easily explained in the zone level and displayed straightforward within a GIS system. In this study, two basic zones, the traffic analysis zone (TAZ) and census collection district (CCD), are required for the analysis of travel behaviour. The TAZ zones are used to identify trip origins and destinations and the CCD zones are used to recognise the zones where households reside. The CCD information, primarily designed by the Australian Bureau of Statistics, usually comprises approximately 300-400 households, and also includes property boundaries and natural features. The TAZ boundaries are not a standard unit in the Australian Standard Geographical Classification (ASGC), but are formulated specifically for transportation survey purposes. One CCD is often crossed by several TAZs, and also one TAZ may reside in two or three CCDs. Lastly, the digital cadastral data base (DCBD) comprises legal land parcels, their identification, geographic position of their boundaries relative to the national map grid and the dwelling count in that parcel. Each household whose travel behaviour is analysed in the model can be identified with their dwelling location in the dwellings table that is produced based on the DCBD. However, as each dwelling in the dwellings table only has the information of the CCD, but not the TAZ, therefor the TAZ data. The CCD data, along with the dwelling table, were processed in GIS ArcView for each dwelling to obtain its associated TAZ.

6. The synthetic population generator

The individual travellers and households are the basic entities for simulating travel behaviour in microsimulation travel demand models. Whilst it would be desirable to work with an entire population of individuals (or households) within the microsimulation model, a complete state estimation is basically impossible, because some data is systematically missing due to privacy issues. Therefore it is necessary to use alternative techniques, such as the development of a synthetic population, and it is even advantageous to have a synthetic population generator within a microsimulation travel demand model system. Most of the microsimulation models developed to date, whether for short-term predictions or for medium and long-term forecasting applications, have required a representative sample of population. However, an updating procedure of the representative sample must be included in a medium or long-term application (Miller, 1996). One of the approaches to deal with this is to use a survey sample as representative of a whole population, such as used in the development of the Activity-Mobility Simulator (AMOS) (RDC Inc., 1995). An alternative approach is to use census data together with a micro-data sample to generate a synthetic population as representative of a study area (for example, Beckman *et al*, 1996).

The approach adopted for this study is the synthetic household method. The basic concept of the approach was discussed by Taylor *et al* (2002), and it involves using a GIS model of the corridor, which includes dwelling

[7] The methods and processes used to repair missing data are detailed in another paper – see Xu *et al* (2002).

types and locations. The synthetic population (households and individual household members) is generated at the census collector district (CCD) level using the Australian Bureau of Statistics' (ABS) census data and one per cent household sample file data. The detailed land use data is then incorporated into the model, including the cadastral database for the study region, which identifies the location and attributes of all dwellings and other building and site developments in the region, by assigning the synthetic population to the available dwellings in the CCD to match the known demographic characteristics of the population. Monte Carlo simulation is employed to generate the synthetic households. The GIS software MapObjects was incorporated into the model for data presentation and display.

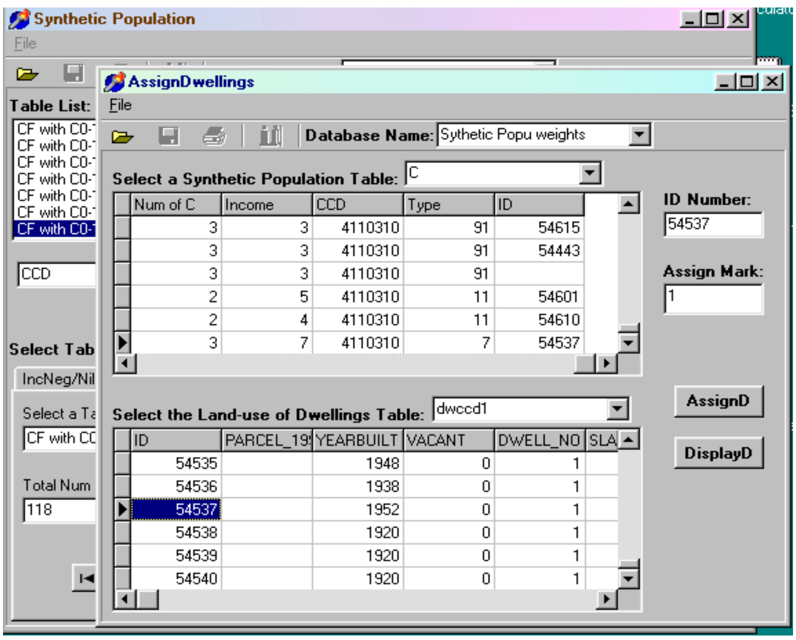
6.1 The synthetic population procedure

There are two procedures within this method. The first is the creation of the synthetic population (households and individual household members) at the CCD level. After acquiring a synthetic population, the second procedure is carried out, involving assigning the synthetic households into available dwellings at each CCD to match the known demographic characteristics of the study area population.

The synthetic population generator first reads the number of households in each CCD by their family type and family income, and then randomly picks a household from an appropriate category (family type by family income) in the household sample data file, and assigns it to a household that has the same category in the CCD.

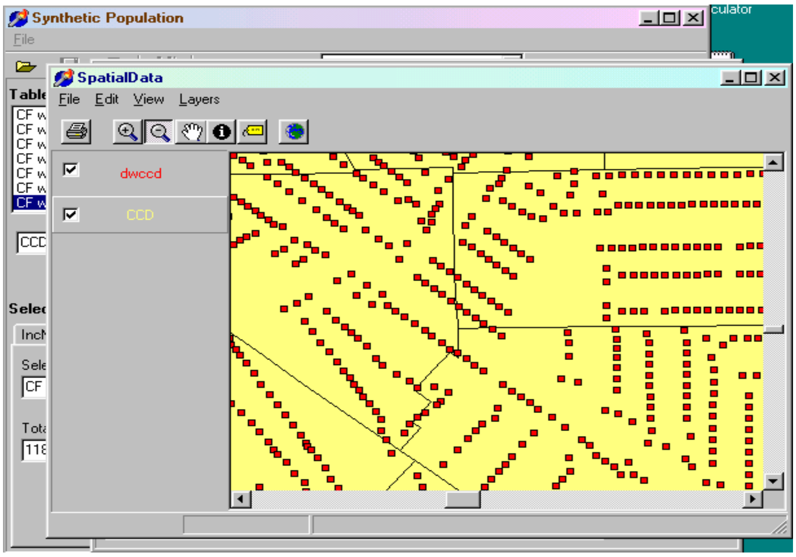
After all the households in each CCD have received their detailed household with their household members, then the second procedure is carried out by the program. As the land use data was incorporated into the model, according to the available dwellings and their locations and attributes, each synthetic household is selected and assigned to a unique dwelling. This process is repeated until all households receive their dwellings. Figure 2 shows the model interface of synthetic population generator and dwelling assignment.

Figure 2 The interface of the synthetic population generator and dwelling assignment



When all the synthetic households have been assigned to their suitable dwellings, the households with their dwelling locations can be displayed within the model or in GIS ArcView. Figure 4 shows the interface of dwelling display within the model.

[8]Figure 3 The interface of the dwelling display



6.2 Validation

For effective use of the model, several different synthetic populations are sampled, to allow validation of the model with observed behaviour. Each synthetic population may be regarded as an independent sample of the real population.

In this paper, only the results from the synthetic households of adult couple families with children are presented. Ten different synthetic populations were sampled. The Chi-square goodness of fit test was used to test each synthetic population with the observed characteristics of the population, in terms of the number of cars and the number of persons by the family type and family income. All the results showed that the synthetic households have similar characteristics with the observed households. Tables 4 and 5 show the results of the goodness of fit test in terms of the number of persons and the number of cars between the observed and synthetic population.

Table 4 Results from goodness of fit test by number of persons in the population

Number of persons	3	4	5	6+	Total	Chi-square χ^2
Observed	6284	7245	2846	1056	17431	
SynPop1	6298	7244	2875	1014	17431	2.00
SynPop2	6380	7182	2847	1022	17431	3.11
SynPop3	6346	7204	2809	1072	17431	1.57
SynPop4	6383	7194	2795	1059	17431	2.84
SynPop5	6361	7277	2775	1018	17431	4.22
SynPop6	6431	7125	2845	1030	17431	6.07
SynPop7	6369	7223	2840	999	17431	4.31
SynPop8	6281	7265	2836	1049	17431	0.14
SynPop9	6371	7165	2857	1038	17431	2.44
SynPop10	6316	7148	2913	1054	17431	3.04

Five per cent critical value of χ^2 with three degrees of freedom is 7.81

[8] In this display the black lines represent the CCD boundaries and the red squares indicate dwelling locations

Table 5 Results of goodness of fit test by number of cars in the population

Number of Cars	0	1	2	3	4+	Not state	Total	Chi-square χ^2
Observed	406	4350	8529	2762	984	400	17431	
SynPop1	391	4370	8496	2760	1012	402	17431	1.58
SynPop2	405	4266	8619	2729	1015	397	17431	3.97
SynPop3	381	4377	8522	2718	1041	392	17431	5.88
SynPop4	375	4205	8659	2782	1002	408	17431	9.82
SynPop5	380	4304	8618	2707	1025	397	17431	5.91
SynPop6	436	4253	8592	2708	1047	395	17431	10.00
SynPop7	363	4333	8523	2790	1033	389	17431	7.65
SynPop8	409	4289	8529	2743	1063	398	17431	7.36
SynPop9	385	4363	8631	2670	990	392	17431	5.61
SynPop10	370	4341	8571	2748	1019	382	17431	5.54

Five per cent critical value of χ^2 with five degrees of freedom is 11.07

7. The activity travel simulator

The activity travel simulator is an activity-based microsimulaton of daily household activity and travel over time and space. It is envisaged as an analytical tool for policy makers to assist their planning policy decision-making. The travel behaviour could be best understood by such a model through exploring the decisions made on engagement in optional activities, the allocation of flexible times to activities, the choice of location for activities to be participated in and the interaction between household members in undertaking their activities and travel, as well as the decision to trade off in-home and out-of-home activities. The model procedure is based on the theory of decomposition of a daily activity/travel pattern into segments to which certain aspects of observed activity/travel behaviour are matched. Key features of the simulator are as follows:

- It schedules each member of the household's daily activities and travel within a study area, with a consideration of the possible interactions between the members of a household in the choice of their activity, duration and mode, trip chains are also modelled at the same time.
- It supports decision-making with issues of public transport provision and level of service. For example, it can simulate changes of travel behaviour when a policy scenario is applied.
- The output of the model is flexible with respect to the level of aggregation or disaggregation. The output can be a selected socio-demographic group, a selected CCD or a whole study area. The output result includes the type of activity, trip frequencies by purpose by mode, trips frequencies by purpose by time of day, household trip rates, etc.

As the household is used as a primary entity in the analysis of travel behaviour within the context of this study, several elements that could influence household members participation in activities and travel have been taken into account in modelling the activities/travel patterns. These elements include the resources that may be shared by household members (for example, income, vehicle), constraints and conditions that may be imposed on other household members in participating in activities and travel, as well as modal choice. The constraints include coupling constraints, spatial-temporal constraints, modal constraints and activity constraints. For example, in general, a full time employee would only participate in a social recreation activity after an eight-hour work activity or during the lunch time period of work. While if a person who is a parent has a trip chain of dropping off a child to school on the way to work, then the child has a travel mode for going to school as a car passenger.

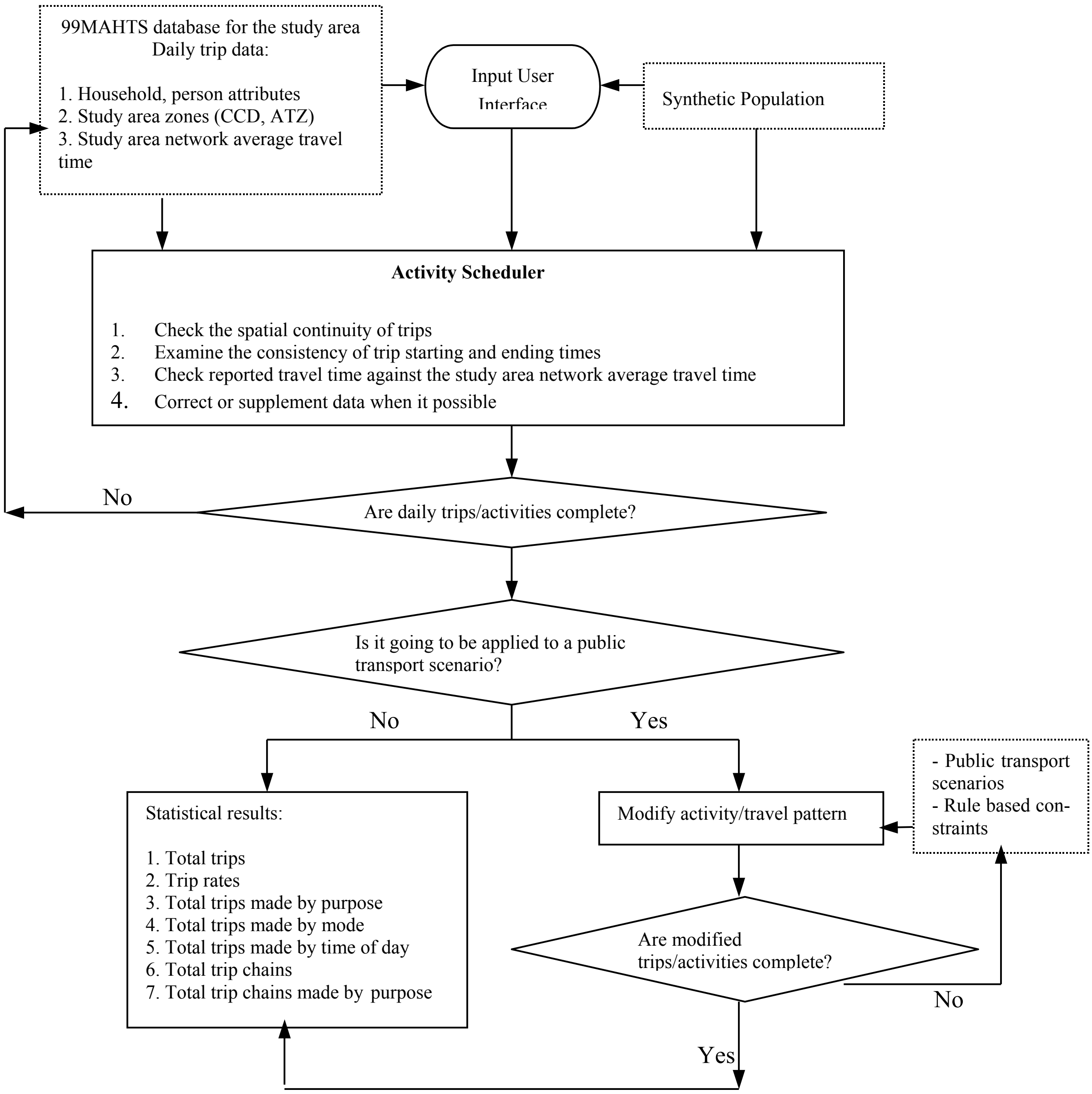
Monte Carlo simulation was employed in the procedure. The flow chart in Figure 4 shows the modelling procedure. The steps of the modelling process are as follows:

- Step 1: The simulator reads the socio-demographic information of each household from the input data of a synthetic population to recognise the family type.
- Step 2: According to the family type, the simulator looks for a suitable representative activity/travel pattern in the processed 99MAHTS database, and schedules a 24 hour activity/travel pattern to a member of the

household. From the start time point, the simulator simulates an activity type, its duration and location based on the observed activity distribution aligned with the activity/travel pattern and time step. A new activity and its characteristics are simulated at the finish time of the preceding activity based on the activity/travel pattern. This process continues until the complete 24-hour activity/travel pattern is determined for each individual. This process usually starts with one of the parents, if it is an adult couple family household or a single parent family household, until all the members of a household receive a daily activity/travel pattern, including in-home activities and trip chains.

- Step 3: The simulator then checks each person's activity/travel pattern, including the spatial continuity of trips, the consistency of trip starting and ending times, the travel time against the network travel time, and the availability of vehicles in the household against the mode being used.
- Step 4: It then corrects or supplements data if required.
- Step 5: If a public transport policy scenario is not being considered, it calculates the statistical output.
- Step 6: If a public transport policy scenario is being considered, the simulator reschedules activity/travel patterns for each person within the whole synthetic population, and Steps 3 to 5 are then repeated.

Figure 4 The flow chart of the modelling process



The results displayed in Table 6 are one of the outputs, aggregated statistics from a selected socio-demographic group, adult couple family with children 0-14 years old.

The mode choice decision results from the simulator are based on rules and observed data. In addition, a nested logit modal choice model has been incorporated into the model system and could be used for mode choice as well. In support of the primary aim of the study, the nested logit modal choice model has been designed with three hierarchies. The first level is the choice between motorised and non-motorised transport modes; the second level is the choice between public transport and private motor vehicle transport or between walking and cycling; and the third level is the choice between bus and train and other public transport modes or between car driver and car passenger. The model is still under development, and thus there are no results available at present. We expect to present some initial results at the IATBR conference in Lucerne in August 2003.

Table 6 Aggregated statistic results from adult couple family with children 0-14:

Total number of households:	8779	Total trip chain	16575
Total number of person:	34873	1 stop only:	12832
Total trips:	123847	2 stops:	2997
Households not making any trips:	104	3 stops:	559
Persons/per hhld:	3.9723	4 stops:	113
Trips/per hhld/per day:	14.107	5 stops:	0
Trips/per person/per day:	3.55	6 stops +:	74
Trips made by purpose:		Trip chains made by purpose:	
Home-based work:	17617	Change mode of travel/vehicle:	4823
Home-based education:	18069	Dropped-off/picked-up someone:	7422
Home-based shopping:	12389	Accompanied someone:	9228
Home-based social/recreation:	21256		
Home-based personal business:	36169		
Home-based other:	2095		
Non home-based employers business:	1570		
Non home-based personal business:	14682		
Trips made by mode:		Trips made by time of day:	
Train:	516	23:01 - 05:00:	627
School Bus:	0	05:01 - 07:00:	1881
Metro Ticket Bus:	897	07:01 - 08:00:	5416
Other Bus:	0	08:01 - 09:00:	19151
Car Driver:	60508	09:01 - 11:00:	13409
Car Passenger:	47492	11:01 - 13:00:	17889
Truck:	536	13:01 - 15:00:	14956
Taxi Passenger:	242	15:01 - 17:00:	27485
Motor Cycle:	78	17:01 - 18:00:	10370
Bicycle:	824	18:01 - 19:00:	4466
Walk, Wheelchair:	12752	19:01 - 21:00:	5524
Other:	2	21:01 - 23:00:	2673

8. Conclusion and future development

In this paper, we have proposed a framework for modelling household activity and travel behaviour based on the notion that travel is derived from participation in activities. The data used, along with the processing procedures are discussed and demonstrated. Two main components of the model, the synthetic population generator and activity travel simulator, are presented in detail.

The synthetic population generator creates synthetic populations that can be used in any activity-based microsimulation travel behaviour model. Two procedures are included in this approach. Firstly, a synthetic population (households with their household members) is generated at the census collector district level, using census data and the household sample file data, and by applying Monte Carlo simulation. Secondly, the synthetic households are assigned to available dwellings by incorporating detailed land use data into the model to match the known socio-demographic characteristics of the study area population. In this paper, only the generation of synthetic adult couple family with children households in a study area is presented. However, the same procedures can be used to generate synthetic populations of other household types. The advantage of this method is that a GIS technique is used in the model. This provides a clear picture of the location of the synthetic households. Travel behaviour and travel patterns of people living in any location can be analysed through the microsimulation model if the synthetic populations are input as representative of the observed population. The results from the initial validation indicate that this procedure has performed very well as synthetic populations generated by this method have similar characteristics to the observed population.

The activity travel simulator schedules a daily activity and travel pattern for each member of the synthetic population based on socio-demographic characteristics, observed activities and travel characteristics, as well as a set of rules and constraints using Monte Carlo simulation. The output from the model can be at either an aggregated or disaggregated level. The results of the analysis include total trips, trips rates, total trips made by purpose, total trips made by mode, total trips made by time of day, total trip chains and total trip chains made by purpose. In this paper, only the results from the model by a socio-demographic group adult couple family with children 0-14 years are demonstrated. The model has been designed to be sensitive in favour of assisting policy makers to their planning policy decisions and, therefore, activity travel patterns can be modified while a travel demand management measure scenario is applied. [This function is currently under construction.]

Since the model system is an ongoing research project, future development of the sub-models will be carried out, as well as some experiments to establish. For the practical use of the model, the validation of the activity travel simulator and further validation of the synthetic population generator are also considered important. Current research activities and future research activities contributing to the development of the model system include the following:

- Further validation by the use of more variables, such as the age of head of households, will be performed. The synthetic households were generated at the CCD level, while the validation was carried out for the whole study area by aggregating households from all CCDs, it is possible that some deviations exist in some CCDs. It would be desirable if the validation could be carried out at the CCD level. A comparison of the socio-demographic characteristics between the synthetic households and observed households at some CCDs will be considered in future research;

- Further synthetic populations, for alternative planning scenarios within the corridor, will also be generated in order to test their likely travel outcomes;

- Development of the full function of the activity travel simulator, including the nested logit modal choice model, and applying public transport policy scenarios to estimate and analyse the possible changes in travel behaviour; and

- Development of the proposed network model.

9. Acknowledgement

The research project described in this paper is supported by a ‘Linkage-Industry’ collaborative research grant provided by the Australian Research Council (ARC), with the industry partner being the South Australian Department of Transport and Urban Planning, which includes the agencies Transport Planning Agency, Planning SA and the Passenger Transport Board. The authors wish to thank the staff of these agencies for their kind cooperation and interest in the project.

10. References

Beckman, R.J., Baggerly, K.A. and McKay, M.D (1996). Creating synthetic baseline populations. *Transportation Research Part A,* 30, pp. 415-429.

Ben-Akiva, M., Bowman, J.L. and Gopinath, D. (1996). Travel demand model system for the information ear, *Transportation,* **23** (3), pp.241-266.

Bowman, J.L., Bradley, M., Shiftan, Y., Lawton, T.K. and Ben-Akiva, M. (1998) Demonstration of an activity-based model system for Portland, *8th World Conference on Transport Research*, Antwerp, Belgium pp.15p.

Cervero, R. and Radisch, C. (1996). Travel choice in pedestrian versus automobile oriented neighbourhoods. *Transport Policy* **3** (3), pp.127-141.

Dornemann, M. (2001). Microsimulation of travel behaviour in a congested urban area - MOVER. In Hensher, D A (eds). *Travel Behaviour Research: The leading Edge*. (Pergamon: Oxford), pp.677-688.

Golob, T.F. (1998). A model of household choice of activity participation and mobility. In *Theoretical Foundations of Travel Choice Modeling* (Eds, Gärling, T., Laitila, T. and Westin, K.) Elsevier, pp.365-397.

Goulias, K.G. (1997). Activity-based travel forecasting: what are some issues. *Activity-Based Travel Forecasting Conference Proceedings*, Travel Model Improvement Program (TMIP), U.S. Department of Transportation, Washington, D.C.

Hensher, D.A. and Johnson, L.W. (1981). *Applied Discrete Choice Modelling*. London, Croom Helm.

Kitamura, R., Pas, E.I., Lula, C.V., Lawton, T.K. and Benson, P.E. (1996) The Sequenced Activity Mobility Simulator (SAMS): an integrated approach to modelling transportation, land use and air quality, *Transportation,* 23, pp.267-291.

Kitamura, R. (1997). Activity-based travel forecasting: what are some issues Applications of models of activity behavior for activity based demand forecasting. *Activity-Based Travel Forecasting Conference Proceedings*, Travel Model Improvement Program (TMIP), U.S. Department of Transportation, Washington, D.C.

Kitamura, R., Chen, C., Pendyala, R.M. and Narayanan, R. (2000). Micro-simulation of daily activity-travel patterns for travel demand forecasting, *Transportation,* 27, pp.25-51.

Louviere, J.J., Hensher, D.A. and Swait, J.D. (2000). *Stated Choice Methods: Analysis and Application*. Cambridge , Cambridge University Press.

Mannering, F., Murakami, E. and Kim, S.G. (1994). Temporal stability of travellers' activity choice and home-stay duration: some empirical evidence. *Transportation* 21 (4), pp.371-392

Miller, E.J. (1997). Microsimulation and activity-based forecasting, *Activity-Based Travel Forecasting Conference Proceedings*. Travel Model Improvement Program (TMIP), U.S. Department of Transportation, Washington, D.C.

Miller, E.J. and Salvini, P.A. (2001). The integrated land use, transportation, environment (ILUTE) microsimulation modelling systems: description and current status. In Hensher, D.A. (ed). *Travel Behaviour Research – The Leading Edge*. Oxford , Pergamon-Elsevier, pp711-724.

Nagel, K. and Axhausen, K. (2001). Microsimulation. In Hensher, D.A. (ed). *Travel Behaviour Research – The Leading Edge*. Oxford, Pergamon-Elsevier, pp239-258.

Pas, E.I. (1997). Recent advances in activity-based travel demand modelling. *Activity-Based Travel Forecasting Conference Proceedings*, Travel Model Improvement Program (TMIP), U.S. Department of Transportation, Washington, D.C.

Ramsay, E. (1994). Corridor level modelling of TDM measures. *ARR No 252, Proc workshop on models for travel demand management*, ARRB Transport Research Ltd, Melbourne, pp.75-78.

RDC Inc (1995). Activity-based modeling system for travel demand forecasting. RDC, Inc, San Framsisco.

Roy, J.R, Marquez, L.O, Taylor, M.A.P. and Ueda, T. (1996). SUSTAIN-a model investigating sustainable urban structure and interaction networks, In Hayashi, Y and Roy, J R (eds) *Transport, Land-Use and the Environment*. (Kluwer: Dordrecht), pp.125-145.

Stephanedes, Y.J. and Kwon, E. (1988). Optimisation strategies for transit systems in urban corridors. *Transportation Research record 1165,* pp.75-85.

Stopher, P., Bullock, P. and Rose, J. (2002). Simulating household travel survey data in Australia: Adelaide case study. *Papers of the 25th Australasian Transport Research Forum*, Canberra, October 2002. (Bureau of Transport and Resource Economics: Canberra).

Stopher, P and Pointer, G (2003). Monte Carlo simulation of household travel survey data with Bayesian updating. *Proc 21st ARRB Conference*, Cairns, May. (ARRB Transport Research Ltd: Melbourne), paper no 47.

Taylor, M.A.P. (1999).Dense network traffic models, travel time reliability and traffic management. I: general introduction. *J Advanced Transportation 33 (2),* pp.218-233.

Taylor, M.A.P., Hamnett, S., Xu, M. and Page, A. (2002). A microsimulatioon approach to modelling travel choice: a case study from Adelaide's north-west corridor. *Proceedings, An International Conference on Seamless and Sustainable Transport*. Singapore, November 2002.

Xu, M. and Taylor, M.A.P. and Page, A. (2002). A comparison of two methods for imputing missing income from household travel survey data. *Papers of the 25th Australasian Transport Research Forum*, Canberra, October 2002. (Bureau of Transport and Resource Economics: Canberra).

Yagar, S. (1988). Generating partial origin-destination tables for streamlined application of corridor models. *Transportation Research Record 1194,* pp.135-138.