PERGAMON

Computers, Environment and Urban Systems
28 (2004) 9–44

Computers, Environment and Urban Systems

www.elsevier.com/locate/compenvurbsys

Microsimulating urban systems

Eric J. Miller[a,*], John Douglas Hunt[b], John E. Abraham[b], Paul A. Salvini[c]

[a]Joint Program in Transportation, University of Toronto, 35 St. George Street, Toronto, Ontario M5S 1A4, Canada
[b]Department of Civil Engineering, The University of Calgary, 2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada
[c]Department of Civil Engineering, University of Toronto, 35 St. George Street, Toronto, Ontario M5S 1A4, Canada

Abstract

This paper presents a status report concerning on-going research and development work by a team of Canadian researchers to develop a microsimulation, agent-based, integrated model of urban land use and transportation. It describes in some detail the overall design and current status of the ILUTE (Integrated Land Use, Transportation, Environment) modelling system under development. The overall purpose of ILUTE is to simulate the evolution of an entire urban region over an extended period of time. Such a model is intended to replace conventional, aggregate, static models for the analysis of a broad range of transportation, housing and other urban policies. Agents being simulated in the model include individuals, households and establishments. The model operates on a "100% sample" (i.e., the entire population) of agents which, in the base case, are synthesized from more aggregate data such as census tables and which are then evolved over time by the model. A range of modelling methods are employed within the modelling system to represent individual agents' behaviours, including simple state transition models, random utility choice models, rule-based "computational process" models, and hybrids of these approaches. A major emphasis within ILUTE is the development of microsimulation models of market demand-supply interactions, particularly within the residential and commercial real estate markets. In addition, travel demand is modelled explicitly as the outcome of a combination of household and individual decisions concerning the participation in out-of-home activities over the course of a day. Spatial entities in the model include buildings, residential dwelling units and commercial floor-space, as well as aggregate "spatial containers" such as traffic zones, census tracts or grid cells.

Keywords: Microsimulation; Agent-based; Urban systems; Location processes; Land development; Activity; Travel behaviour; Automobile transactions; Land use; Transportation

* Corresponding author. Tel.: +1-416-978-4076; fax: +1-416-978-5054.
E-mail addresses: miller@civ.utoronto.ca (E.J. Miller), jdhunt@acs.ucalgary.ca (J. D. Hunt), jabraham@acs.ucalgary.ca (J.E. Abraham), pas@sidefx.com (P.A. Salvini).

0198-9715/03/$ - see front matter
doi:10.1016/S0198-9715(02)00044-3

1. Introduction

Microsimulation, in which the dynamic behaviour of individual agents is explicitly simulated over both time and space to generate aggregate system behaviour, has been applied with increasing frequency over the past decade or more in the field of transportation systems analysis. Perhaps the best developed application is in the area of transportation network simulation models, in which a number of operational (and often commercially supplied) software packages exist which model second-by-second operations of individual road and/or transit vehicles over very high fidelity representations of urban transportation networks. Examples (but a far from complete list) of such models include TRANSIMS (Barrett et al., 1995), PARAMICS (Quadstone, 1999), INTEGRATION (Van Aerde & Yager, 1988a, 1988b), DYNASMART (Hu & Mahmassani, 1995; Mahmassani, Hu, & Peeta, 1994), DynaMIT (Ben-Akiva, Bierlaire, Koutsopoulos, & Mishalani, 1998; Ben-Akiva, Bierlaire, Koutsopoulos, & Yang, 1999) and VISSIM (PTV, undated). Models of urban travel behaviour, in which the temporal, spatial and modal (e.g. car versus transit) distributions of trips in an urban area are predicted, similarly are increasingly being developed and applied in a microsimulation framework (Arentze & Timmermans, 2000; Axhausen, 1990; Bradley & Bowman, 1998; Bradley, Outwater, Jonnalagadda, & Ruiter, 2001; Goulias & Kitamura, 1992; Harvey & Deakin, 1996; Jonnalagadda, Freedman, Davidson, & Hunt, 2001; Kreibich, 1979; Miller, Roorda, Eberhard, & Litwin, 2002; RDC, 1995; Spear, 1994). Other transportation-related microsimulation applications include auto ownership (Bonsall, 1982), residential location choice (Mackett, 1985, 1990; Miller, Noehammer, & Ross, 1987; Oskamp, 1997; Otter, van der Veen, & de Vriend, 2001; Wegener, 1983), and demographic evolution (Cumpston & Sarjeant, 1998; Kitamura & Goulias, 1991). More extensive reviews of microsimulation applications in transportation are provided in Miller (1996) and Miller and Salvini (2001a), which also place transportation-related microsimulation modelling within the larger context of the microsimulation modelling literature as a whole.

A long tradition exists in the transportation planning field of trying to bring these various transportation-related components of urban systems (networks, travel behaviour, location choice, etc.) into an integrated modelling framework, in which the two-way interactions between urban form (land development, building supply, location choices of households and finns, etc.) and transportation (auto ownership, travel demand, transportation network performance, etc.) are modelled together in as comprehensive, logically consistent and theoretically sound fashion as is practically possible, given current theory, data, modelling methods and computational capabilities. Miller, Kriger, and Hunt (1998) provide an extensive discussion of both the technical and policy rationales for developing such integrated "transportation – land use" or urban system models (we will use the latter term throughout this paper, although the former is the more common term in the transportation literature), as well as a detailed review of the state of best practice in this field. Other reviews of the integrated modelling literature include Hunt, Kriger, and Miller (1999, in press), Wegener (1994, 1998) and Southworth (1995).

The ‘‘track record’’ of operational integrated urban models is a very mixed one, with many (particularly early) modelling efforts either failing to become operational or failing as useful policy analysis tools. Lee's seminal critique of ‘‘first generation’’ comprehensive modelling efforts, ‘‘Requiem for Large-Scale Models’’ (Lee, 1973) is still relevant in considering such models today: there is no doubt that integrated urban models are computationally intensive, data hungry, make extreme demands on our theoretical understanding of urban spatial processes and our methodological capabilities for capturing that understanding within operational computer code, and are difficult to estimate and validate. Nevertheless, nearly 30 years have passed since Lee's damning critique. Extraordinary progress has been made over this time period in computational capabilities, modelling methodology, empirical and theoretical understanding of spatial processes, and data resources to support modelling activities. Research and development with respect to integrated models has proceeded over this time period, to the point that operational models are in use in a number of locations worldwide, particularly outside of North America (Wegener, 1994, 1998).

The usefulness for credible, operational integrated models in urban policy analysis is, hopefully, virtually self-evident. As is discussed in greater detail in Section 2, a fundamental, two-way transportation—land use interaction exists. Ignoring the longer-run ‘‘feedbacks’’ within this interaction (as is routinely done in conventional urban transportation planning analysis) is likely to be a serious mis-specification of the urban system and may well lead to serious mis-estimates of the response of this system to policy initiatives. Conder and Lawton (2002) provide an instructive example of this, in which they compare travel forecasts for Portland, Oregon generated in the traditional manner with those generated using an integrated land use–transportation model. They conclude that the integrated approach generates improved (i.e. more credible, plausible, policy-sensitive) results relative to their previous results generated through traditional, non-integrated means.

Currently operational integrated models, including several commercially available software packages such as MEPLAN (Hunt & Simmonds, 1993) and TRANUS (Modelistica, 1995), are all ‘‘conventional’’ in the sense that they are spatially aggregate (typically using fairly large zones to represent the urban area), temporally static (usually with strong and explicit equilibrium assumptions for all but the land development market) and incorporate very little socio-economic disaggregation of the population being modelled (Hunt et al., 2001). More recently developed and emerging models, however, tend to be more spatially and socio-economically disaggregated, while still retaining the basic zone-based modelling structure [e.g. MUSSA (Martinez, 1997) and UrbanSim (Waddell, 1998)]. UrbanSim is also noteworthy in that it adopts a dis-equilibrium approach to modelling land use and location choice processes.

As in many modelling applications, a more disaggregated approach to modelling socio-economic processes such as travel behaviour, residential location, etc. is generally desirable in order to reduce model aggregation bias, enhance its behavioural fidelity, etc. (Goulias & Kitamura, 1992). Similarly, it is increasingly recognized that the dynamic evolution of urban systems must be explicitly captured if future system states are to be properly estimated. That is, it can be strongly argued that urban systems evolve in a path-dependent fashion (especially in the presence of significant policy interventions into the system) that may not be well captured by conventional static equilibrium models. Putting these two observations together leads inevitably to the adoption of a microsimulation approach to modelling such systems (Miller, 1996).

Agent-based models are ones in each individual actors within the system of interest are modelled as autonomous "agents", each one of which possesses identity, attributes and the capability to "behave", i.e. to make decisions and to act within the system. Agent-based modelling has been recognized as an extremely powerful design paradigm across virtually the gamut of socio-economic modelling (Weiss, 2001), including travel-related behaviour (Polak & Huang, 1999; Cuena & Ossowski, 2001; Timmermans et al., 2002), in that it provides an extremely efficient, effective and "natural" way of both conceptualising and implementing complex, dynamic, disaggregate models of human decision-making. Agent-based models have exploded as a practical, operational possibility over the last decade or more with the emergence of object-oriented software design principles (Booch, 1994; Taylor, 1990) and programming languages (Java, C++, etc.) as industry-standard approaches to designing and programming complex software systems.

Attempts to develop true agent-based microsimulation models of transportation – land use interactions have been relatively rare to date, either because such models were developed prior to the agent/object-based approach was widely understood and operationally practical [e.g. early integrated microsimulation models such as Mackett (1985, 1990), Miller et al. (1987), Wegener (1983), etc.], or they have evolved over time out of more aggregate, non-agent-based approaches (e.g. UrbanSim is evolving in terms of its level of disaggregation and its conversion to an object-based system, but the current version of the model probably should not be classed as being truly agent-based). One notable exception to this rule is the Dutch RAMBLAS model (Veldhuisen, Kapoen, & Timmermans, 2000) which is explicitly agent-based in its design and implementation.

The purpose of this paper is to describe on-going research and development work by a team of Canadian researchers to develop a fully agent-based, integrated microsimulation model of urban land use and transportation. The overall objective of this research program is to investigate in detail the suitability and feasibility of agent-based microsimulation for integrated urban systems modelling. It is hypothesized that agent-based microsimulation, in fact, represents the best approach currently available to modelling large, complex, dynamic, open-ended socio-economic systems such as an entire urban region. That is, it is believed that microsimulation may prove to be the most computationally efficient, practical approach to modelling highly complex systems. Partial support for this hypothesis is provided by the recent development of a new travel demand forecasting model for the New York metropolitan region. While not a fully integrated model and not agent-based, this modelling system is fully microsimulation-based. Results to date with this model are extremely promising, and the model developers are explicit in stating that they view microsimulation as the best practical way of trying to model one of the world's largest and most complex urban areas (Voshva, Peterson, & Donnelly, 2002). Nevertheless, "classic" issues in developing and applying integrated models (which include data requirements, computational feasibility, model parameter estimation, model validation, robustness of model results, etc.) clearly exist and must be addressed in detail before the "hypothesis" of the agent-based, microsimulation approach can be accepted.

Given this extended preamble, this paper describes in some detail the overall design and current status of the ILUTE (Integrated Land Use, Transportation, Environment) modelling system under development by the Canadian team. The overall purpose of ILUTE is to simulate the evolution of an entire urban region over an extended period of time e.g. (10–20 years into the future). Such a model is intended to replace conventional, aggregate, static models for the analysis of a broad range of transportation, housing and other urban policies.

2. Overview of the ILUTE modelling approach

It is important to understand the transportation planning context which motivates the development of a model such as ILUTE and within which the model is to be applied. The fundamental purpose of transportation is to provide *accessibility* between people and businesses which are located at diverse points within the urban region. In providing this accessibility, the transportation system inevitably bestows locational advantages upon certain sites for certain purposes (good access to shopping facilities, jobs, schools, customers, etc.) while other sites not as well served by the transportation system will be at a relative disadvantage. The transportation system consists of a network of physical infrastructure (roads, rail lines, control systems, etc.), vehicles (automobiles, buses, etc.), modes (drive oneself, transit passenger, walking, etc.) and services (private car, taxis, public transit, etc.). The performance of the network depends both on its physical and operational design (roadway physical capacity, operating rules and control systems, etc.) and the usage or load that is imposed on the system (e.g. roadway travel times increase as usage—and hence congestion—increases). System usage (i.e. the demand for travel), depends, in turn on the spatial distribution of people and activities (e.g. how many people live here and work there) and the quality of service provided to them to execute their daily schedule of activities (*ceteris paribus*, the faster and cheaper the transportation system is, the more travel that is likely to occur). Thus, as summarized in Fig. 1, there is a fundamental two-way interaction between the transportation system on the one hand and the "land use" or "urban activity" system on the other, where the urban activity system essentially refers to all other spatial–economic–social–environmental elements of urban regions. Further, these interactions occur over both the long term, as both urban form and the transportation system evolve over time, and in the short term as people and businesses plan and then go about executing their daily and weekly activities.

Urban systems models should be able to evaluate the impact of a wide variety of different policies on the transportation system, and the full impact of transportation

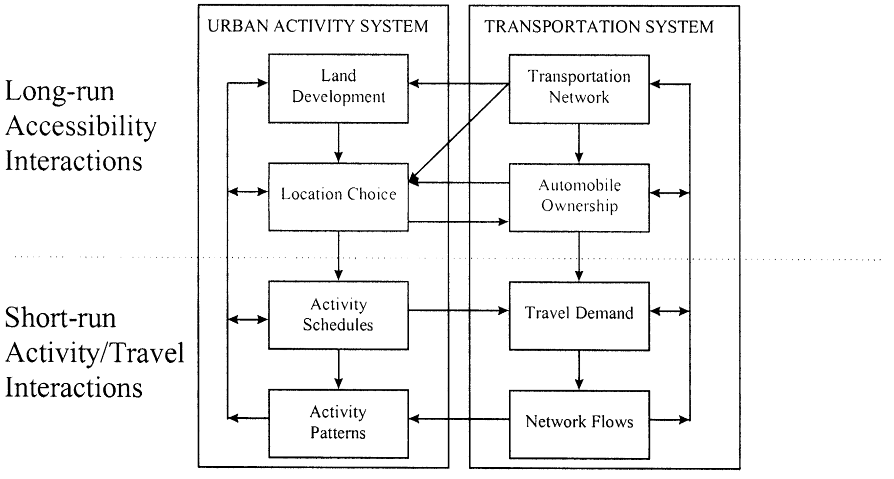

Fig. 1. Transportation–Urban–System–Interactions.

policies on the entire urban system. Thus, the full range of short- and long-term interactions shown in Fig. 1 should be considered. This, in turn, implies the need for comprehensive, integrated models such as ILUTE to explore "what if" scenarios concerning alternative policy options, which may well include land zoning regulations and property tax regimes, major infrastructure investments (build a new transit line or expressway), pricing policies (fuel taxes, transit fare policies, etc.), "managing" the supply/performance characteristics of the system (through the use of "intelligent transportation systems" or other operational controls), and "managing" the demand for transportation services (flexible working hours, rideshare programs, etc.), among others. In virtually all cases a detailed understanding of both the "transportation side" and the "land use side" impacts is important in projecting the net impact of a given policy. As one example, a "congestion pricing" scheme for an urban central area which is intended to discourage driving into the downtown in favour of transit usage may, in the long run, encourage businesses to relocate to more suburban locations which possess better (and, especially, cheaper) accessibility by auto.

Given these observations, an integrated urban systems model concerned with assessing transportation or other broader urban planning policy impacts should include:

1. evolution of the built environment (houses, commercial buildings, etc.);
2. evolution of population demographics over time, both in response to endogenous population changes (fertility, mortality, household formation and dissolution) and to migration into and out of the region;
3. simulation of the location choices of households and firms within this built environment;
4. simulation of the internal economy of the urban area within an explicit spatial framework (this includes the labour market, the internal exchange of goods and services, and the import/export of goods and services; that is, all

of the economic processes which generate person travel and goods movements within the urban area);
5. given the simulated urban structure and economic interchanges, simulation of the activity/travel patterns of the resident population and the internal flow of goods and services, by mode, route and time of day;
6. simulation of the performance of the road and transit systems serving these person and goods movements (travel times and costs, congestion levels, etc.); and
7. estimation of the atmospheric emissions (CO, CO_2, etc.) generated by the transportation system and by the 'fixed-point' production processes within the larger urban economy

Within this overall simulation process, modelling the location choices of households and firms are particularly critical, since if the spatial distributions of people and activities is fundamentally "wrong" in the model, then clearly the predicted travel patterns and transportation system performance characteristics will be equally wrong. Many factors influence household residential location choice, however, which should be included in the model, including: access to a wide variety of activities (family, employment, services, social activities, etc.), site specific influences not related to transportation (neighbourhood quality, types of neighbours, etc.), and housing characteristics (structure type, age, quality, price, etc.). Similarly, commercial location choices depend upon a range of factors such as: access to and from customers, owners, labour, suppliers; the location of competitors; agglomeration effects; non-transport-related site attributes; and building characteristics.

In focussing the discussion to this point on explicitly spatial decision processes such as travel and location choices, one should must not lose sight of item 4 above: simulation of the internal urban economy. Transportation and land use are but two markets within a much larger urban economy which must be understood and modelled if the urban area as a spatial entity is to be properly captured. Labour markets obviously determine work trip commuting patterns, while the distribution of "points of exchange" of retail goods and services (e.g. stores) determine "shopping" travel, to take two examples. At its core, an integrated model should be as much a model of urban economic (and demographic) processes as it is of travel and land development. The key difference between an integrated urban model in the sense being discussed here and other models of urban economics and demographics is the need to explicitly incorporate a spatial dimension into the analysis. This introduces an enormous new level of computational and theoretical complexity into the problem relative to aspatial models, and, indeed, represents the single biggest challenge in model development and implementation.

Obviously, not all aspects of such a complex modelling system can be presented in a single paper. In Section 3, we discuss key design issues concerning the treatment of space, time and the representation of the actors ('agents') in the system. Section 4 discusses agent decision processes modelled in ILUTE. We then deal with several of the key sub-models included in the overall ILUTE system in Section 5, while Section 6 briefly discusses some "generic" issues in the development of a large-scale model such as ILUTE. Additional discussion of ILUTE can be found in Miller and Salvini (1998, 2001b).

3. Agents, space and time: choosing a disaggregate representation

3.1. Decision-making agents

A "first principle" motivating a microsimulation modelling approach is that the behaviour of the actors within the system being modelled is significantly non-linear in nature, and, hence, significant bias can exist if one attempts to model such a system using arbitrary aggregations of these actors, such as spatial zones. Given this, a disaggregate, reductionist approach to system modelling is adopted in which system behaviour is the sum of the behaviour of the individual actors or agents comprising the system and through whose actions the system actually evolves (i.e., the urban area evolves over time through the decisions of developers, employers, home owners, workers, shoppers, etc.). However, having accepted this principle, in any practical model one must always define the level of disaggregation/aggregation to be employed. That is, there is always some level of disaggregation (with respect to space, time and agent attributes) beyond which it is infeasible, inefficient and/or unnecessary to go, given the model objectives and practical limitations of data, computing power, etc.

In ILUTE the primary agents being simulated in the model are individuals, households, businesses and establishments. Each agent, in general, plays a number of "roles"; that is, will have a variety of decisions for which it is responsible. A person, for example, can be a student, a worker, an employer, a child, a parent, a spouse, a home-owner, etc. A business similarly may be an employer, a producer of goods and services, a consumer of goods and services, a renter, a landlord, etc. A major challenge within the ILUTE framework is dealing with the complexity of roles (and the inter-agent interactions which they imply) in the development of an appropriate taxonomy and organizational structure to represent these agents. For example, both households and firms can be either property owners or renters, or, in some cases, both at the same time (with respect to two different properties).

Both individual persons and households are required as distinct types of agents in the model. Many decisions/processes of interest occur at the person level (birth, death, schooling, employment, most travel, etc.), but several critical decisions are fundamentally household-level in nature (e.g. household formation/dissolution, auto ownership, residential location). In addition, household level constraints play a major role in activity/travel decision-making, even if the actual travel is executed by individual household members (if you use the car, I can't; if I pick up our child from daycare, you don't have to), as well as quite possibly other "individual" decisions (a spouse's decision to quit work and go back to school generally has household-level ramifications concerning income, lifestyle, etc. and so presumably usually is not purely the spouse's decision alone to make).

Similarly, in ILUTE a business (or firm) is an economic enterprise which is organized into one or more establishments, where each establishment is located in a single physical location. In this case, decision-making may lie primarily at the business (firm) level, but the realization of these decisions is largely manifested at the individual site or establishment level [employees are hired/fired within a given establishment;

growth/contraction of the business will translate into changes in the size(s) and/or location(s) of the businesses establishment(s), etc.].

The model operates on a "100% sample" (i.e. the entire population) of agents which, in the base case, are synthesized from disaggregate sample data together with aggregate data such as census tables (Beckman, Baggerly, & McKay, 1996; Wilson & Pownall, 1976; Miller, 1996) and which are then evolved over time by the model. Synthesized individuals within the model are born, age, obtain schooling, select occupations and pursue required training, enter/leave the labour force, die, etc. Households form, evolve and eventually dissolve as their individual members move through their lifecycles. Businesses similarly follow an evolutionary pattern within the model, thus establishing a 'firmography' to go along with the population demography. Maintaining detailed representations of agent attributes is an essential component of the model, given the important role which these attributes play in determining agent behaviour (income, household structure, age, etc. all play major roles in determining household location choices, personal travel choices, etc.).

3.2. Spatial representation

A major issue of on-going research within the ILUTE project is the question of spatial representation. Many different spatial elements exist which are relevant to ILUTE processes, each of which possess strengths and weaknesses in terms of their appropriate use within the model. These include:

1. *Buildings* are the fundamental "containers" of activities and the primary occupiers of land. Buildings, however, can change size, be demolished, etc. and so do not provide a persistent, universal, invariant method for defining space.
2. Land is owned on a *parcel* by parcel basis, but parcels also are not immutable, and even in today's world of reasonably well developed GIS databases, working with parcel-level data can be an extraordinarily frustrating and difficult task due to incomplete/inconsistent information, etc.
3. *Zones* (census tracts, traffic zones) provide a stable, consistent spatial representation, and much of the data we use in constructing even disaggregated models inevitably comes in zone-based formats from sources such as the census. But, as has already been discussed, arbitrary spatial aggregations introduce aggregation bias into models which use them as their unit of analysis.
4. *Grid cells* provide a stable representation of space, are generally computationally more efficient to work with than zones, and, if defined on a sufficiently fine scale, will largely avoid aggregation bias problems. On the other hand, they are a very artificial construct and their use can (perhaps) lead to an overly abstracted representation of space and spatial processes. In the limit, as the grid size goes to zero, grid cells become individual points in space, and the spatial representation becomes fully continuous. Despite current raster GIS capabilities, fully continuous representations of spatial distributions and

processes is not a practical possibility at this time, at least for an ILUTE type model.

5. Firms and households do not only occupy land area, they occupy a certain amount of *floorspace* which is situated upon this area of land. Floorspace maps into land area in terms of floorspace to land densities or "coverage ratios". In many practical senses, particularly for businesses, it is floorspace which is the key measure of what it is they own or are renting, not how much land that floorspace happens to cover.
6. While households also occupy floorspace, they do so within a relatively well defined entity which is the *dwelling unit* (generally either a house or an apartment). That is, while a firm may say "we need 10,000 m^2 of office space for corporate headquarters", and this 10,000 m^2 of floorspace can be obtained in a variety of configurations, a household simply says "we need a house (apartment) to live in". Dwelling units, of course, come in different sizes (and, hence, will be more or less attractive to a household, given its space requirements, ability/willingness to pay, etc.), but the fundamental decision facing the household is a "lumpy" or discrete one of taking this dwelling unit or that one, not the more "continuous" one of taking x or $x+1$ m^2 of floorspace.

Clearly, both floorspace and dwelling units are contained within buildings. Buildings are located upon land, which can be divided into parcels, zones or grid cells. That is, a hierarchy of spatial elements exists, consisting of "occupancy units" (floorspace and/or dwelling units), which are contained within physical structures (buildings), which occupy (sit upon) a finite area of land. In developing ILUTE, we must make design decisions about how to represent space at all three levels of this hierarchy: occupancy, built structures and land.

At the moment two parallel approaches to spatial representation are being investigated by the ILUTE team: a grid-based system and a building-based system.

3.2.1. Grid-based system

The first approach involves the use of a fine grid system (30 m square) to represent space. Occupancy is represented by floorspace, by type (commercial versus residential). Buildings are not explicit entities of the model, but enter implicitly through the supply process which determines the type of floorspace in each cell (single-family housing, apartment, etc.) and the density of development in the cell. One of the principal benefits of this approach is a uniformity in treatment that requires very little overhead in handling and storing the information about the land system while still working at a very detailed level. It is thereby possible in a practical sense to establish a representation that is very close to continuous (although not entirely so) that allows appropriate aggregate behaviour to emerge.

3.2.2. Building-based system

The second approach is a much more direct representation of the elements in the system. Dwelling units, floorspace and buildings all exist as explicit objects. Each residential building contains one or more dwelling units, while each commercial building contains a certain amount of floorspace, which is sub-divided among one or more occupants (mixed-use buildings containing both residential and commercial occupants are also possible). Buildings have both type (high-rise structure, bungalow, etc.) and usage (condominium, office building, mixed use, etc.). Dwelling units have tenure (own or rent) and size (floorspace). Each building has a location, which means it has a geocode attached to it. At the moment, this version of the model is ambiguous in terms of the explicit representation of land. In principle, each building geocode could correspond to a specific building site (i.e. parcel of land or lot), representing a very fine level of spatial disaggregation. On the other hand, the geocode may be the centroid of the census tract within which the building is located—a far more aggregate spatial representation. One advantage of this design is that it is relatively indifferent to what level of spatial aggregation is selected. It is the intention of the ILUTE project team to investigate the effects of different levels of spatial aggregation on model performance within this second version of the model, as well as to compare the first and second approaches with respect to computational performance, empirical accuracy, and theoretical soundness.

3.3. Temporal representation

Another fundamental assumption underlying adopting a simulation approach to urban systems modelling is that in an open, dissipative system such as an urban region system states are path dependent. That is, the system is not, in general, in an equilibrium state (for which equilibrium conditions can be stated and then solved for in order to determine the future system state). Rather it is in a state of dynamic disequilibrium, in which the system is constantly adapting to changing conditions. If this assumption is correct, then the only way that a future system state can be predicted is by explicitly "stepping the system through time" from some known base state, that is, by evolving the system state one time step at a time.

Two classic approaches to handling time in simulation models are discrete event (in which a queue of timed events is maintained and the simulation clock advances to the time of the next event in the queue) and discrete time (in which the simulation clock advances by a fixed increment). In a real urban system, time and decision-making are continuous processes—in an urban systems model, these continuous properties must be discretized. The continuous nature of real-world decision-making makes it ill-suited to a purely discrete event approach since regularly-occurring decisions would need to be entered as "events". The addition of such events would essentially turn the discrete event simulation into a discrete time simulation with the added overhead of maintaining an event queue.

In ILUTE, an adaptation of the discrete time approach is used. In this approach, several levels of "temporal aggregation" can coexist in the simulation since each decision/update process is assigned an update interval. For example, residential location decisions may be evaluated monthly whereas firm location decisions could be evaluated annually. This approach provides great flexibility to the modeller in controlling the level of temporal aggregation.

Within this framework, there is still considerable choice in how decisions and events are processed. Two approaches for processing decisions and events are currently being explored:

3.3.1. Temporal Aggregation Approach 1 (agent-based approach)

In this approach, all decisions for each agent are processed before moving to the next agent. The processing order for the agent types (e.g. households, establishments, etc.) could be random or sequential. Similarly, the order in which the agents of a given type are processed could also be random or sequential. An illustration of this approach is as follows: when a household is selected, the births and deaths of household members are generated, as well as other changes in household composition (e.g. individuals leaving households). Auto ownership decisions are made, as is the choice of whether to look for a new home or not. Once all household decisions have been evaluated, the simulation moves the next household. Once all households haven been updated, the simulator moves to update the establishments. In this example, the possibility of going bankrupt is determined, as are the decisions to expand or contract, to hire or fire employees, and decisions regarding physical location.

3.3.2. Temporal Aggregation Approach 2 (decision-based approach)

In this approach, the simulation processes decisions every N months (e.g. monthly, semi-annually, annually, etc.) as need, data availability, etc. dictate. Each procedure (or decision type) is considered in turn, so, for example, the births are all considered for the month at once, separately from the decisions of household members to leave their household, and separately from the decisions of household to change their level of vehicle ownership. This has the advantage of allowing separate investigation into (and model estimation of) each of the processes being modelled. Household decisions (or person-level decisions for each member of the household) considered include births, deaths, household composition changes, labour force participation and job changes, residential location, and auto ownership. Models of firms' decision processes are not currently implemented in this version of the model.

The activity and travel processes of interest within the model occur on a much finer time-scale of the hour, the day and the week, given the current state of the longer-run attributes such as auto ownership, residential location, employment location, household structure, etc. At the moment in ILUTE, a "typical day" is being modelled for each household in the system, once per year. While clearly raising theoretical and practical issues (how does one define or predict a "typical" day out of the continuous time stream of a person's life?), this is representative of the current travel demand modelling state of the art and is what is most practical to achieve at this point in time. The travel conditions that are calculated as a result of this detailed consideration are assumed to influence the longer term urban processes over the next 12 months. Approaches to modelling activity/travel are discussed further in Section 4.3.

On-going work within the ILUTE project is also exploring the potential for developing activity/travel models for a "typical" week, drawing on novel survey methods developed within the ILUTE group which permit detailed information concerning household activity/travel scheduling behaviour to be gathered over a week's period (Doherty & Miller, 2000). It is felt that such a approach would significantly improve upon our ability to represent household travel dynamics, if the computational burden of such a model is not excessive (Doherty, Miller, Axhausen, & Gårling, 2001; Scott, 2000).

4. Agent decision processes

A range of modelling methods are employed within the modelling system to represent individual agents' behaviours, including simple state transition models, random utility choice models, rule-based "computational process" models, and hybrids of these approaches. A major emphasis within ILUTE is the development of microsimulation models of market demand-supply interactions, particularly within the residential and commercial real estate markets. In addition, travel demand is modelled explicitly as the outcome of a combination of household and individual decisions concerning the participation in out-of-home activities over the course of a day. Each of these issues is discussed in more detail in the following sub-sections.

4.1. Alternative evaluation and choices

Several approaches to modelling human decision-making are employed in ILUTE. The first is random utility theory, derived from micro-economic consumer theory, in which rational decision-makers are assumed to make utility-maximizing choices. As modellers of these choices, however, we are unable to observe these utility-based calculations with certainty. This observation leads to the development of a wide range of probabilistic models of decision-making which have been applied to a variety of travel, location and other choice situations. The most common functional forms for operational random utility models are *logit* and *nested logit* models which are notable for their computational efficiency, ease of parameter estimation, and successful application in numerous operational settings. Ben-Akiva and Lerman (1985) provide a comprehensive discussion of random utility theory and its implementation within logit and nested logit models.

Random utility models have substantial advantages for practical policy analysis. They are continuous, in that an infinitesimal change in policy inputs results in an infinitesimal change in the probability of various actors making various choices. This property is valuable when using the model to understand the influence of small policy changes. The continuous response also allows model parameters to be easily estimated by maximising the likelihood that the model will reproduce observed choice behaviour, with such maximisation performed using standard real variable search processes (Abraham, 1994). Also, for the simpler multinomial logit model, the equation describing the choice from among a subset of alternatives is invariant in the presence of additional alternatives. McFadden (1978) has shown that this equation is identical to the equation that would result if the only alternatives that were available were the ones selected for the subset. This allows random utility models to be estimated from simplified data describing chosen alternatives and only a sample of the available unchosen alternatives. Random utility models also make explicit the value of the various attributes of alternatives, and the willingness of different types of agents to make trade-offs in one attribute of an alternative to obtain a benefit in another attribute. These trade-off rates are useful for understanding behaviour at a conceptual level, and understanding the operation of the model. Equally importantly, the explicit measuring of the value of attributes allows the resulting model to be used for benefit analysis, to understand how different policy scenarios cater to the needs and wants of the full disaggregate sample of individual agents within the model.

In application, random utility models can be quite complex, since the "random" component of utility can be selected from arbitrary distributions, allowing the entire simulation to behave as a Monte Carlo investigation of aggregate behaviour. Standard estimation processes for random utility models, however, rely on integrating over specific distributions assumed for the random components of utility, giving a closed form mathematical representation of choice probabilities. Microsimulation models are thus poised to benefit substantially from new developments in random utility model estimation, but, currently, practical models will typically use standard logit and nested logit formulations and be subject to the limitations thereof.

Random utility models alone are difficult to apply in situations involving very large and/or complex choice sets, such as typically exist in location choice problems and activity-based approaches to travel demand modelling (discussed further below). In such cases, the assumption of joint choice over a large, complex set of alternatives (which is inherent in the random utility approach) is both behaviourally implausible (people tend to be problem-simplifiers, and do not simultaneously consider large number of alternatives, many of which are often minor permutations/combinations or one another), statistically questionable (due to complex correlations among the alternatives which can only be partially handled by the typical model formulations), and computationally inefficient (due to the need to explicitly consider every one of the available choices.) In such cases, psychologists and others argue for rule-based approaches to decision-making (Tversky & Kahneman, 1981; Gårling, Laitila, & Westin, 1998). Many rule-based approaches to decision-making exist, many of which have been applied to various decisions of interest within ILUTE. While theoretically attractive in many respects, a major practical concern is how to develop statistically reliable rule sets for heterogeneous sets of actors. One recent example of a fully rule-based approach to activity/travel modelling, including the statistical estimation of rules from empirical data, is provided by Arentze and Timmermans (2000).

In the ILUTE project, as a general rule we have adopted a hybrid approach. Rule-based search processes are used to reduce large "universal" choice sets to a relatively small set of alternatives which are being actively considered by a given agent. Random utility-based models are then used to provide detailed evaluation of this sub-set of actively considered alternatives, leading to the choice of the "best" (i.e. maximum utility) alternative from among those considered. Note that this represents a "boundedly rational" model of decision-making, since it does not involve explicit search over all feasible alternatives and, as a result, does not guarantee that the optimal alternative (in an abstract, absolute sense) will be chosen. A well designed hybrid model will retain most of the advantages of the random utility theory, but within a more behavioural framework.

4.2. Activity/travel scheduling

The demand for travel is derived from the need to participate in out-of-home activities. Following on from this very simple observation, one can readily build the case for an *activity-based approach* to modelling travel demand, in which the scheduling of and participation in out-of-home activities is explicitly modelled, from which the need for and characteristics of trip-making are then derived. This can be contrasted with conventional, aggregate travel demand models, which start directly with the individual trip (or more typically, the flow of trips from zone to zone) as the unit of analysis and which pay at most cursory lip-service to the derived nature of travel demand. The case for activity-based travel models has been made in detail elsewhere (see, among many others, TTI, 1996), while the state-of-art in activity-based methods is reviewed in TTI (1996), and Arentze and Timmermans (2000), among others.

Agent-based microsimulation is ideally suited to activity-based travel modelling given the disaggregated, dynamic, complex nature of the phenomenon (Goulias & Kitamura, 1992; Miller, 1996; Miller & Salvini, 2001a). In ILUTE, two models of activity/travel behaviour are currently under investigation. For ease of reference, one is labelled a "pattern-based" approach, while the other is labelled an "activity scheduling" approach. Both methods are briefly discussed below.

4.2.1. Pattern-based approach

The pattern-based approach for activity and travel scheduling involves simulating, by selecting from the probabilities of random utility theory models, the choice of individuals from amongst a range of options at various steps, to "build up" the pattern of travel for each individual during each day. The hybrid approaches of using rules or imitation to limit the choice set in a decision (see Section 4.1) are not currently used in the pattern-based approach, so a zonal representation of space is necessary to simplify the location choice sets to a reasonable number of alternatives. Thus, the location choices are represented using two sequential models—a random utility model of the choice of zone, and a simple model of the choice of detailed location (e.g. grid cell) within a zone.

The first step involves selecting for each household member an activity pattern for the day. The activity pattern is a listing of the sequence of activities undertaken by the household member as a series of tours made out from the home (and from the workplace as appropriate). For example, one such activity pattern is 'Home-Shop-Home-Work-Lunch-Work-Recreation-Home', which includes two 'home-based' tours and one 'work-based' tour. Each household member is assigned an activity by selecting from the probabilities of a pattern choice random utility model, with the probabilities influenced by the age and gender of the household member, household income, work and school status, expenditure level and transport accessibilities at both the home location and the workplace location as appropriate. The activities in a given pattern are then assigned durations based on hazard models that are functions of the general nature of the rest of the activity pattern assigned the person, the age and gender of the person, household income, work and school status, expenditure level and transport accessibilities at both the home location and the workplace location as appropriate.

Following this, each home-based tour and work-based tour is considered separately. The tour is assigned a 'primary' destination location using a random utility model influenced by the type of activity the attributes of travel and related accessibilities. After the 'primary' destination is assigned, then the tour mode is selected based on the travel attributes for the round trip from the zone containing the home (or workplace) to the zone containing the 'primary' destination, with tour modes including driving, non-motorized modes, and transit, auto passenger and combinations of auto-passenger with transit.

The networks of links available for the individual trips in the tour, considered in the transport supply module, are dictated by the tour mode that is assigned. For example, if the tour mode is auto drive alone for both the outbound and return components of the tour, then just the road network links are available for each of the trips in the tour.

The number and locations (zones) of any intermediate stops are then assigned to each tour, up to a maximum of one for each of the outbound and return portions of the tour, again using a Monte Carlo process with the selection probabilities determined using logit functions with utility values that reflect the nature of the tour, the tour mode, and the accessibility at the home and workplace. For each auto trip, including both driver and passenger tour modes, a further specification of the number of people in the vehicle is performed using a Monte Carlo process with the selection probabilities determined using logit formulations with utility values that are functions of the relevant travel attributes. This is in order that the trip assignment process can make the right set of link types (with or without HOV) available for the specific trips in each case.

The start time of each trip is then established according to the durations previously assigned the corresponding activities. The origin and destination of each trip are attributed to particular grid cells by randomly selecting from the set of cells with space types consistent with the activity at the stop. In this way the link for the start of the trip and the link for the end of the trip are identified, as required by the transport supply module.

4.2.2. *Activity scheduling approach*

In this approach the activity pattern for each individual in each household is constructed "from scratch", rather than selected from a set of representative patterns. Each person has a series of "projects" (work, school, shopping, in-home activities) which collectively define the universal set of possible daily activities in which this person might engage (Axhausen, 1998). The household containing the person will also have a set of household-level projects, such as child-care, home maintenance, etc. Each project has an ''agenda'' of specific activity ''episodes'' which are candidates to be actually scheduled and executed within a person's activity/travel pattern for the day. Activity episodes are randomly generated for each project for each person based on episode frequencies derived from large sample surveys. Attributes of each episode include type, start time, duration and location. Work and school episode locations are assumed to be known a priori in most cases (workers without a usual place of work do not fall into this case). For other episodes, locations are determined based on a logit destination choice model.

Episodes from the various agendas are ''scheduled'' into each person's plan for the day so as to maintain feasibility (e.g. episodes can't overlap, and it must be feasible to travel from one episode location to the next in the time available) and priority (in the current implementation, priority is assigned by a fixed ordering which schedules ''high priority'' projects first; in the longer term a more dynamic, context-dependent priority assessment is anticipated (Doherty et al., 2000) using a set of scheduling rules. Episode start times and/or durations may be modified during this process in order to maintain scheduling feasibility.

Associated with each activity episode are travel episodes, representing trips from one episode location to another. As with activity episodes, travel episodes have start time, duration (i.e. travel time) and location (in this case, location is two-valued, consisting of the origin and the destination of the trip/episode). In addition, the travel episode has a mode of travel (auto-drive, passenger, etc.).

Mode choice occurs within this model as an integral part of the scheduling process, since the feasibility of a given schedule alternative depends on the travel times to/from activity episodes, which, in turn, depend on the mode chosen to execute each trip. At the same time, the ''utility'' of a given mode depends upon the activity episodes it serves, etc. As in the pattern-based approach, the concept of the tour or trip-chain plays a key role in the combined mode choice/episode scheduling process. Travel modes are determined for home-based and work-based tours, given constraints on feasible modes for each trip on a tour and given auto availability for the given tour. That is, if, for example, two drivers exist in a household which owns one car, and if home-based tours for these two drivers overlap, then only one of the two will be able choose the drive option. In such cases, a ''household vehicle allocation'' model is used to determine which household member gets to use the car and which must use non-drive modes to execute his/her tour (this process generalizes to three drivers, two cars, etc.).

This model is an example of the hybrid decision-making approach, in that random utilities are used to characterize modal options for each trip and tour, and decisions are made so as to maximize household utility. This utility maximization, however, occurs within a relatively complex rule-based scheduling procedure. The full model design and preliminary implementation results are presented in Miller et al. (2002) and Miller and Roorda (2002a, 2002b).

4.3. Route choice and network performance

Many approaches exist for the assignment of vehicles and persons to paths in the transportation network, and thereby also establishing the performance of the networks of transportation supply under these loads. ILUTE is being designed to interface with a variety of these route choice/network performance models. Miller and Hassounah (1993) discuss the interface between integrated land use–travel demand models and network models in some detail. They argue that a dynamic, "mesoscopic" approach to network modelling might be appropriate for such applications, which in terms of detail and complexity lies somewhere between current conventional static equilibrium methods on the one hand and second-by-second vehicle microsimulators on the other. Peiravian (in press) is one example of this approach, in which the existing network modelling software package DYNASMART is being adapted to this purpose.

One contrast between the ILUTE approach to route choice and some of the more detailed "traffic" microsimulations, is that ILUTE adopts a conventional view regarding how individual travellers respond to congestion. Travellers are assumed to know, from past experience, the congested travel conditions on various links, and they make travel choices consistent with this knowledge. Thus the route choice decisions of travellers are assumed to be in some sort of Wardrop equilibrium with the route choices of other travellers (Wardrop, 1952).

One approach being explored involves a new travel assignment algorithm, termed 'micro-assignment', which is an assignment of individual vehicles and heterogeneous individual travellers and trips in a manner which does not require second-by-second vehicle calculations.

Micro-assignment involves loading the individual trips (with randomly varying individual route choice sensitivities represented in the utility values calculated for the links) from specific origin point locations to destination point locations one at a time in order to obtain a dispersion of trips consistent with the range of sensitivities in the population and to avoid the various disadvantages of zonal-trip-matrix loading, such as the artificial overloading of trips at the points where the zone centroid connectors join the road network. Each trip is considered in turn, but to ensure a Wardrop User Equilibrium it is necessary to process the trip list more than once, so that earlier trips have the opportunity to adapt to congestion caused by later trips. Dispersion in routes occurs because individual trip utility function coefficients are presampled from distributions, and perceived travel times on links can also be presampled from distributions. This is consistent with the larger ILUTE approach, in which random utility theory is adopted, but once the error terms on random variables are sampled the choices make by individual agents are deterministic. The approach converges partly because there is no need to resample error terms when a trip is being reconsidered, only to process the route choice decision again based on the new congestion levels.

To ensure conservation of the total number of trips and the corresponding network loadings, it is necessary to retain the path choice for each trip, to allow the removal the previous loading if the agent chooses a different route when a trip is reconsidered as part of the iterative process.

The micro-assignment approach does take substantial computing resources, as a path-building algorithm must be run repeatedly for each trip. To speed up the algorithm the Frank-Wolfe assignment process (Ortúzar & Willumsen, 1994) is currently being used for a sufficient number of iterations to achieve an acceptable level of equilibrium convergence for private vehicle trips. The Frank-Wolfe process requires a trip table based on zones, which is constructed by aggregating the original trip list and averaging some of the characteristics of individual travellers and trips, and adding in the pre-scheduled transit vehicles. Based on the equilibrium set of link flows and travel times from the Frank-Wolfe assignment, the micro-assignment procedure is implemented for each trip without iteration, to allow the unique origin, destination and characteristics of each trip to be respected (as is appropriate in an agent-based microsimulation), and to reduce the reliance on the more abstract "zone" representation of space.

The loading to the transit network is done using a form of the optimal strategies procedure as implemented in the EMME/2 modelling software (INRO, 1996). With this procedure all 'reasonable' paths through the transit network are identified, and the flow of travellers from origin zone to destination zone is allocated among these paths according to their relative attractiveness. It is assumed that travellers arrive at a transit station uniformly during the headway period and board the first reasonable transit route that becomes available. Thus, if competing routes are available, wait time is determined by a composite headway of the various competing lines that provide a path through to the final destination. Currently, the impact of transit user loads on transit services is not considered, eliminating the need to iterate in the transit assignment.

4.4. Microsimulating market interactions

A major research thrust within ILUTE is the investigation of how to model market demand-supply interactions on an agent-by-agent basis; i.e. how to microsimulate market interactions rather than solve for aggregate market equilibrium conditions. This approach is assumed to be an attractive and perhaps even necessary one when dealing with markets (such as housing markets) which may never be in a state of equilibrium, but it is also at this point in time an open hypothesis as to whether totally disaggregated transaction-by-transaction market simulations will robustly reproduce aggregate market performance. As is typical of the ILUTE program, two approaches to this problem are being investigated, where each approach is associated with one of the temporal aggregation schemes discussed in Section 3.3.

4.4.1. Market Micorsimulation Approach 1

In the first temporal aggregation procedure, as households and firms make decisions, they create availability in various "markets" by making "offers". When a household vacates a location it creates a housing opportunity in that location for another entity (an offer of floorspace). When a firm decides to hire, it creates an offer of employment. When people leave households, they also offer themselves as potential housemates (or spouses) in new households. As each household or establishment is considered in turn the decisions in each market are queued up to be considered at a later time.

The frequency at which these queues of offers are considered depends on the market. A lower frequency allows the queues to grow larger, resulting in a larger range of possibilities. For the resale residential housing market, for instance, the selling queue size should be allowed to grow to be approximately as large as the number of outstanding real estate listings at any one time, so that the behavioural units within the model have the correct quantity of opportunities to choose from.

For each market, the queue of offers that create availability are processed first (for example, offers to sell real estate, offers to hire employees). The queue of offers that take up availability are then considered, in random order, and they choose from among the offers that are available.

At the end of the consideration of queues in a market there may be some actors who are left "unsatisfied", and there may be offers that were not taken. The price of the "commodity" in each zone is adjusted based on the unsatisfied demand and supply, so that over time the spatial markets will move toward balancing demand and supply, but without necessarily reaching a point of balance.

This currently operational strategy relies on zonal average prices, and is only used to consider the longer term markets, such as the housing market. Eventually, the strategy will be expanded to consider all markets in all goods and services, including the daily transactions between households and firms, between households and households, and non-market interactions. As well, offers can have individual prices associated with them, to avoid the need for zonal average prices. The price offered by each actor would be based on the actor's past experience with similar offers, and the number and type of similar outstanding offers at the time (Abraham & Hunt, 2001).

4.4.2. Market Microsimulation Approach 2

Similar to the first approach, in the second temporal aggregation approach, as households decide to enter the housing market they become both demanders for a new housing location and potential suppliers of their current dwelling units. Each month, the currently "active" households in the market search over available dwelling units. If a suitable unit is found, the household may decide to offer a bid for this unit. The current owner then decides to accept this bid or wait for a better offer. If the bid is accepted, then the unit is transferred to its new owner, otherwise the bidder and seller move on to consider other possibilities. In each time step (month), currently active participants in the market decide to continue searching or to return to a "passive" state in which they are no longer actively searching (Miller & Haroun, 2000).

In this approach a specific selling price is determined for each dwelling unit transacted, as the outcome of the bid submission/acceptance "game" in which the prospective buyer and seller are engaged. At the end of each time step, average market prices as a function of housing and location attributes are determined through the construction of a hedonic price function, which is regressed over the sales for the given time step. This hedonic price function is then used to influence market activity and bidding behaviour in the next time step. That is, it provides a means of "integrating over" the individual transactions of the previous time step in order to obtain market-level indications of market activity in this time step which can be "fed back" into the next time step's decision-making.

This approach also applies, with appropriate changes in language, etc. to the markets for rental housing markets, commercial real estate, and labour. In the case of labour markets, wages rather than prices are the issue, and these may well be determined by a somewhat different process than the one sketched above (this issue has not yet been addressed in detail in this version of the model).

5. Other model elements

A few key elements in the overall ILUTE system are discussed in this section. These include: modelling the behaviour of the intra-urban economy, modelling the land development process, modelling household auto ownership, and synthesis procedures for creating baseline population inputs into the model. Each of these is discussed in turn in the following sub-sections.

5.1. Establishments in the urban economy

Commercial market interactions play a central role in determining the evolution of urban areas, in at least two fundamental ways. First, economic interactions of consumption and production across the full gamut of economic sectors provide a primary and absolutely essential driving force for city formation and evolution. These interactions play out within an explicit spatial context in that both production and consumption occur at fixed and diverse points in space, and goods and services must be physically transferred from points of production to points of consumption. Thus, to model the flow of goods and services within an urban area (i.e. the demand for transportation services for both people and freight) one must model the workings of the intra-urban economy on a spatial level. Note that these flows of goods/services include the flow of labour services from workers' residential locations to places of employment (i.e. commuting to work) and the flow of consumer goods from retail outlets to residential locations (i.e. shopping travel).

Some current integrated models such as TRANUS and MEPLAN include explicit representations of the intra-urban economy using a spatially disaggregated input/output model (Hunt & Simmonds, 1993) (actually, a "social accounting matrix" since households are included) that explicitly describes the nature of money flows between various aggregate sectors of the economy. To build on the work of these previous modelling frameworks the ILUTE framework synthesizes a population of establishments so that the aggregate flow of money between establishments matches the totals in the input–output matrices that are available. The individual establishment sizes, and the specific technical coefficients describing the money flows, are sampled from distributions, to give a variety of unique establishments of different sizes, in much the same way that a household synthesis procedure produces individual household records while preserving aggregate population characteristics from census data.

The money flows in the input–output representation imply the shipment, sale and consumption of various commodities. Extensive commodity flow surveys (descriptions of the nature of goods and services movement of a sample of establishment) have been conducted, to allow an analysis of truck and commercial vehicle movements and "tours". The intention is to build the goods-and-services movement module to be as functionally as close as possible to the activity scheduling approach to personal travel, while retaining the differences between establishments and households.

The success of business establishments in shipping commodities (and having commodities shipped to them) will determine their success and growth or decline over time. Most establishments rely on customers and/or employees coming to their place of business, so the success of an establishment in attracting travellers is often more important than their success in sending shipments. The representation of heterogeneous spatial markets thus must include offers to sell goods and services, and offers to purchase labour, and the adjustment of spatial prices in those market simulations are critical to understanding the business conditions of establishments, and the resulting success, failure, and location decisions of establishments.

Establishments are not primarily defined based on their employees. They are defined based on the business they conduct, and they purchase labour from employees as part of this business. This means that establishments are more abstract than households, and need to be able to form "out of thin air" based on a business idea. Thus, there needs to be a continual generation and testing of business ideas. To preserve the disaggregate, microsimulation approach, and to limit the number of "bad business ideas" that make it into the model in one year, only to fail miserably in the next, the framework is being designed to support "proto-establishments". A "proto-establishment" is like an entrepreneur pretending to operate a business, observing how the markets would respond to presence of the business, but without actually affecting the business environment. Proto-establishments that seem as if they would have been successful in one year will be introduced as "real" establishments in the following year.

The primary advantage of the concept of the "proto-establishment" is that the full complexity of the model can be used to evaluate alternatives *before* committing the agent to a particular action. If this strategy is successful in the simulation of establishments it can be applied to household decisions as well, so that, for example, households could imagine living in a community for a year to see how their lives would be in that community, before making the choice to attempt to locate in that community. This would provide a richer and, arguably, more behavioural approach to evaluating complex trade-offs between all the facets of important lifestyle decisions.

5.2. Land development

The land development portion of the model must predict the growth over time of the built environment within which household and business activities locate. As with many other portions of the ILUTE effort, two approaches to modelling land development are underway, one corresponding to each of the two spatial representations being investigated: the grid-based approach and the building-based approach. Each of these two approaches is briefly described in the following two sub-sections.

5.2.1. Grid-based approach

As described in Section 3.2, in the grid-based approach, the built environment is described in terms of the amount and type of floorspace located in each grid in the system. The primary task of the land development module in this approach is to adjust the quantity and type of this floorspace over time in response to changes in price. This is done in a highly disaggregate manner, one grid cell at a time.

Three decisions are simulated for each grid cell. The first is whether the cell should be redeveloped or not. The second is choice of the category of space that the cell should be redeveloped into. The third is the quantity of new development in the cell. These three decisions are represented using logit models. The utility functions for each of the alternatives in each choice set are based on prices and vacancy rates established in the modelling of location choices.

The land development model is behavioural, simulating the decisions of land owners regarding how to improve their properties. Land owners (developers) make their decisions based on current prices and vacancy rates, leading to behaviour consistent with the notion that developers react to changes in conditions because they expect future conditions to be similar to the most current conditions.

The changes to grid cells are assigned based on probabilities from random utility models that are functions of the current space type, the age of the development (number of years since the more recent change), the prices for space of difference categories nearby (e.g. in the same zone) and both local (e.g. zonal) and model-area-wide vacancy rates. This leads to a patchwork of different space types and densities, providing a synthetic representation of the diversity in the urban form (which then helps simulate the diversity of trip origins and destinations.) Exact site characteristics (such as the adjacency to specific transportation infrastructure and the content of adjacent cells) can also influence grid transition possibilities.

5.2.2. Building-based approach

In the building-based approach, individual buildings must be generated by the land development model, where each building is characterized by its type, floorspace, location, etc. To date, model development in this approach has focused on the residential housing market, in which a two-stage approach to modelling housing starts has been adopted. In the first stage a "macro" model is used to predict region-wide housing starts at time t as a function of socio-economic variables and exogenous cost shifters. A combination of regression and time series forecasting techniques are used in the macro model, in which housing starts are predicted on a monthly basis so as to capture seasonal effects (Haider, 2002).

Given this regional control total, the second stage consists of a micro model which predicts the spatial distribution the new starts as the outcome of a spatial choice process of real estate developers. This is a microsimulation model in which each developer-agent must decide in each time step:

- whether to start a new building project in this time period or not
- if starting a new project, what type/mix of housing to build, in what amount
- where to build this new project

In the current implementation, a nested logit model is used to model these interconnected decisions, although in future work it is expected that other decision models will be investigated (Haider, 2002).

5.3. Household auto ownership

Auto ownership is an extremely important determinant of travel behaviour. Not surprisingly, households with cars tend to use them, and the more cars in a household, the more travel in total which occurs, with more of this total travel being made in cars. Since many transportation policies are aimed at reducing car usage, including auto ownership decisions within the model is very important. Based on data gathered through a retrospective survey of household auto transactions in the Greater Toronto Area (Roorda, Mohammadian, & Miller, 2000), a dynamic model of household auto transactions suitable for use within ILUTE has been developed (Mohammadian & Miller, 2002a, 2002b). In this model, each year each household evaluates its auto holdings, in terms of the number, type and age of vehicles it currently owns/leases. At this time step, the household may do nothing, trade one of its current vehicles for another vehicle, add a vehicle to its holdings, or delete (sell/scrap) a vehicle from its current fleet. Thus, the household's vehicle holdings evolve over time in response to these yearly decisions. If trading or scrapping a vehicle, the household must decide which one (if it currently owns more than) to trade/delete. If obtaining a new vehicle through a trade or a straight purchase, the household must decide what type (sub-compact, compact, etc.) and vintage (new, used, etc.) of vehicle to buy. In addition to being dependent on current vehicle holdings (i.e. current state), dynamic elements within the model include dependency upon past behaviour (e.g. how long since the last transaction by type of transaction), as well as on changes in other household state attributes (e.g. changes in household size and number of workers in the household). The auto transactions model has been implemented within the random utility (nested logit) framework, although portions of the model have also been developed within a neural network formulation in order to compare the relative merits of the two modelling approaches (Mohammadian & Miller, 2002a).

5.4. Population synthesis

As briefly discussed in Section 3.1, a fundamental input to ILUTE is a "list" of every agent (and its attributes) existing in the system in the base year whose behaviour is then to be simulated over time. This list must be synthesized from more aggregate data in a manner which is statistically valid; i.e. if the list is re-aggregated it should reproduce the original aggregate data, and the probability of generating a specific realization of a synthesized list of agents should be maximally consistent with the known information.

A common procedure currently used in many microsimulation applications is the "Iterative Proportional Fit" (IPF) procedure developed by Beckman et al. (1996). In ILUTE we have found that this procedure can not be used for at least three reasons:

1. The Canadian Census effectively lacks the equivalent of the U.S. Census PUMS files. The IPF procedure is based on the availability of these files which provide detailed, joint distributions across the person and household attributes being synthesized. Some "microdata" information is provided by Statistics Canada, but these files consist of very small samples defined over very large areas (e.g. an entire Census Metropolitan Area), and so are of little direct use.
2. The complexity of the entities being synthesized is very high relative to many applications in which the IPF procedure has been used. ILUTE requires that for each household we have full information about every person in the household (age, sex, employment status, education level, role in the household, etc.), household vehicles (number, type), and the household's dwelling unit (type, tenure, size, price, etc.). In addition, for each worker in the household employment location, occupation, income, etc. must be known. The sheer dimensionality of the synthesis problem seems to us to be well beyond what the IPF procedure can readily handle, both computationally and in terms availability of "full, joint" information. That is, no one data set provides joint distribution information across the full set of variables which must be synthesized. Rather, the full set of information required must be synthesized from a variety of sources, each one of which provides conditional information concerning the relationships between two or more of the variables of interest.
3. An issue associated with the use of Census data for population synthesis that does not seem to be well discussed in the literature is that of random rounding. That is, in order to ensure confidentiality in published data, Statistics Canada randomly rounds all tabulated cell values to have a right-most digit of either 0 or 5. The result of this practice is that two tables for the same variable (e.g. number of families in a census tract with exactly two children) may not contain the same value. This in general makes for additional complexity in synthesizing population attributes from multiple tables, as is required in the ILUTE application and, in particular, further complicates the application of the IPF procedure.

Thus, in ILUTE a more sequential process, such as the one pioneered by Wilson and Pownall (1976), is employed, in which the full attribute set for an agent (say a person) is constructed in a series of steps. In each step, typically an additional attribute is specified for the agent, conditional on the "currently known" attributes for the agent by sampling from an aggregate table which defines a set of conditional probabilities. Fig. 2 illustrates this process for the simple case of three attributes, X_1, X_2, and X_3. The primary advantage of this approach is its practicality: it can be made to work for applications of virtually any complexity. Its possible disadvantages are

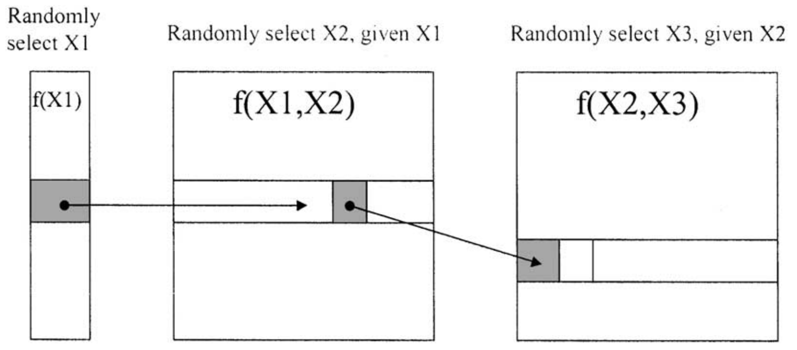

Fig. 2. Population Synthesis Procedure.

that it is inevitably a somewhat ad hoc, cumbersome procedure, and key correlations among variables may be missing (leading to potentially biased results) due to the idiosyncrasies of the available data (i.e. a cross-tabulation between key variables which would provide this correlation information may simply be unavailable).

To date, a base 1986 set of synthesized households, families (ILUTE uses both households and families in its class design) and persons has been generated by this procedure. The procedure used is documented in Guan (2001, 2002). The work to date is viewed as preliminary, and we expect to investigate the performance of the current procedure as well as the potential for improving it in subsequent work.

6. Model development issues

Several general issues exist in developing any large-scale modelling system. These include assembling the base data required to build the model, parameter estimation and model validation, and the computational requirements of the modelling system. Each of these issues is briefly discussed in this section for the case of ILUTE.

6.1. Data

An extensive time-series database must be assembled from a wide variety of sources to support the development of an integrated urban systems model. While this is, indeed, a formidable task, given modern computer-based data systems, it is not an impossible one. In the case of ILUTE the data which have been assembled to support model development activity include the following.

6.1.1. Census data

The fundamental, "base" data for the development of ILUTE is Canadian Census data. Several reasons for this exist, including: it is a highly reliable, large-sample (100 or 20% sample, depending on the variable in question) data set; it is comprehensive, containing a wide variety of demographic, economic, building stock, and (work) travel-related variables; and it is uniformly available for all urban areas in Canada (thereby improving the potential for model transferability). The major limitations of Census data are: it is only available in aggregate, tabulated form for various combinations of variables (i.e. the full, joint distribution of attributes for, say, a census tract is not available, thereby complicating the agent synthesis problem); it possesses limited information about travel (only the place of work and the usual mode of travel to work is known in the Canadian Census; work trip rates, auto ownership, driver's licence rates, and any non-work travel information all are not available through the Census); and the random rounding procedures used in constructing tabulated data (which further complicates agent synthesis). Extensive tabulations of Census data for the Greater Toronto Area (GTA) have been accumulated for ILUTE development for Census years 1971, 1976, 1981, 1986, 1991, 1996 and 2001. This potentially provides a 30-year time period for model development and historical validation. Unfortunately, most other data sets required for model development and validation are typically only available from approximately the mid-1980s onward, so 1986 tends to the base year for ILUTE model development.

6.1.2. Household travel survey data

Canadian cities are fortunate in that most have an on-going program for gathering relatively large-sample (e.g. 5%) one-day travel surveys. A fairly standard telephone interview survey procedure is also used in most cities, thereby ensuring that relatively comparable data is collected from one city to another. While these surveys are relatively simple in content and procedure relative to state-of-the-art activity-based surveys undertaken in many jurisdictions, their large sample sizes provide a statistically robust basis for a wide variety of analyses and model building. These surveys are typically timed to occur in Census years, both so that Census data can be used to validate and weight the sample results, and so that Census and travel survey data can be used in combination as much as possible. In the case of Toronto, the "Transportation Tomorrow Surveys" (TTS) have been undertaken in 1986, 1991, 1996 and 2001 (with an earlier, home-interview-based survey for 1964 also being available for research and modelling purposes), while comparable survey data also exists for Quebec City and Calgary, two of the other "regional laboratories" used in the ILUTE consortium for model development. These travel surveys provide the primary basis for travel-related model development within ILUTE.

6.1.3. Real estate data

The ILUTE group has worked very hard over the last several years to accumulate a comprehensive database on the GTA real estate market. In addition to tapping into publicly available data sets, this has typically also involved entering into data-sharing agreements with private-sector firms active within the real estate market. To date, data sets obtained include:

- Toronto Real Estate Board (TREB) disaggregate resale sales data from 1987 to 2001 with information on housing attributes, prices, and location;

- CanSim database from Statistics Canada, which offers time series data on housing starts, housing and condominium price index, etc.;
- Canada Mortgage and Housing Corporation (CMHC) disaggregate data on freehold housing starts available at enumeration area level, which can be aggregated at the census tract level;
- CMHC aggregate data (by housing type) on housing starts at the local municipality level;
- residential land inventory data obtained from PMA Brethour;
- housing starts/sales data for individual developers from RealNet with information on housing type, prices, and location, and time of sale for the period 1999 to 2001;
- housing starts/sales data by individual developer from PMA Brethour with information on housing type, prices, and location, and time of sale for the period 1994 to 1999; and
- land price and sales data from Teela land sales database.

PMA Brethour, RealNet and Teela are all local companies involved in the acquisition and re-sale of real estate information. Collectively these data sets provide the basis for the development and testing of models of land development and housing market behaviour in the GTA.

6.1.4. Demographic data

In addition to Census demographic data, time-series information on births, deaths, marriages and divorces have been assembled from provincial and federal sources to support the development of demographic sub-models dealing with the "updating" of the resident population (births and deaths) and household formation/dissolution (marriages and divorces). Something which has not yet been acquired in sufficient detail by the ILUTE team is information on in- and out-migration from the GTA. Migration, however, is extremely important in a high-growth region such as Toronto and may prove to be a relatively problematic issue with which to deal.

6.1.5. Special-purpose surveys

Not all information required to develop disaggregate, dynamic models of human decision-making can be obtained from the sources listed above. In particular, it is often necessary to observe actual choices and choice contexts in order to develop individual sub-models (e.g. to develop a model of household auto transactions over time, individual households must be observed over time and their decisions over this time period most be noted). In order to supplement the (typically large-scale) data sets listed above, smaller sample, special-purpose surveys are undertaken as required (and as research budgets permit!). Examples of such surveys which have either been undertaken directly by or are available to the ILUTE team include the following:

1. A 1998 retrospective survey of approximately 800 households in the GTA which gathered information on their automobile transactions during the

period 1990–98, as well as salient household and personal attributes (including employment locations, household composition changes, etc.) over this time period that might be relevant to auto transaction decision-making (Roorda et al., 2000). This data set provided the basis for developing the auto transaction models previously described.
2. A 1998 retrospective survey of 270 households in the GTA which investigated their "housing careers" (i.e. their residential locations over time), along with other changes in household structure, etc. which might influence residential mobility (Haroun & Miller, 1999). These data provide the basis for on-going research into household residential mobility decision-making (e.g. when households decide to move), in combination with the more aggregate real estate data described above.
3. A 1998 survey of 280 households in the GTA which obtained detailed information the spatial search patterns/strategies which these households used during their last residential move (Pushkar, 1998). This data set has been used to test hypotheses about spatial search rules used by households who are active in the housing market (Poon, 2002), and will continue to be of use as the housing market models evolve within ILUTE.
4. Very detailed, computer-based activity diaries have been obtained for very small samples of households (e.g. approximately 40) in both Hamilton (a medium-sized city at the west end of the GTA) and Quebec City using the "CHASE" survey procedure (Doherty & Miller, 2000). The CHASE survey method is now being implemented in a 300-household, 3-year panel survey for the GTA (a reasonably comparable "paper and pencil" activity diary is being applied in parallel to a 300-household sample in Quebec City as well). These data sets not only provide detailed information concerning household-level activity/travel patterns but also have been designed in such a fashion as to provide at least some insight into the dynamics of the household and personal activity scheduling processes which underlie and determine these patterns. These data are being used in the further development of the activity scheduling model described above (Litwin & Miller, 2002).

6.2. Parameter estimation and model calibration/validation

To date, parameters for individual sub-models have been estimated using data appropriate to the given sub-model. A major issue typical of comprehensive, integrated models, however, is the need to "calibrate" model parameters so that the overall performance of the linked sub-models reproduces overall system behaviour satisfactorily. Abraham and Hunt (2000) have dealt with this issue in the case of a conventional, aggregate integrated model (MEPLAN), in which they developed a systematic and robust procedure for searching over a space defined by key model parameters in order to "optimise" model fit at the system level. It may prove necessary as testing of the full-scale ILUTE system proceeds to adopt a somewhat similar approach.

At the moment, however, we are somewhat naively proceeding on the basis that significant errors in overall model performance should be corrected by improving the model structure/specification, rather than "force-fitting" the model to reproduce historical system performance through the "tweaking" of parameters within a fixed model formulation. ILUTE is very much being built as an experimental tool, a "laboratory" if you will, within which we are testing alternative hypotheses about how to best model urban spatial processes. Whether the sheer complexity of the model system being constructed will defeat us in this approach remains to be seen. But certainly one of the strengths we see in an agent/object-based approach to model formulation and programming is that we hope that it will maintain model transparency to far greater depths of model detail/complexity than would be the case for alternative approaches.

Thus, for the time being at least, we are not in the business of "calibrating" (i.e. force-fitting) ILUTE. Validating the model, however, is another and much more fundamental concern. Clearly, such a model will have little credibility with respect to generating long-run forecasts if it cannot at least be shown to reproduce the past with a reasonable level of reliability. A primary motivation for constructing an historical time-series database which is as long as possible is so that we can test the model against historically observed behaviour over as extended a period as possible. At this point in time we have at least a 15-year time period (1986–2001) over which we have sufficiently consistent data to do model test runs. This test period will continue to expand, both as a simple function of time (i.e. we almost certainly will still be in this business by the time the next Census year rolls around in 2006), but also as we work harder at pushing our database back further in time (e.g. ideally to 1971 when our Census time-series first starts).

Thus, the general strategy is to estimate sub-model parameters from specially designed, typically relatively small sample surveys. Where possible, these individual models will be validated against independent, often larger-sample (but also often more aggregate) data. As the overall ILUTE system becomes and operational reality, it will then be tested against large-sample, aggregate time-series system data obtained from the Census, the standard travel surveys, and other independent data sources.

6.3. Computational issues

Despite their considerable computational complexity, agent-based microsimulation models have the potential to be computationally more efficient than conventional zone-based models, especially if even a modest amount of socio-economic disaggregation of the trip-makers, households, etc. is required. The reason for this is that zone-based models must process massive origin-destination matrices (typically further split by each category of each socio-economic variable included in the analysis) to determine system behaviour, while a microsimulation model processes a list of individual agents and their individual attributes. A simple example is the comparison of two mode choice models for the GTA. The conventional, zone-based model used by the City of Toronto computes morning peak-period mode splits for seven modes for all origin-destination pairs in the GTA for 20 worker categories (Miller, 2001). This model currently takes well over an hour to compute on a rather elderly Sun workstation. The prototype activity-scheduling model described above computes not just morning peak-period work trip mode choices, but mode choices for all trips by all purposes for 10 modes of travel for an entire 24-hour weekday in approximately 10 min on the same machine (Miller & Roorda, 2002a).

The single biggest computational challenge that we have encountered in developing large-scale microsimulation models is not processing speed per se, but computer memory limits. The data set describing the urban system state is very large and really must be held in memory at all times. That is, if one is constantly reading from disk to obtain information (e.g. the next batch of household agents to process this time step), then disk I/O time becomes prohibitive. Thus, the primary limitation on the size of the system which can be practically modelled is the number of agents (and other system data) which can be held in memory at one time. Fortunately, cost-effective, very large RAM has become available in the past few years, so that this is becoming less of a serious concern. Very rigorous memory management and very efficient attribute typing (e.g. use of integers or short integers rather than real variables whenever possible, or even byte-level data storage procedures) will, however, remain practical and important issues for the foreseeable future.

7. Summary

This paper has presented an overview of an ambitious research project underway by a group of Canadian researchers to develop "next generation" integrated urban systems models. The "ILUTE" model under development is a fully microsimulation, agent-based modelling system which incorporates all spatial processes affecting land use, location choice, and activity/travel by all actors within the urban system. An explicit part of the research design is to investigate a variety of modelling approaches; that is, to use the microsimulation software being developed as a "laboratory" within which different modelling assumptions and methods can be tested (Miller & Salvini, 1998). This approach is illustrated extensively in this paper, in which typically two approaches to most key sub-models are being activity investigated and tested.

In some sub-models, more traditional "zone based" representations are being used, but always with a view that a zone is a mechanism for averaging and aggregating information for processing, and that relying on aggregates and averages compromises the behavioural and other advantages of a disaggregate, agent-based microsimulation. These compromises are necessary to ensure that the overall vision of a comprehensive and integrated model is obtained, but they are being adopted reluctantly, and with a view to replacement by "purer" versions in the future.

Despite this caveat, the modelling system is highly disaggregated, working with synthesized enumerations of total populations of a complex set of agents (persons, households, firms, vehicles, etc.) within very large-scale urban systems. In the Toronto application of the model, for example, the model is being developed for an urbanized region of approximately 2 million households and 5 million people, providing the model with an extremely strong test with respect to computational feasibility, problem complexity, and data requirements and handling.

The ILUTE model (in its various versions) is still very much a work in progress. Many sub-models are operational, as is much of the overall software system. Results obtained to date are very promising, both in terms of "proof of concept" of the feasibility of large-scale, comprehensive microsimulation models of urban systems, and in the performance of the sub-models developed to date. Much work remains, however, at both the sub-model and modelling system levels before ILUTE will be a fully operational software package suitable for in-the-field planning applications.

Acknowledgements

Early work on the ILUTE project was funded by an NSERC Collaborative Project Grant. The project is currently funded by a SSHRC Major Collaborative Research Initiative grant, the GEOIDE Network of Excellence, the City of Toronto, and NSERC individual operating grants. Contributions to the ideas discussed in this paper from Ron Builing, Sean Doherty, Joel Freedman, Murtaza Haider, Jim Hicks, Pavlos Kanaroglou, Marek Litwin and Matthew Roorda are gratefully acknowledged. Some of the ideas presented here also draw heavily on work being done on the development of a land use transport interaction model of the State of Oregon under the sponsorship of the Oregon Department of Transportation and in conjunction with Parsons Brinckerhoff.

References

Abraham, J. E. (1994). *A nested logit model of the household choice of home location, work locations and modes to work in Calgary*. MSc thesis, Department of Civil Engineering, University of Calgary.

Abraham, J. E., & Hunt, J. D. (2000). Calibration strategies for large-scale urban models. *Transportation Research Record, 1722*, 9–16.

Abraham, J. E., & Hunt, J. D. (2001). Dynamic microsimulation of heterogeneous spatial markets, presented at the *Workshop on Economics with Heterogeneous Interacting Agents*, University of Maastricht, The Netherlands, 7–9 June.

Arentze, T., & Timmermans, H. J. P. (2000). *ALBATROSS: a learning-based transportation oriented simulation system*. Eindhoven: Eindhoven University of Technology.

Axhausen, K. (1990). A simultaneous simulation of activity chains and traffic flow. In P. Jones (Ed.), *Developments in dynamic and activity-based approaches to travel analysis* (pp. 206–225). Aldershot: Avebury.

Axhausen, K. (1998). Can we ever obtain the data we would like to have?. In T. Gårling, T. Laitila, & K. Westin (Eds.), *(1998) Theoretical foundations of travel choice modelling*. Oxford: Pergamon Press.

Barrett, C., Berkbigler, K., Smith, L., Loose, V., Beckman, R., Davis, J., Roberts, D., & Williams, M. (1995). *An operational description of TRANSIMS, LA-UR-95–2393*. Los Alamos, New Mexico: Los Alamos National Laboratory.

Beckman, R. J., Baggerly, K. A., & McKay, M. D. (1996). Creating synthetic baseline populations. *Transportation Research*, *30A*(6), 415–429.

Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H., & Mishalani, R. (1998). DynaMIT: a simulation-based system for traffic prediction, presented at *DACCORD Short Term Forecasting Workshop*, Delft, The Netherlands, February. Available: http://www.rosowww.epfl.ch/mbi/papers/delft_final.pdf.

Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H., & Yang, Q. (1999). Investigation of route guidance generation issues by simulation with DynaMIT. In A. Ceder (Ed.), *Transportation and Traffic Theory, Proceedings of the 14th International Symposium on Transportation and Traffic Theory*, Peregamon Press. http://www.rosowww.epfl.ch/mbi/papers/istttfinal.pdf.

Ben-Akiva, M., & Lerman, S. R. (1985). *Discrete choice analysis: theory and application to predict travel demand*. Cambridge, MA: MIT Press.

Bonsall, P. W. (1982). Microsimulation: its application to car sharing. *Transportation Research A*, *15*, 421–429.

Booch, G. (1994). *Object-oriented analysis and design, with applications* (2nd ed.). Redwood City, CA: The Benjamin/Cummings Publishing Co.

Bradley, M., & Bowman, J. L. (1998). *A system of activity-based models for Portland, Oregon, Travel Model Improvement Program*. Washington, DC: Department of Transportation, May.

Bradley, M., Outwater, M. L. Jonnalagadda, N., & Ruiter, E. R. (2001). Estimation of activity-based microsimulation Model for San Francisco, paper presented at the *80th Annual Meeting of the Transportation Research Board*, Washington, DC, January (proceedings on CD-ROM).

Conder, S., & Lawton, T. K. (2002). Alternative futures for transportation and land use–integrated models contrasted with "Trend-Delphi" methods: the Portland metro results, presented at the *81st Annual Transportation Research Board Conference*, Washington DC, January (proceedings on CD-Rom).

Cuena, J., & Ossowski, S. (2001). Distributed models for decision support. In G. Weiss (Ed.), *Multiagent Systems-a modern approach to distributed artificial intelligence* (pp. 459–504). Cambridge, MA: MIT Press.

Cumpston, R., & Sarjeant, H. (1998). Models of Migration Between Australian States and Statistical Divisions, presented at the *9th National Conference of the Australian Population Association*, September 29–October 2.

Doherty, S. T., & Miller, E. J. (2000). Tracing the household activity scheduling process using one-week computer-based survey. *Transportation*, *27*, 75–97.

Doherty, S. T., Miller, E. J., Axhausen, K. W., & Gårling, T. (2001). A conceptual model of the weekly household activity-travel scheduling process. In E. Stern, I. Salomon & P. Bovy (Eds.), *Travel behaviour: patterns, implications and modeling*, pp. 148–165.

Gårling, T., Laitila, T., & Westin, K. (Eds.). (1998). *Theoretical foundations of travel choice modelling*. Oxford: Pergamon Press.

Goulias, K. G., & Kitamura, R. (1992). Travel demand forecasting with dynamic microsimulation. *Transportation Research Record*, *1357*, 8–17.

Guan, J. (2001). *Population synthesis for the ILUTE model, working paper*. Toronto: Joint Program in Transportation, University of Toronto.

Guan, J. (2002). *Synthesizing family relationships between individuals for the ILUTE micro-simulation model*. BASc thesis, Department of Civil Engineering, University of Toronto.

Haider, M. (2002). *Spatio-temporal modelling of determinants of owner-occupied housing supply*. PhD thesis, Department of Civil Engineering, University of Toronto.

Haroun, A., & Miller, E. J. (1999). A descriptive analysis of a region-wide residential mobility survey, presented at the *46th Annual meeting of the North American Regional Science Association*, Montreal, November.

Harvey, G., & Deakin, E. (1996). *Description of the step analysis package (draft manuscript)*. Hillsborough, N.H: Deakin/Harvey/Skabardonis.

Hu, T. Y., & Mahmassani, H. S. (1995). Evolution of Network Flows under Real-Time Information: Day-to-day dynamic simulation assignment framework. *Transportation Research Record*, *1493*, 46–56.

Hunt, J. D., Johnston, R. A. Abraham, J. E. Rodier, C. J., Garry, G., Putman, S., & de la Barra, T. (2001). Comparisons from the Sacramento Model Testbed, accepted for publication in the *Transportation*

Research Record and presented at the *80th Annual Transportation Research Board Conference*, Washington DC, January.

Hunt, J. D., Kriger, D. S., & Miller, E. J. (1999). Current Operational Urban Land-use Transport Modelling Frameworks, presented at the *78th Annual Transportation Research Board Conference*, Washington DC, January (proceedings on CD-Rom).

Hunt J. D., Kriger, D. S., & Miller, E. J. (in press). A review of current operational urban land-use transport modelling frameworks. Transport Reviews.

Hunt, J. D., & Simmonds, D. C. (1993). Theory and application of an integrated land-use and transport modelling framework. *Environment and Planning*, *20B*, 221–244.

INRO Consultants Inc. (1996). *EMME/2 User's Manual, Software Release: 9.0*, Montreal.

Jonnalagadda, N., Freedman, J., Davidson, W. A., & Hunt, J. D. (2001). Development of microsimulation activity-based model for San Francisco: destination and mode choice models, paper presented at the *80th Annual Meeting of the Transportation Research Board*, Washington, DC, January (proceedings on CD-ROM).

Kitamura, R., & Goulias, K. G. (1991). *MIDAS: a travel demand forecasting tool based on a dynamic model system of household demographics and mobility, Projectbureau Integrale Verkeer-en Vervoerstudies*. the Netherlands: Ministerie van Verkeer en Waterstaat.

Kreibich, V. (1979). Modeling car availability, modal split and trip distribution by monte-carlo simulation: a short way to integrated models. *Transportation*, *8*, 153–166.

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

Litwin, M., & Miller, E. J. (2002). Investigation of household activity behaviour—descriptive Analysis, presented at the *Annual Meeting of the Canadian Regional Science Association*, Toronto, June.

Mackett, R. L. (1985). Micro-analytical Simulation of Locational and Travel Behaviour, *Proceedings PTRC Summer Annual Meeting*, Seminar L: Transportation Planning Methods, London: PTRC, pp. 175–188.

Mackett, R. L. (1990). Exploratory analysis of long-term travel demand and policy impacts using micro-analytical simulation. In P. Jones (Ed.), *Developments in dynamic and activity-based approaches to travel analysis* (pp. 384–405). Aldershot: Avebury.

Mahmassani, H. S., Hu, T. Y., & Peeta, S. (1994). Microsimulation-based procedures for dynamic network traffic assignment, Proceedings of the 22nd European Transport Forum, PTRC. *Seminar H: Transportation Planning Methods*, *II*, 53–64.

Martinez, F. J. (1997). *MUSSA: a land use model for Santiago City*. Santiago: Department of Civil Engineering, University of Chile.

McFadden, D. (1978). Modelling the choice of residential location. In A. Karlqvist, *et al.* (Eds.), *Spatial interaction theory and planning models* (pp. 75–96).

Miller, E. J. (1996). Microsimulation and Activity-Based Forecasting, in Texas Transportation Institute (Eds.), *Activity-Based Travel Forecasting Conference, June 2–5, 1996 Summary, Recommendations, and Compendium of Papers,* Washington, DC: Travel Model Improvement Program, US Department of Transportation and U.S. Environmental Protection Agency, pp. 151–172.

Miller, E. J. (2001). *The Greater Toronto Area Travel Demand Modelling System Version 2.0, Volume I: Model Overview*. Toronto: Joint Program in Transportation, University of Toronto.

Miller, E. J., & Haroun, A. (2000). A microsimulation model of residential housing markets, presented at the *9th International Association for Travel Behaviour Research Conference*, Gold Coast, Queensland, Australia, July 2–5.

Miller, E .J. & Hassounah, M. I. (1993). *Quantitative analysis of urban transportation energy use and emissions: phase I final report*. Report submitted to Energy, Mines and Resources Canada, Toronto: Joint Program in Transportation, University of Toronto.

Miller, E. J., Kriger, D. S., & Hunt, J. D. (1998). *Integrated urban models for simulation of transit and land-use policies*. Final Report, Transit Cooperative Research Project H-12, Toronto: University of Toronto Joint Program in Transportation, September (Web Document 9 at www4.nas.edu/trb/crp.nsf).

Miller, E. J., Noehammer, P. J., & Ross, D. R. (1987). A micro-simulation model of residential mobility. *Proceedings of the International Symposium on Transport, Communications and Urban Form, Volume 2: Analytical Techniques and Case Studies* (pp. 217–234).

Miller, E. J., & Roorda, M. (2002a). *Estimating CO_2 emissions and other impacts of alternative urban forms, report I: summary of modelling methods*. Report to the City of Toronto, Toronto: Joint Program in Transportation, University of Toronto.

Miller, E. J., & Roorda, M. (2002b). *Estimating CO_2 emissions and other impacts of alternative urban forms, report I: summary of base network model test results*. Report to the City of Toronto, Toronto: Joint Program in Transportation, University of Toronto.

Miller, E. J., Roorda, M., Eberhard, L., & Litwin, M. (2002). A microsimulation model of daily household activity and travel, presented at the ''Understanding and Modeling Travel Behavior'' session of the *98th Annual Meeting of the Association of American Geographers*, Los Angeles, 19–23 March.

Miller, E. J., & Salvini, P. A. (1998). The integrated land use, transportation, environment (ILUTE) modeling system: a framework, presented at the *77th Annual Meeting of the Transportation Research Board*, Washington, DC, 12–15 January, (proceedings on CD-ROM).

Miller, E. J. & Salvini, P. A. (2001a). Activity-based travel behavior modelling in a microsimulation framework, invited resource paper, forthcoming in the proceedings of the *Eighth Meeting of the International Association of Travel Behavior Research*, Austin, TX, 1997.

Miller, E. J., & Salvini, P. A. (2001b). The integrated land use, transportation, environment (ILUTE) microsimulation modelling system: description and current status, Chapter 41. In D. Hensher (Ed.), *The leading edge in travel behavior research*, selected papers from the *9th International Association for Travel Behaviour Research Conference*, Gold Coast, Queensland, Australia, 2–5 July, 2000.

Modelistica, (1995). *TRANUS: integrated land use and transport modeling system; version* 5.0. (available on the Oregon Department of Transportation website).

Mohammadian, A., & Miller, E. J. (2002a). A comparison of the performance of nested logit models and artificial neural networks for predicting household automobile choices, presented at the *81st Annual Meeting of the Transportation Research Board*, Washington, DC, January (proceedings on CD-ROM).

Mohammadian, A., & Miller, E. J. (2002b). Understanding and modeling dynamics of household automobile ownership decisions: a Canadian experience, presented at the *ASCE Applications of Advanced Technology in Transportation Conference*, Boston, 3–5 August.

Oskamp, A. (1997). *Local housing market simulation: a micro approach*. Amsterdam: Thesis Publishers.

Ortuzar, J. de D., & Willumsen, L. G. (1994). *Modelling transport* (2nd ed.). New York: John Wiley and Sons.

Otter, H. S., van der Veen, A., & de Vriend, H. J. (2001). ABLOOM: Location Behaviour, Spatial Patterns, and Agent-Based Modelling. *Journal of Artificial Societies and Social Simulation*, *4*(4).

Peiravian, F. (in press). *Road network modelling for environmental impact analysis*. PhD thesis, Department of Civil Engineering, University of Toronto.

Polak, J. K., & Huang, X. (1999). Agent-based modeling system for activity-based models, presented at the *78th Annual Meeting of the Transportation Research Board*, Washington, DC, January (proceedings on CD-ROM).

Poon, W. (2002). *An investigation into residential relocation rules and processes*. MASc thesis, Department of Civil Engineering, University of Toronto.

PTV (undated). *VISSIM traffic flow simulation technical description*. Karlsruhe, Germany: Planung Transport Verkehr AG. Available: http://www.itc-world.com/.

Pushkar, A. (1998). *Modelling household residential search processes: methodology and preliminary results of an original survey*. MASc. thesis, Department of Civil Engineering, University of Toronto.

Quadstone. (1999). *Paramics modeller V3.0 reference manual*. Edinburgh, U.K: Quadstone Ltd.

RDC, Inc. (1995). *Activity-based modeling system for travel demand forecasting (DOT-T-96–02)*. Washington, D.C: U.S. Department of Transportation.

Roorda, M., Mohammadian, A., & Miller, E. J. (2000). Toronto area car ownership study: a retrospective interview and its applications. *Transportation Research Record*, *1719*, 69–76.

Scott, D. M. (2000). *Toward an operational model of daily household activity-travel behavior*. PhD thesis, School of Geography and Geology, McMaster University.

Southworth, F. (1995). *A technical review of urban land use—transportation models as tools for evaluating vehicle travel reduction strategies (ORNL-6881)*. Knoxville: Oak Ridge National Laboratory.

Spear, B. D. (1994). *New approaches to travel demand forecasting models, a synthesis of four research reports (DOT-T-94–15)*. Washington, D.C: U.S. Department of Transportation.

Taylor, D. A. (1990). *Object-orented technology: a manager's guide*. Reading, MA: Addison-Wesley.

Texas Transportation Institute (Eds.). (1996). Activity-based travel forecasting conference, June 2–5, 1996 summary, recommendations, and compendium of papers. Travel Model Improvement Program, US Department of Transportation and US Environmental Protection Agency, Washington, DC.

Timmermans, H., Arentze, T., Dijst, M., Dugundji, E., Joh, C. H., Kapoen, L., Krijgsman, S., Maat, K., & Veldhuisen, J. (2002). AMADEUS: a framework for developing a dynamic multi-agent, multi-period activity-based micro-simulation model of travel demand, presented the *81st Annual Meeting of the Transportation Research Board*, Washington, DC, January (proceedings on CD-ROM).

Tversky, A., & Kahneman, D. (1981). The framing of decision and the psychology of choice. *Science*, *211*, 453–458.

Veldhuisen, K. J., Kapoen, L. L., & Timmermans, H. J. P. (2000). RAMBLAS: a regional planning model based on the micro-simulation of daily activity patterns. *Environment and Planning A*, *31*, 427–443.

Waddell, P. (1998). *Final report on the Oregon prototype metropolitan land use model, phase II*. Report to the Oregon Department of Transportation.

Wardrop, J. G. (1952). Some theoretical aspects of road traffic research. *Proceedings of the Institute of Civil Engineers*, *1*.

Wegener, M. (1983). *The Dortmund housing market model: a Monte Carlo simulation of a regional housing market (Arbeits Paper Number 7)*. Universitaet Dortmund, Dortmund: Institut fuer Raumplanung.

Wegener, M. (1994). Operational urban models: state of the art. *Journal of the American Planning Association*, *60*(1), 17–28.

Wegener, M. (1998). Applied models of urban land use, transport and environment state of the art and future developments, in. In D. Batten, T. J. Kim, L. Lundqvist, & L. G. Mattson (Eds.), *Network infrastructure and the urban environment: recent advances in land use/transportation modelling*. Springer VerlagBerlin.

Weiss, G. (Ed.). (2001). *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. Cambridge, MA: MIT Press.

Wilson, A. G., & Pownall, C. E. (1976). A new representation of the urban system for modelling and for the study of micro-level interdependence. *Area*, *8*, 246–254.

Van Aerde, M., & Yager, S. (1988a). Dynamic integrated freeway/traffic signal networks: problems and proposed solutions. *Transportation Research A*, *22A*(6), 435–443.

Van Aerde, M., & Yager, S. (1988). Dynamic integrated freeway/traffic signal networks: a routing-based modeling approach. *Transportation Research A*, *22A*(6), 445–453.

Vovsha, P., Peterson, E., & Donnelly, R. (2002). Micro-simulation in travel demand modeling: lessons from the New York 'best practices' model, presented at the *81st Annual Meeting of the Transportation Research Board*, January (proceedings on CD-Rom).