Implementation of a “Next Generation” Activity-Based Travel Demand Model: The Toronto Case

Eric J. Miller, Ph.D.
Professor, Department of Civil Engineering, University of Toronto
Director, University of Toronto Transportation Research Institute

James Vaughan
Software Designer, Travel Modelling Group, University of Toronto

David King
Transportation Network Modeller, Travel Modelling Group, University of Toronto

Matthew Austin
M.A.Sc. Candidate, Department of Civil Engineering, University of Toronto

Paper prepared for presentation
at the Travel Demand Modelling and Traffic Simulation Session
of the 2015 Conference of the
Transportation Association of Canada
Charlottetown, PEI

The work reported in this paper was funded in part by the Travel Modelling Group Partners: Metrolinx, Ontario Ministry of Transportation, the Cities of Toronto, Hamilton, Mississauga and Brampton, and the Regional Municipalities of Durham, Halton, Peel and York. It was also funded in part by a NSERC Discovery Grant.
ABSTRACT

Disaggregate, activity-based travel demand models have been promoted for several decades as being more behaviourally sound and, as a result, more policy sensitive demand forecasting tools than conventional aggregate, trip-based, “four-step” procedures. In the past decade numerous operational models have been implemented in a number of US metropolitan regions as well as in Europe. This paper discusses the development and implementation within operational planning practice of the first fully activity-based travel demand model in Canada, the GTAModel V4.0 model system for the Greater Toronto-Hamilton Area (GTHA). Based on over a decade of research at the University of Toronto, “V4.0” has recently been used by the City of Toronto in a major study of transit infrastructure investment strategies. It has also been adopted by the City of Mississauga for use in future planning studies.

The paper discusses the advantages of the disaggregate activity-based approach to travel demand modelling. It then provides a concise overview of the key features and procedures of the V4.0 model system, as well as a detailed bibliography of more detailed documentation of the model system. The paper then discusses the operational implementation of the model system and briefly describes the on-going first application of the implemented model system in the analysis of major rail transit investment alternatives for the City of Toronto. The paper concludes with a few “lessons learned” that may be useful for other Canadian urban regions considering the evolution to activity-based model system formulations for their operational use.
1. INTRODUCTION

An activity-based approach to modelling urban travel demand has long been advocated as more behavioural and policy-sensitive than the conventional trip-based four-step modelling paradigm (Jones, et al., 1990; Axhausen and Garling, 1992; Goulias and Kitamura, 1992; Spear, 1994; TTI, 1997; Miller, 2005d). The basic proposition is quite simple: people travel so that they can participate in out-of-home activities; hence, it is decisions about activity participation that constitute the fundamental behaviour, not travel per se. Travel is a derived demand, with the purpose, timing and destinations of trips (as well as possibly mode) being determined by the activity to be participated in. Further, within the activity-based paradigm travel becomes just another activity (with start time, duration, location, etc., just like any other activity episode). It is argued that the activity-based approach provides a better framework for dealing with complex trip-making (tours), non-work/school and off-peak travel, the dynamics of travel, etc.

Starting with basic research into activity participation and patterns in the 1970s, activity-based models have gradually emerged as operationally practical alternatives to the trip-based four-step approach, particularly over the past decade or so. Loosely speaking, these models fall into two broad categories:

- Tour-based models, in which the tour (a sequence of linked trips that begins and ends at the same location, usually the trip-maker’s home) is the basic unit of analysis (replacing the individual trip as the unit of analysis). Tour-based models are the dominant form of “activity-based” model in the US, with models in operational use in Columbus, Atlanta, Denver, Sacramento, Chicago and San Francisco, among possibly others. Tour-based models have also been developed in Stockholm and Israel, among others (Algers, et al., 1997; Jonnalagadda, et al., 2001; Fosgerau, 2002; Shifman, et al., 2003).

- Activity-scheduling models, in which out-of-home activity schedules/patterns are explicitly generated, with the trips (and trip chains/tours) required to access these activities being generated as needed. Examples of activity-scheduling models include ALBATROSS (The Netherlands; Arentze and Timmermans, 2000, 2005), ADAPTS (US; Auld and Mohammadian, 2009), CEMDAP/SimAgent (US; Bhat, et al., 2004), FAMOS (US; Pendyala, et al., 2004), MATSIM (Germany, Singapore, Switzerland; matsim.org), PCATS (Japan; Kitamura, et al., 2000), and TASHA (Canada; Miller and Roorda, 2003; also see below).

Replacing an operational four-step model with some form of activity-based model is clearly a major decision for a transportation planning agency. Section 2 of this paper discusses a number of issues associated with this decision, including the reasons for considering taking such a step and concerns that must be addressed in doing so. Sections 3, 4 and 5 then present overviews of:

- TASHA, the activity-scheduling model that has been developed for the Greater Toronto-Hamilton Area (GTHA).
- GTAModel V4.0, the overall travel demand model system for the GTHA within which TASHA has been implemented.
- XTMF, the custom software system developed at the University of Toronto for travel demand modelling. GTAModel V4.0 is implemented with XTMF.
Section 6 discusses the on-going operational implementation of GTAModel V4.0 within the Cities of Toronto and Mississauga. Section 7 then concludes the paper with a brief discussion of key lessons learned through the implementation process that may be helpful to other transportation agencies contemplating implementing an activity-based travel demand model system.

2. THE CASE FOR ACTIVITY-BASED TRAVEL MODELS

Many important criticisms of conventional four-step models exist. The first of these is the trip-based nature of four-step models, in which individual trips are modeled in isolation, ignores the interconnections between trips that can be important in terms of determining mode choice, location choice, trip timing, etc. Figures 1 and 2 illustrate these effects in which a transit improvement results in different responses by trip-makers, depending on their activity schedules for the day. In both cases, a trip-based, four-step model will almost certainly not capture the trip-makers’ responses since it ignores tour-effects and activity participation constraints.

![Figure 1: Tour Effects on Travel Choice in Response to a Transit Improvement, Case 1](image)

Second, four-step models focus on the individual trip-maker and ignore household-level constraints and interactions. In Figure 2, the need for the parent to drop off and pick up the child at/from daycare constrains the worker’s mode choice, over-riding standard time-cost trade-offs which form the basis for conventional trip-based mode choice models. Similarly, in Figure 3 competition among household members for the household car results in Person 1 not being able to use the preferred choice (car) for a tour since it overlaps with Person 2’s tour and (in this case) Person 2 has priority over Person 1 for use of the car. Given this Person 1 will need to travel by another mode (e.g., transit) or reschedule the tour to another time when the car is available.
Figure 2: Tour Effects on Travel Choices in Response to a Transit Improvement, Case 2

Figure 3: Within-Household Car Allocation

Third, significant heterogeneity exists in trip-makers’ socio-economic attributes and their associated tastes, preferences and behaviour. As illustrated in Figure 4, significant aggregation
bias can exist if these heterogeneities are ignored or mishandled in the model. Segmenting trip-makers to account for heterogeneity becomes increasingly cumbersome and computationally burdensome as one attempts to add detail to the model system while retaining an O-D trip matrix model formulation. As an example, the previous City of Toronto model uses 1717 traffic zones, yielding 2.948 million O-D zone pairs that need to be processed in the mode choice model. The work trip mode choice model is segmented by 4 occupation groups, 4 age categories and 5 “mobility” (car ownership plus driver’s license combinations), resulting in 80 socio-economic categories that need to be processed for each O-D pair. This means that 235.847 million nested logit model evaluations need to be computed for each iteration of the model to compute the mode choices for approximately 1.5 million morning peak period work trips. It would be far more efficient to enumerate the 1.5 million workers making an AM-peak work trip with their individual attributes and compute mode choices for these individual workers.

Figure 4: Aggregation Bias in Mode Choice Models

Other issues with four-step models include:

- Four-step models do not impose tour constraints on individual trips within a given tour. Most notably, if a car “leaves the driveway” for a tour’s first trip, then this car must eventually return home. Four-step, multi-time period models do not generally impose this constraint, since they do not track the movements of individual trip-makers, only zone-to-zone aggregate flows.
- Tours can only be practically modelled at the level of the individual trip-maker, not within an O-D matrix framework.
- Non-work/school and non-home-based trips are generally poorly handled by four-step models, since the “context” (explanatory factors) for these trips (particularly non-home-based trips) is difficult to determine.
Similarly, off-peak trip-making and peak-spreading are generally poorly handled in four-step models due to their lack of tour linkages and their very aggregate temporal representation. As with socio-economic aggregation/segmentation, their limits to what can be efficiently accomplished through time segmentation of four-step models. Four to five time periods are the most that can be reasonably achieved. Regardless of the number of time periods used, boundary effects between time periods always exist, as do within-period aggregation effects.

Four-step models do not support dynamic route choice modeling well due to their very aggregate temporal representation.

Summarizing, the four-step paradigm is inherently limited in its ability to deal with:

- Tours and tour-based effects on individual trip-making.
- Socio-economic heterogeneity among trip-makers. In particular, the matrix-based nature of four-step approach is increasingly dominated by totally disaggregated list-based (agent-based) approaches in which each individual trip-maker (and his/her attributes) are explicitly modeled.
- Travel dynamics / temporal representation.
- Explicit representation of households (and their impacts on individual trip-making) within the model.
- Non-work/school and off-peak travel.

All of these limitation argue for a tour-based, microsimulation modelling approach, in which the travel behaviour of individual “agents” (persons) is modelled within a household context.

Further, it is well recognized that travel is a derived demand, in which trip-making arises from the need to participate in out-of-home activities. The frequency, timing, purpose and location of trips are fundamentally determined by activity participation processes and decision-making, not by any “need” to travel per se. Activity-based models, in which the focus of the model is on explaining out-of-home activity participation are behaviourally fundamentally more sound than trip-based models that focus on the outcome of activity participation decisions (i.e., the trip to the activity), rather than on the primary process, which is activity participation.

3. **TASHA: Travel/Activity Scheduler for Household Agents**

Key features of TASHA include:

- It is an agent-based microsimulation (ABM) model, which simulates the scheduling of out-of-home activity and travel episodes for each individual within each household within the GTHA for a typical 24-hour weekday.
- It is household-based in that each person explicitly resides within a household and interacts with other household members in terms of: sharing household resources (cars, income), sharing household-level responsibilities (child care, etc.), participating in joint activities, and within-household ride-sharing (see Figure 5).
- It has a very fine temporal resolution, with activity episodes and trips being scheduled continuously (second-by-second) over the course of the 24-hour weekday period.
- It is an activity-scheduler in which individual activity episodes are generated and scheduled over the course of the day. Trips to access each activity episode are generated
as needed. Thus, tours emerge out of the activity scheduling process and can be of arbitrary complexity and type, rather than simply selected from a pre-specified set of activity/tour patterns (see Figure 6).

Figure 5: Households in TASHA

Figure 6: Activity Episode Generation and Scheduling in TASHA
- Mode choice is based on a random utility tour-based model that can accommodate arbitrary tour lengths (number of activity episodes) and complexity. Sub-tours (e.g., a
work-based sub-tour) are accommodated within overall home-based tours. Tour-based constraints are imposed; e.g., cars that leave a tour “anchor point” (e.g., home) must return to the anchor point at the end of the tour. Non-auto-drive tours need not use the same mode on all trips within the tour (see Figure 7).

- Within household ridesharing is explicitly modelled for a wide variety of use cases (see Figure 8).

TASHA was originally developed using 1996 Transportation Tomorrow Survey (TTS) data (Miller, 2002; Miller and Roorda, 2003; Miller, et al., 2005; Miller, et al., 2006; Roorda and Miller, 2006; Roorda, et al., 2006; Roorda, et al., 2009a,b) for the GTHA and was subsequently validated using 2001 TTS data (Roorda, et al., 2008).

TASHA has been designed to interface with various transportation network modelling software and currently interfaces with both EMME and MATSIM (Gao, et al., 2010; Hao, et al., 2010), and can be used as either a replacement for the first three stages of a four-step model (the current application, discussed in this paper), or as a fully integrated component within an agent-based integrated urban model system such as ILUTE (Miller, et al., 2011). In particular, note that TASHA was developed using standard trip survey data (TTS) (although it could benefit from more detailed activity-based data to supplement the standard trip data) and can be run using exactly the same inputs as a standard four-step model: total population and employment by traffic zone and road and transit networks for the scenario being analyzed.

4. GTAMODEL V4.0

The University of Toronto has worked closely over the past 25 years with the City of Toronto (as well as the Ontario Ministry of Transportation and other GTHA transportation agencies) on the development of best-practice regional travel demand forecasting models for the GTHA. As shown on Figure 9, the outcome of this collaboration has been a series of operational regional travel demand model systems labelled GTAModel. In addition to being used by the City of Toronto, the GTAModel development program has had a strong influence on other travel demand models in the GTHA, developed by various consulting firms for the client agencies. GTAModel V2, a four-step morning peak period model system, has been the operational demand forecasting system for the City of Toronto since 2001. V3 was an experimental prototype developed circa 2005-07 that was never operationally implemented. The new V4.0 model system has been developed over the past two years by the University of Toronto Travel Modelling Group (TMG)\(^1\) as a collaborative project funded by all transportation planning agencies within the GTHA.

As shown in Figure 10, V4.0 is a hybrid model system, with the TASHA ABM at its core, supplemented by additional modules required to provide a fully operational travel demand forecasting system. This hybrid approach minimized risk in the adoption of TASHA by building it into the overall pre-existing GTAModel framework. It also accelerated the model system

\(^{1}\)TMG is a travel modelling R&D program at the University of Toronto that is funded by provincial and GTHA municipal transportation agencies. Its mandate is to advance the state of modelling practice in the GTHA. http://tmg.utoronto.ca/About.aspx
Figure 9: GTAModel Historical Evolution

Figure 10: GTAModel V4.0
development process since not all components needed to be developed “from scratch”. The key non-TASHA components are:

- A synthesis module that synthesizes persons, households and cars\(^2\) given the input of total population per zone. It also synthesizes jobs by occupation category and employment status (full- or part-time) from the input of total employment per zone.
- Place-of-Residence – Place of Work (PORPOW) and Place-of-Residence – Place of School (PORPOS) models that allocate work places and school locations to all workers and students, respectively.
- Disaggregate destination choice models to determine locations of non-work/school activity episodes, as these episodes are generated by TASHA.
- Very simple external trip and other special trip generators,
- Road and transit network model components, which include:
  - Emme road and transit assignment modules, applied to five time periods comprising the 24-hour weekday being modelled.
  - A new, disaggregate access/egress station choice model that can deal with auto access/egress to/from rail stations at any point in a tour.
  - Surface transit speed module.

### 4.1 Transit Assignment
Arguably transit assignment has historically received far less attention than road assignment in the literature, whether we are talking about conventional aggregate static models, mesoscopic dynamic models or microsimulation. This undoubtedly reflects the relative unimportance of transit in many auto-dominated urban regions, as well as perhaps the assumption that transit path choice is straightforward to model. In major metropolitan areas such as Toronto, however, in which complex transit systems currently exist and transportation policy largely focuses on issues of transit investment and service operations, behaviourally sound, computationally attractive transit path/route choice modelling is absolutely critical to the overall quality of modelling-based evidence that can be produced for input into policy analyses and decision-making.

Given this, much of our work over the past several years has focussed on developing improved transit route choice models and transit network representations for the GTHA (Kucirek, et al., 2014; Kučirek and Miller, 2014; Miller, et al, 2015). The key design decision underlying this work was to adopt an “integrated”, “technology neutral” representation of the transit network. Transit “sub-modes” such as commuter rail, subway, LRT, buses, etc. are all treated as alternative paths through the transit network, not as potential alternative modes of travel. This representation possesses several major advantages, including:

- It permits construction of a much simplified mode choice model, which contains only two transit modes: “transit with walk access” and “transit with auto access”.
- New modes are readily introduced into the analysis without having to restructure either the mode choice or the route choice model. This is extremely critical in many applications, including Toronto’s, in which new LRT, BRT and RER services are all under very active (and contentious) consideration and in which none of these modes currently exist with the region. This is both a very practical modelling issue and a very

\(^2\) An endogenous household auto ownership model will be added to the next version of the model system, thereby making this an endogenous, policy-sensitive decision.
important political one – it is very important to not to be seen to be biasing the analysis in favour one type of technology or service relative to competing concepts within the model formulation.

- It forces the model-builder to think deeply about the attributes needed to explain trip-maker behaviour in both mode and route choice and about how these can be captured within the route choice / network performance (supply) model.

The big concern with this approach obviously is whether the qualitative differences among transit sub-modes can be adequately captured, especially within the route choice model. At a minimum, it clearly means that a stochastic route choice model is required. As noted in the last point immediately above, it also means taking care to capture as many systematic elements of route choice as possible in the model – i.e., “making observable” as many terms that we usually do not systematically measure and hence are usually content to leave in the error term. While there are clear limits in terms of what one can accomplish in this regard within an aggregate transit assignment model, elements that we have successfully managed to incorporate into GTAModel V4.0 include:

- Fare-based assignment, in which a “generalized time” impedance term incorporates the time-value of fares paid on different routes.
- Congestion/crowding effects onboard transit vehicles. This is very critical in very congested transit systems such as Toronto’s, both from a model estimation/calibration point of view and in order to adequately predict the net impacts of the additional of new capacity to the system.
- A crude “reliability” effect has been incorporated using sub-mode specific boarding penalties.
- Care has been taken to model wait/transfer time as a function of headway as best as possible within the static, non-schedule-based assignment procedure.

5. XTMF: eXtensible Travel Modelling Framework

XTMF is a software system which has been developed by TMG to support the rapid, flexible and extensible development of a wide range of travel demand (and other) model systems. Written in C#, it consists of hundreds of modules that can be combined to build model systems, model parameter estimation software, model run output displays, and interfaces with network modelling software such as Emme. A convenient user interface exists for both developing model systems and then running these systems. XTMF has a fully developed interface with Emme and an extensive toolbox of Emme modules to support road and transit network modelling and analysis. It can be readily extended to similarly interface with other network modelling packages.\(^3\)

TASHA, GTAModel V4.0 and ILUTE are all implemented as XTMF applications. XTMF and the TMG Emme Toolbox are both open source software under the GPLv3 licence and are available on GitHub.\(^4\)

\(^3\) http://www.ecf.utoronto.ca/~miller/TMG-XTMF-Documentation.pdf
\(^4\) https://github.com/TravelModellingGroup/XTMF
6. OPERATIONAL IMPLEMENTATION

All parameters for GTAModel V4.0 and the TASHA “core” of the model system have been estimated and calibrated using 2011 Transportation Tomorrow Survey (TTS) data (DMG, 2014). Documentation of the model system is available in (TMG, 2015), while a validation report is under preparation at the time of this paper’s submission.

The model system has been adopted as the operational travel demand forecasting system by the City of Toronto. It is currently being used to analyze ridership associated with a wide variety of major rail transit proposals for the City. The model system will also be used in a major transportation corridor study within the City of Mississauga, which will be commencing in the spring/summer of 2015.

This operational implementation of an activity-based, microsimulation based travel demand modelling system did not, of course, “happen overnight”. Rather it is the outcome of nearly a 15-year R&D effort. The initial impetus for the development of TASHA was a policy question that the City of Toronto needed to address in 2001. The City wished to intervene in a major Provincial land use decision concerning whether to develop or protect the Oak Ridges Moraine – an environmentally sensitive area north of the City. As part of this intervention they asked the University of Toronto to assess the transportation greenhouse gas (GHG) emissions associated with various development scenarios. In order to do this we said that we needed to build a new model system which:

- Models all trip-making over a 24-hour time period in order to produce a full accounting of GHG emissions (the current GTAModel V2 only modelled the morning peak period).
- Is tour-based, in order to realistically model auto usage over the course of the day.
- Is household-based, in order to realistically model within household auto allocation among household drivers.
- Can be developed within a short time period (6 months) using available data (TTS) given the decision-making time lines and available budgets (i.e., no time or money was available to support new data collection efforts or protracted model development efforts).

Fortunately, we had already been working on a schedule-based and household-based conceptual framework for activity-based modelling (Miller, 2005a,b,c) that generated daily tours as an emergent outcome of the activity-scheduling process. We were thus well positioned to enter directly into detailed model system design based on this conceptual framework in response to the City’s request. This schedule/activity/tour/household-based formulation required an agent-based microsimulation (ABM) formulation – there simply is no other practical way of implementing the disaggregate interrelationships (among activity episodes and trips and among household members).

The urgency for timely input into the policy debate drove many model design decisions that have shaped TASHA to this day:

- Activity episode generation is based on empirical probability distributions derived from the large (5%) samples available in TTS.
- A relatively simple population synthesis is used, also based on TTS distributions.
Use of Emme as the road and transit assignment module. Emme is used by all transportation agencies in the GTHA; implementation of any sort of dynamic and/or microsimulation assignment procedure simply was not a feasible proposition at the time.

This need to respond to policy-driven requests as a motivation for model development is very typical. It is often difficult to motivate planning agencies, etc. to invest in tools before they have a “real and present” need for such tools. Too often it is then too late to respond to the request. In this case, in addition to the conceptual framework already developed, we fortunately also had a massive database in hand with which we were very familiar, calibrated base year road and transit networks for the region (albeit for static, equilibrium assignment models), and a strong working relationship with the sponsoring agency. The result was that we were able to develop the original TASHA model from scratch in six months for a ridiculously small amount of money, and, indeed, did provide input into the Oak Ridges Moraine debate.

Over the next few years we gradually improved the original TASHA prototype in terms of its software implementation, validation of the scheduling algorithm (Roorda, et al., 2008) and enhancement of the tour-based mode choice model (Roorda, et al., 2006, 2009b). We also gradually built “the case” for agent/activity-based microsimulation modelling with regional transportation planning agencies, largely through small environmental modelling exercises that demonstrated the inherent strengths of the approach for emissions modelling, (Miller and Roorda, 2002; Hatzopoulou, et al., 2007; Hatzopolou and Miller, 2010; Hao, et al., 2010; Hatzopolou, et al., 2011), for which it is particularly well suited given its dynamic, agent-based framework. Figure 11 illustrates the use of TASHA for environmental analysis.

![Figure 11: Environmental Analysis Using TASHA](image)
As a result of this string of events, when the City of Toronto was finally in a position in 2012 to support the development of a 24-hour travel demand model to replace their now very old V2 model system, they “of course” wanted a activity/tour-based model and “of course” the TASHA ABM framework made perfect sense to them and did not seem to be an excessively risky approach to which to commit. Further, when the City elected a new mayor in the fall of 2014 with a mandate to move forward quickly with very much needed major transit investments, and when he announced that these investments would be based on sound modelling-based evidence, we were in the position to quickly convince senior planning staff (and through them the mayor) that the new ABM model system was a far superior policy analysis tool for this purpose than the “tried and true” old model system that the City had been using for the past decade.

7. LESSONS LEARNED

Key points to be drawn from this brief history include:

- Building credibility/trust with planning agencies takes considerable time and effort. When the time came, the decision to support the implementation of TASHA in GTAModel V4.0 happened very quickly and easily – but this decision was built on a 10+ year R&D program as well as a 20+ year working relationship with regional agencies working on conventional model systems.
- Once technical credibility is established, ABM is actually a very “easy sell”. The intuitive, “simple story” nature of the model structure – we are simulating a “day in the life” of persons and households is something that senior bureaucrats and politicians can readily grasp. They don’t care about the niceties of random utility theory, system equilibration, etc., etc. – they trust us to worry about these things. What they want is a model that can credibly tell them something about the policies of interest to them in ways that they can readily understand.
- The ability to analyze a wide variety of scenarios (network, land use, fares, etc.) quickly and effectively is essential. Support for modelling can quickly dissipate if questions can’t be answered at all, or if results are not available to meet deadlines, or if the range of options that can be investigated is overly restricted by model run times or other resource constraints. In our current study we are anticipating having to do at least 50-60 model system runs in approximately a month’s time. Run times of a day or more simply will not cut it.\(^5\) ABMs can result in improved computational performance relative to conventional models, providing that computational efficiency is considered at all steps in the model development process.
- The ABM approach actually simplifies many model components. It also makes feasible the modelling of many components of travel behaviour (trip-chaining, ride-sharing, household-level decision-making, etc.) that simply cannot be modelled well using conventional methods.
- A credible, large travel behaviour data base is essential to the development of any advanced model, ABM or otherwise.
- Strong agency support, both in terms of technical staff and management support for best-practice modelling is also essential to the development of advanced models.

\(^5\) A full V4.0 run currently takes about 50 minutes per run. This can be compared “overnight” V2 model runs of 8 hours or more.
REFERENCES


Miller, E.J. 2005c. “Project-Based Activity Scheduling for Household and Person Agents”, in Transportation and Traffic Theory, Flow, Dynamics and Human Interaction, Proceedings of


