A PROSPECTUS FOR IMPROVED LONG-DISTANCE TRAVEL DEMAND MODELS[1]

Eric J. Miller, University of Toronto

Introduction

Intercity/long-distance travel demand models (as opposed to much more common intra-urban models) have always been the neglected orphan of the travel demand modelling world. This is despite the importance of long-distance travel to many policy questions such as: high(er) speed/frequency rail proposals, economic viability of common carrier (rail, bus, air) operations, energy/GHG concerns in passenger transportation, inter-regional/mega-regional economic development, tourism, accessibility/equity issues for rural/remote communities, impacts of new technology (autonomous vehicles, hyperloops), safety, etc. These issues are of particular relevance to Canada, given our needs to not only knit together our dispersed population from sea to sea to sea, but also for strong north-south connections to US and Mexican economic centres.

Over the past 35+ years the author has reviewed the state of art/practice in intercity travel demand modelling approximately once a decade (Rice, et al., 1981; Miller, 1990, 1992, 2004). The findings of these reviews have been depressingly similar: very little progress in modelling methods has occurred over this very long time period, despite considerable advances in intra-urban modelling during this time, and despite continuing interest throughout North America in possible high-speed rail implementations and other major intercity transportation policy issues. Indeed, the 2004 paper ("The Trouble with Intercity Travel Demand Models")could be reissued today without needing to make major modifications to its critique of the field.

Rather than provide yet another detailed critique of current practice, this paper represents a "fresh start" towards developing significantly improved long-distance travel demand models to support policy analysis in Canada and elsewhere. This "fresh start" is predicted on the observation that very significant progress in data, modelling methods and computing technology has occurred, particularly over the past 10+ years, which collectively provide the potential to support significant advances in the field, if public and/or private sector agencies wish to do so. The paper develops its argument in three parts. First, the travel market – long-distance travel – that is to be modelled is explicitly defined. Second, a number of key characteristics of the long-distance travel market that pose significant modelling challenges are enumerated and briefly discussed. Third, a few key propositions are advanced that, it is argued, provide an appropriate starting point for the design of advanced, practical, policy-relevant models of long-distance travel demand.

Problem Definition

The first issue with respect to the design of "non-urban" travel demand models, is the actual definition of the travel behaviour to be modelled. At least four classes of models exist:

- Single-mode models.
- Intercity (corridor) models.
- Province/state-wide models.
- Long-distance travel models.

Single-mode models estimate travel by a single mode with the given mode's network, without fully accounting for competing modes. A very typical and well-developed example of such models are models

---

[1] 54th Annual Meetings of the *Canadian Transportation Research Forum*, May 26 - 29, 2019 at Vancouver, British Columbia.

of air travel demand (Garrow, 2010) that are used by airlines in their planning and operations. While often quite sophisticated in design, these models are generally proprietary and lack the ability to model multi-modal competition, and so are generally not available and/or suitable for public policy analysis purposes. They are, therefore, not considered further in this paper.

Intercity models estimate city-to-city (urban region to urban region) travel, typically within a well-defined, quasi-linear travel corridor. Classic examples of such corridor-based applications include the Boston-New York-Washington “North-East Corridor” (NEC), the San Francisco – Los Angeles Corridor and the “Texas Triangle” (Dallas – Houston – San Antonio – Austin) in the US and the Quebec- Windsor Corridor (QWiC) in Canada. Such models represent the “classic” approach to the modelling of multi-modal long-distance (non-urban) travel, and are the focus of the critical reviews mentioned above. Their origins date back to the NEC studies of the 1960’s. These models are typically custom-built for a specific corridor, often for a specific study, more often than not dealing with high-speed rail projects.

Province/state-wide models attempt to model all travel within a province (Canada) or state (US), including trips to/from neighbouring provinces/states. Beginning in the US in the 1990’s (TRB, 1998), these large-scale models have been developed for a number of US states, and, much more recently, for Ontario (Damodaran, 2017)and Alberta (Karam, 2017) in Canada. These models vary considerably in structure, but generally can be viewed as urban models “writ large” across much larger study areas. To the extent that progress in non-urban modelling has been made over the past twenty years, it has been in the development of such models. In particular, notable US examples are the Oregon (Donnelly, et al., 2018) and Ohio state wide models (Erhardt, et al., 2007) and the California High-Speed Rail Authority model (Cambridge Systematics, 2014). They do, arguably, still struggle to deal with urban versus non-urban (rural, etc.) travel on the one hand as well as short- versus long-distance travel on the other within a single modelling framework. They also may not be well-designed to address typical “corridor” planning issues (such as high-speed rail investment) in cases where corridors are multi-provincial/state (such as QWiC and NEC).

Long-distance travel models. An emerging area of research is on long-distance travel *per se*. Most research to date has been empirical and conceptual in nature, rather than involving the construction of formal, operational models (LaMondia, et al., 2010, 2016). Such a model, however, would represent a generalization of a traditional intercity model, in that it would not be confined to city-to-city travel within a quasi-linear corridor. At the same time, in some respects it would be a restriction of scope relative to current province/state-wide models in that it would not deal with “daily”, short-distance trip-making (whether urban or rural). But it might also be at least spatially an extension of such models in that it need not be restricted to a single province/state. This paper argues for adopting an explicit “long-distance travel” modelling framework, distinct from both the current “intercity” and “statewide” modelling paradigms.

Long-Distance Travel Modelling Challenges

Modelling urban travel demand is, of course, a complex and complicated undertaking. Long-distance travel brings additional challenging issues that need to be addressed within any model design, if this model is be behaviourally sound and policy sensitive. Difficulty in effectively and parsimoniously dealing with these issues has been a major historical barrier to the development of improved models. These include:

- Visitor/non-resident travel.
- Multi-day travel.
- Multi-party travel.
- Luggage.
- Seasonality.

- Trip induction.
- "Sparse" frequencies of trip-making.
- Heterogeneity in trip types and travellers.
- Spatial representation and long-distance network modelling.

Each of these issues is very briefly discussed below.

Visitor/non-resident travel: Visitor/non-resident travel constitutes a greater percentage of long-distance travel within a given network than in the urban case (wherein non-resident travel is often ignored completely). It is not feasible to "model the world", however, and so ways need to be found to model trips to/from and through the long-distance network (however it is defined) that are generated by visitors/non-residents. Modelling visitor/non-resident travel is a major challenge, regardless of modelling framework adopted, that must be addressed in any practical model implementation. It will not, however, be discussed further within this short paper, which focusses instead on the more traditional problem of modelling travel by the residents of the study area being studied.

Multi-day travel: Many long-distance trips involve one or more overnight stays at non-home locations; i.e. the "return-home" trip often does not occur on the same day as the "leave-home" trip. Virtually all existing models are "one-day" models, which are not well formulated to deal with this problem.

Multi-destination travel (tours): Similarly, many long-distance journeys involve visiting more than one non-home destination; i.e., they consist of multi-trip *tours* or *trip-chains*. Best-practice urban models are now generally tour-based in nature, but current long-distance models generally are still single-trip-based. Introducing tours into the model (combined with multi-day travel) adds significant complexity to the model and requires quite a different model structure than is found in current conventional models.

Multi-party travel: Further, many long-distance trips involve two or more people travelling together (business associates, families, etc.). Multi-party trip decision-making is at best very crudely modelled within current models. Note that joint travel within urban models is also generally not well handled. But this issue is much more critical in the long-distance travel case, since it dramatically affects the perception of travel costs (e.g., "sharing" the cost of an auto trip among the family members on a vacation trip versus buying individual airplane or rail tickets for each family member), among other concerns.

Luggage: Multi-day and multi-party trips also have the associated complication of the need to bring luggage on the trip. This can significantly influence mode choice.

Seasonality: Seasonality is generally ignored in urban models, with most models dealing only with a "typical" weekday. This is defensible, since such models are usually focussed on dealing with peak load conditions on the road and transit networks, which occur during the morning and afternoon weekday peak periods. Seasonality is of far greater importance in long-distance travel since demand varies considerably both across days of the week as well as seasonally, with peak loads being seasonally determined (e.g., major holidays and vacation periods) rather than daily (although daily ebbs and flows in demand also occur).

Trip induction: A unique feature of most intercity models that is distinct from urban models, is that they usually include an explicit "trip induction" component, in which the volume of travel demand is sensitive to the levels of service (accessibilities) provided by the intercity transportation system. I.e., unlike the urban case in which "trip generation" is generally inelastic with respect to service levels (congestion, travel times and costs, etc.), intercity models generally predict increased total travel with improved service levels. Thus, for example, introduction of a new high-speed rail line in a corridor will not only divert existing trips from the other modes serving the corridor but will induce new trips to be made that would not have been made in the absence of the high-speed train being implemented. While theoretically very attractive, the nature of these trip induction terms in most current models is typically very *ad hoc* and challenging to defend.

"Sparse" travel frequencies: While cumulatively a large amount of long-distance travel occurs, trip-rates at the level of the individual trip-maker (or even household) are low. While "uber-business travellers" exist who may make long-distance travel on a weekly (or even more frequent) rate (the George Clooney character in the 2009 movie "Up in the Air" comes to mind as the arch-typical example), for most people long-distance travel is a relatively rare event: perhaps a once-a-year vacation, or visiting grandparents on major holidays, or the occasional business trip. This can be contrasted with the urban case, in which the majority of people in an urban region travel every day. A modelling structure is required that can efficiently and appropriately deal with the low probability nature of long-distance trip-making at the daily personal level.

Heterogeneity: As with urban models, the considerable heterogeneityin population socio-economic attributes (and, hence, trip-making propensities) and trip purposes pose huge challenges to developing robust, policy-sensitive models. Typical intercity model trip purposes ("business" and "vacation", etc.) are massively heterogeneous and so are difficult to model with precision, especially given the very simple and aggregate "trip attraction" explanatory variables typically used (often just total population and/or employment at a destination zone).

Spatial Representation and Network Modelling: The large spatial scale and extensive, multi-modal transportation networks involved in modelling long-distance travel pose many practical, challenging problems in the implementation of any operational model. Current intercity models typically use very large traffic zones (sometimes treating entire municipalities as single zones) and often struggle to construct accurate representations of the road and common carrier networks and their service attributes (fares, frequencies, etc.). A sample of issues associated with zone and network design modelling include:

- Properly modelling local access/egress to/from long-distance services (airports, rail terminals, etc.).
- Modelling very long distance auto travel.
- Modelling the effects of relatively low frequency common carrier services (e.g., perhaps only a few flights or trains per day connecting a given origin-destination trip pair).

Addressing any of these important issues is well beyond the scope of this short paper. But it is important to note that, along with possible limitations in data available concerning base travel behaviour for model development, these network modelling issues may pose the greatest challenge to developing significantly improved long-distance models and will require significant R&D effort in any new model development.

Long-Distance Travel Demand Modelling: A Few Propositions

This section presents a few "propositions" that might provide the conceptual starting point for developing a "next generation" of long-distance travel models. The starting points for developing these propositions are the observations that:

- "Classic" intercity/corridor type models are a dead-end that are fundamentally inadequate to the task at hand. These need to be abandoned and replaced with a much more "modern" and sound modelling paradigm.
- Provincial/state-wide models in their current form also are not optimized to deal explicitly with long-distance travel policy questions and network/services design. They may, however, provide an overall "modelling environment" within which a long-distance travel model might "reside".

Defining Long-Distance Travel

Given these two observations, the first proposition is that a clear distinction must be drawn between "local" and "long-distance" travel. It must be recognized that these two travel markets are quite different in terms of travel behaviour, service supply, policy questions, etc., and that they must be modelled separately, although interactions between these two markets – on both the demand and supply sides – may well exist.

Here "local" is used to describe "daily", "routine", "relatively short" travel in which people engage during the execution of their daily activity patterns. Work/school participation, daily/weekly shopping, recreation, socializing, etc. are all well understood activities which we have spent the last 60 years developing increasingly sophisticated and useful models of intra-urban travel. This definition of "local", however, extends to rural areas that historically not been subject to formal modelling, but are being increasingly addressed through provincial/state-wide models. Although trip rates and distances, etc. may vary from the urban to the rural case (e.g., the distance required to visit a grocery store may well be longer, and hence fewer grocery shopping trips may occur per week; travel will be almost completely auto-dominated relative even the auto-orientation experienced in typical suburban areas), the fundamental behaviour remains the same and is subject to being modelled using the same methods as in the classical urban case.

"Long-distance" travel is then travel which is not routine on a daily/weekly basis, often involves multi-day "journeys" between the time the trip-maker leaves and returns home, and, relative to "local" travel, involves a different set of:

- Trip purposes ("business", "vacation", "visit friends and relatives", etc.).
- "Common carrier" services (air, long-distance bus, rail, etc.).
- Behavioural and modelling issues (as discussed in the previous section).

Developing a "crisp" definition of long-distance travel is a non-trivial problem. How long is a trip before it is "long" rather than "short"? Is a GO Train trip from Kitchener, Ontario to Toronto an "intercity" rail trip or an 'intra-urban" commuter rail trip? What of people who live outside the typically defined commuter shed of an urban region but who work within this urban region and who regularly commute to work over a "long distance"?It is the premise of this paper that many of these definitional difficulties can be addressed by asserting that all travel modelled by an intra-regional travel demand model are "local". A similar definition applies to rural travel (whether this travel is explicitly modelled by a formal model or not): if the trip is for "daily/weekly/routine" work, school, shopping, etc. purposes, then it is "local", even if these trips involve travelling distances that appear to be "long" by typical urban modelling standards.

"Long-distance trips" then consist of inter-regional trips, which include such use cases as:

- Trips from one urban region to another urban region.
- Trips from a rural area to/from an urban region.
- Rural-to-rural trips for "non-daily" trip purposes and/or for which long-distance common carrier services are viable options.

This is still a loose definition that would require further tightening in any practical model implementation. It is, in particular, dependent on the definition of urban regional model study boundaries. In the case of the Toronto region, for example, one of the two major urban models in current use is "GTAModel", which takes the Greater Toronto-Hamilton Area (GTHA) as its primary modelling area, while the second major model, "GGH Model" takes the Greater Golden Horseshoe (GGH) as its modelling area – a much larger region. Given the definition above, the Kitchener-Toronto GO Train trip would be classified as a long-distance rail trip by GTA Model and as an intra-regional commuter rail trip by GGH Model. Despite these practical challenges, it is argued that this distinction between "local" and "long-distance" is behaviourally sound and essential if improved models of long-distance are to be developed.

Also note that this definition avoids the "trap" of using a hard distance threshold (50km? 100Km?) to differentiate between "local" and "long-distance" travel. This is a common pitfall of many existing models since it inevitably creates "boundary" or "cliff" effects in the model: a 49.9km trip is "local" while a 50.1km trip is "long-distance". The proposed definition also breaks free of the "intercity/corridor" framing of the study area, to allow for a more general "many-to-many" formulation. At the same time, it does not necessarily require a full province/state-wide model system for its implementation. I.e., one can imagine developing a long-distance model separately from a province/state-wide model, although if such a model exists, it might well represent a sound implementation platform for the long-distance model.

Modelling Framework: Agent-Based Microsimulation

Agent-based micro simulation (ABM) is a modelling approach in which the disaggregate behaviour of a set of agents (persons, households, firms, etc., depending on the behaviour being modelled) are individually simulated over time. It requires the construction of a synthetic list of agents that are statistically representative of an actual population, and then the simulation of these agents' decision-making over a modelled time period as they interact with their environment and other agents. Micro simulation models are the *de facto* standard for best-practice urban travel demand models (Miller, 2018a,b). The second primary proposition of this paper is that an ABM approach is very well suited to addressing the wide range of long-distance modelling issues and needs outlined in this paper. In particular, it is proposed that an ABM model of long-distance travel should involve:

- Synthesizing a complete population of persons within households for the model's study area, which might, for example, be an entire province or state. Although continuously evolving, population synthesis is a very well-studied problem and synthesis techniques are routinely used to generate complete populations for large urban regions involving millions of people and households (Müller and Axhausen, 2010). The GTAModel V4.0 for the GTHA region (Miller, et al., 2015; TMG, 2015), for example, routinely generates populations of up to 10.0 million people depending on the model run forecast year – a number that significantly exceeds the population of every Canadian province or territory except Ontario itself. Synthesizing individual persons with unique socio-economic attributes (income, occupation, age, etc.) enables the model to deal with heterogeneity in trip-making preferences and behaviours. Synthesizing households as well as persons provides the basis for modelling within-household multi-party long-distance trip-making.
- Simulating long-distance travel of each person within each household for all long-distance trip purposes over an extended multi-week time period. Indeed, it is recommended that a year-long (365-day) time period be used in which long-distance travel occurring each day of a forecast year is explicitly simulated. This approach provides a natural and efficient framework for dealing with multi-day and multi-destination trips, as well as seasonality effects.
- An activity-based approach to generating and scheduling long-distance travel as the emergent outcome of the need to travel to engage in activities in distant locations: visit family in the Niagara Peninsula; attend the CTRF conference in Vancouver; spend a weekend in a Muskoka cottage; visit university colleagues in Montreal for a research meeting; etc. As in the urban case, where it has long been recognized that travel is a derived demand and can be best modelled by modelling the participation in out-of-home activities, long-distance travel is similarly activity-based and similarly can only ultimately be modelled well within an activity-based framing.

While many approaches to designing an agent- and activity-based micro simulation model of long-distance travel are conceivable, one "ready-made" solution exists in the TASHA (Travel/Activity Scheduler for Household Agents) (Miller and Roorda, 2003; Miller, et al., 2005; Roorda, et al., 2008). TASHA is currently implemented within the GTA Model V4.0 model system to model intra-urban weekday daily travel for the GTHA region and is in operational use (or in the process of being adopted) by the Cities of Toronto, Mississauga and Burlington and the Regional Municipalities of Durham, Halton and Peel. It is also currently be applied for research purposes in the Helsinki, Finland and Melbourne, Australia urban regions. But the basic conceptual framework of TASHA can be readily translated into the long-distance travel market and it readily extends to multi-day (including year-long) scheduling applications.

A full description of TASHA is well beyond the scope of this short paper. But briefly, TASHA is based on the notion that each person and each household has a set of *projects* corresponding to the different types of activities(and associated goals/objective/motivations) in which they might possible engage. In the case of a long-distance activity/travel model, projects might include: business (for employed persons); vacation/recreation; visit friends/relatives; etc. Each project generates zero to "n" activity "episodes" per time period, where each episode has type (business meeting), start time (in the long-distance case, the day the trip starts), duration (in the long-distance case this can be days/weeks as well as hours/minutes), etc. The activity episodes generated by the projects are then arranged within an *activity schedule* by a *scheduler* for each person for the analysis time period (which in this case may be a calendar year). Trips to/from activity episode locations are generated as part of the scheduling process, including choice of travel mode for each trip. Joint trips by household members are also explicitly modelled.

While such a micro simulation approach may sound very computationally intensive to the point of impracticality, in fact such models, when well designed (including careful parallelization of calculations whenever possible), can be very computationally efficient. Further, micro simulation is the only modelling framework within which tour-based travel, joint household-level interactions and inter-temporal interactions (e.g., if I went on vacation last month, I am not very likely to do so again this month), among many of the other modelling needs discussed above, can be modelled at all.

The other major concern with a micro simulation approach is the question of what base travel behaviour data are available to develop the activity-based demand model components. This is, indeed, a very serious (and potentially limiting) concern. But data availability has been a challenge throughout the history of more aggregate, simplified intercity models, and so arguably is not a decisive objection to rule out the approach *a priori*. Further, new possibilities for obtaining data to support ABM model development are emerging, such as massive (typically third-party) "big data" obtained from cellphone records (and other commercial location-tracking smartphone apps), credit card transaction records, etc., as well as new survey methods for collecting travel information, such as custom smartphone apps and web-based survey methods (Srikukenthiran, et al, 2018). Finally, given the policy importance of the topic and the potentially billions of dollars involved in major intercity transportation infrastructure investment, the historical failure/unwillingness on the part of both public and private stake-holders to gather the data needed to build better models that would help lead to better decisions – and more compelling business cases to justify such massive investments – is both disappointing and mystifying. We can do better if we really want to.

References

Cambridge Systematics (2014) *California High-Speed Rail Ridership and Revenue Model, Version 2.0 Model Documentation*, prepared for the California High-Speed Rail Authority, Cambridge, MA.

Damodaran, S. (2017) "An Integrated Transport – Economics Model for Ontario", presented at the 2017 Annual Conference of the Transportation Association of Canada, St. John's, Nfld.

Donnelly, R., W.J. Upton and B. Knudson (2018) "Oregon's Transportation and Land Use Model Integration Program: A Retrospective", *Journal of Transport and Land Use*, 11(1): 19-30.

Erhardt, G., A. Stryker, R. Anderson, H. Fujioka and J. Freidman (2007) "Ohio long-distance travel model", *Transportation Research Record*, 2003: 130 – 138.

Garrow, L.A. (2010) *Discrete Choice Modelling and Air Travel Demand: Theory and Applications*, Ashgate Publishing: Aldershot, United Kingdom.

Karim, A. (2017) "Alberta Spatial Economic and Transportation (ASET) Model: Transportation Model, March 13.
https://www.cea.ca/files/TriParty-2017-Presentations/Plenary%20C.1%20-%20Ahsanul%20Karim.pdf

LaMondia, J.J., M. Moore and L. Aultman-Hall (2016) "Matching Household Lifecycle Characteristics to Clustered Annual Schedules of Long-Distance and Overnight Travel. *Transportation Research Record* 2594: 11-17.

LaMondia, J., T. Snell and C. Bhat (2010) "Traveler Behavior and Values Analysis in the Context of Vacation Destination and Travel Mode choices: European Union Case Study, *Transportation Research Record, 2156: 140-149.*

Miller, E.J. (1990) "Session II, Intermodal Substitution: Comment", in (D.W. Gillen, ed.) *Canadian Transportation Policy, Policy Forum Series - 18*, Kingston: John Deutsch Institute for the Study of Economic Policy, Queens University, 63-66.

Miller, E.J. (1992) "Intercity Passenger Travel Demand Modelling: Present State and Future Possibilities", *Proceedings of the 27th Annual Meeting of the Canadian Transportation Research Forum*, Saskatoon: University of Saskatchewan Printing Services, 378-389.

Miller, E.J. (2004) "The Trouble with Intercity Travel Demand Models", *Transportation Research Record*, 1895: 94-101.

Miller, E.J. and M.J. Roorda (2003) "Prototype Model of Household Activity-Travel Scheduling", *Transportation Research Record*, 1831: 114–121.

Miller, E.J. (2018a) "The Case for Micro simulation Frameworks for Integrated Urban Models", *Journal of Transport and Land Use*, 11(1): 1025-1037.

Miller, E.J. (2018b) "Agent-Based Activity/Travel Micro simulation: What's Next?", in Briassouli, et al. (eds), *Spatial Analysis: Tools and Land Use, Transport and Environmental Applications*, Springer, 119-150.

Miller, E.J., M.J. Roorda and J.A. Carrasco (2005) "A Tour-Based Model of Travel Mode Choice", *Transportation*, 32(4): 399-422.

Miller, E.J., J. Vaughan, D. King and M. Austin (2015) "Implementation of a 'Next Generation' Activity-Based Travel Demand Model: The Toronto Case", presentation at the Travel Demand Modelling and Traffic Simulation Session of the 2015 Conference of the Transportation Association of Canada. Charlottetown, PEI, Canada.

Müller, K. and K.W. Axhausen (2010) *Population Synthesis for Micro simulation State of the Art*, Working Paper, Zurich: ETH Zurich Research Collection.

Rice, R.G., E.J. Miller, G.N. Steuart, R. Ridout and M. Brown (1981)*Review and Development of Intercity Passenger Travel Demand Models*, Research Report No. 77, Toronto: University of Toronto - York University Joint Program in Transportation.

Roorda, M.J., E.J. Miller and K.M.N. Habib (2008) "Validation of TASHA: a 24-Hour Activity Scheduling Micro simulation Model", *Transportation Research A*, 42: 360-375.

Srikukenthiran, S., P. Loa, S. Hossain, B. Chung, K. Nurul Habib and E.J. Miller (2018)*Final Report*, *Transportation Tomorrow Survey 2.0 research report*, Toronto: University of Toronto Transportation Research Institute, December.

Travel Modeling Group (2015) *GTAModel V4.0, Design, Background and Calibration as of 10/11/2015.* https://tmg.utoronto.ca/documents/#gtamodel-documentation

TRB (1998) Statewide Travel Demand Forecasting Conference Proceedings, Transportation Research Circular E-C011, Washington, DC: Transportation Research Board.