

A Practitioner's Guide to Urban Travel Forecasting Models

David Boyce
Department of Civil and Materials Engineering
University of Illinois at Chicago
842 W. Taylor Street, Chicago, IL 60607 USA
dboyce@uic.edu

Reprinted from the **Proceedings, Metropolitan Conference on Public Transportation Research**, Chicago, June 5, 1998, pp. 309-319.

Introduction

The past five years have witnessed many new developments on the urban travel forecasting model scene. Driven in part by ISTEA, our field has awoken from a lethargy induced by its neglect during the 1980s. A new program of model improvements was launched in 1991, only to be emasculated by a large-scale modeling initiative more related to the end of the Cold War than urban transportation planning. Despite some strange developments, this period has turned out to be an exciting one, perhaps as much so the heyday of the 1960s when our field really got its start.

One result of this increased level of activity is that our field is becoming much more complex, and it never was simple! Until recently, practitioners were mainly concerned with the four-step travel forecasting procedure, and could leave other matters to researchers. But now, several competing paradigms of forecasting models appear to be reaching an operational status. In addition, the environmental community continues to be active in urging and cajoling that improved models actually be used. Finally, computers are expanding in power and size at a seemingly ever-faster rate. In this increasingly complex milieu, both planners and administrators need to make wise choices in the short and long run concerning staff, computing hardware and software systems.

Given this situation, and my perspective of nearly 40 years in our field, I feel challenged to try to offer guidance, if not outright advice. The present situation is confusing enough to researchers who ought to understand the issues, and sometimes even do. My heartfelt sympathies are extended to the practitioner and administrator who appear to have little chance of achieving real understanding, but must make difficult decisions that may haunt them for a long time. I would like to think that they can do better than choosing an expert to rely on, and engaging in debates using concepts and opinions s/he provides, but cannot really explain.

For the past year, from time to time, I have thought that I ought to try to offer a *guide* to the bewildering alternatives confronting practitioners. This paper is a first attempt to do so. If I receive some encouragement, this version may not be the last, but I doubt that it will ever be truly finished. Hopefully, it will not lead to divisiveness and arguments, although I fully expect that it may. My intent is simply to be factual. But, of course what I see as facts the next person sees as my biased opinions.

This version of the paper is organized in this way. First, I try to describe several perspectives on forecasting urban travel. Any given model is likely to adopt one of these perspectives as its primary organizing framework, but probably embodies some aspects of all three. Second, I seek to describe the requirements for forecasting models from a practitioner's viewpoint, as I understand them. Finally, I try to relate the three approaches, as represented by specific models, to the requirements. Earlier, I had imagined that I would then examine the implications for travel forecasting software, but it is already clear that this is too ambitious for this version, at least.

Perspectives on Travel Forecasting

I distinguish three approaches to travel forecasting, which I will designate as follows:

1. *individual travel choice behavior* based on consumer choice theory as found in micro-economics and psychology; this approach is largely from the perspective of the trips made by each individual, taking into account constraints imposed by others.
2. *daily activities incorporating time-space constraints*, as found in human geography; this approach emphasizes more the activities performed by the household as a decision unit, and the interactions of its members.
3. *equilibria of aggregate travel and location choices*, as found in operations research; this approach is from the transportation and spatial systems perspective, especially as it pertains to the joint effects of individual choices such as congestion.

As a form of shorthand, I will refer to these perspectives as *trip-based*, *activity-based* and *equilibrium-based*, respectively. Next I attempt to describe my understanding of these three perspectives, and offer some references to recent research and past models.

Individual travel choice behavior defines the individual traveler as the decision maker, and the trip as the decision unit. The attributes of the trip are regarded as given, often with some perception errors. The basic decision rule is that the traveler seeks to maximize her/his utility in choosing among alternatives. If the perception errors are distributed in a certain manner, the familiar multinomial logit function is the resulting model. This all seems very familiar now; however, as recently as 20 years ago it was reasonably novel in our field.

In contrast to the other perspectives, this one focuses strongly on the individual and views the effects of other travelers as constraints or as fixed. If the other travelers are household members, then their effect may be represented as explicit constraints such as auto availability. If the other travelers are the traveling population in general, then their effects are regarded as exogenous. An example is the way that congested travel times are represented exogenously in this class of models.

The *daily activities* perspective focuses on the activities performed by the household, and its associated travel requirements. These activities, which are distributed through geographical space and 24-hour time during a typical weekday, are considered as an interlocking system of

constraints at the household level. Analyses and forecasting models mainly focus on gaining an understanding of choices made by surveyed households and their members. As with the trip-based perspective, travel times are exogenous to the framework.

In contrast to the first two perspectives, the *equilibrium* perspective begins with a systems view of travelers and the transportation network. The interactions of travelers making trips on the network, and network flows, travel times and costs, are the focus of this approach. Hence, the travel times and costs experienced by individual travelers are endogenous to the model, in contrast to the first two perspectives. Sometimes, the numbers of trips by purpose departing from and arriving to each small area are regarded as exogenous in this perspective. In other versions, however, trip frequencies are also related to travel times, costs and other variables.

To grasp further the differences and similarities among the three approaches, two additional points are offered. The first concerns the time-of-day period for which the forecast is intended. A basic tenet of urban travel analysis is that travel conditions differ during the various time periods of the day. The morning and afternoon peak periods, during which travel to and from workplaces tends to dominate, are the most congested periods. The mid-day and evening periods are more concerned with business, shopping and recreational travel, as well as goods movement, and may also be moderately to heavily congested. In the trip-based and equilibrium approaches, trips are readily classified by time periods on the basis of travel surveys. Then, the various travel choices can be represented and the models solved. These choices can, in principle, include time-period choice, thereby representing the peak-spreading phenomenon. In the traditional trip-based approach, however, forecasting 24-hour travel patterns continues to be the prevailing paradigm. This practice can only be regarded as an anachronism from a former era when computer limitations required that all travel in a 24-hour period be forecast on a network with link capacities equivalent to those of the peak-period. In contrast, the organization of daily activities and travel between pairs of activities is the basic premise of the activity-based approach. A more dynamic orientation seems implicit in this perspective. Whether it is achieved remains to be examined.

Another dimension of this discussion, and source of confusion, is the difference between model *formulations* and *solution procedures*, or *algorithms*. *Model formulations*, which may result from derivations from more basic principles and assumptions, generally consist of systems of equations and inequalities. A simple example is the multinomial logit choice model, a mainstay of the trips-based approach. Alternatively, a model formulation may consist of a set of decision rules and functions describing travelers' behavior, or a mathematical problem whose solution properties corresponds to behavioral assumptions. A *model solution procedure* is a series of steps, usually computerized in some manner, which yields the desired forecasts. A spreadsheet solution of a logit model is a simple example. Likewise, *microsimulation* is a solution procedure. And an *algorithm* for solving a constrained optimization problem whose solution corresponds to certain route choice behavior assumptions is also a solution procedure.

Having described the three approaches, and distinguished between formulations and solution procedures, it is interesting to ask what is the venerable four-step travel forecasting procedure. On the one hand, the individual steps are model formulations, each with its own solution procedure. On the other hand, the entire four-step procedure may be regarded as a

solution procedure for an unstated model formulation, when viewed from the viewpoint of satisfying requirements for consistency among the steps, also known as *feedback*. Although the first interpretation may be regarded as the prevailing one, I believe the second interpretation is more accurate and helpful. When viewed as a solution procedure, the meaning and limitations of the entire four-step procedure become more transparent.

Before turning to a statement of the requirements for urban travel forecasting models, a brief review of current sources of the three perspectives may be helpful. The perspective of individual travel choice behavior is represented by the book of Ben-Akiva and Lerman (1985), *Discrete Choice Analysis: Theory and Application to Travel Demand*. The activity analysis perspective is currently represented by the papers of Kitamura and Pas, of which Kitamura et al (1996) is one recent example. The equilibrium perspective is represented at present by Patriksson (1994), *The Traffic Assignment Problem: Models and Methods*, and by Boyce and Daskin (1997), "Urban Transportation," a chapter in *Design and Operation of Civil and Environmental Engineering Systems*. Oppenheim (1995), *Urban Travel Demand Modeling*, effectively combines the individual travel choice behavior and equilibrium perspectives. The conventional four-step procedure is described by Ortúzar and Willumsen (1990), *Modelling Transport*.

The fact that three approaches are needed to describe our field also suggests that no one individual modeler or school of modelers knows all three approaches well, or is even comfortable with all three. Clearly, ideas are borrowed by each perspective from the others, both in model formulations and solution procedures. Nevertheless, substantial differences exist among the three. This seems especially true for the networks-based equilibrium approach.

Requirements for Urban Travel Forecasting Models

To assess effectively the various alternative models available to professional practitioners, it is useful to define the requirements that should be met by each model. In my view, these requirements can be stated in the following way.

1. The dimensions of choices to be represented in a complete model are: locations of urban activities; choice of level of expenditure for travel by individuals and households, including the number and type of autos owned; frequency of travel between activities; choice of travel destinations/activities; choice of mode, submode and route; and choice of time period of when to travel.
2. Individual and household choices should depend on the aggregate effects of those choices with regard to travel times and costs on the transportation networks, demand for parking and access to destinations.
3. The solution variables of the model should be appropriate for subsequent analyses of land requirements, transportation system performance, atmospheric emissions, financial feasibility and related issues.

4. The model should be capable of being estimated (calibrated) with existing or prospectively available data, and validated independently of that data.
5. The model should be capable of being solved repetitively with current or prospective computing platforms and software systems that are appropriate for professionals in terms of their acquisition cost and solution time requirements.

I know of no modeling capability that satisfies all of these requirements at present. This is important since the choice of a model will undoubtedly always be a compromise between what is required or desired and what is available at the present state of the art.

Description of Models in Terms of the Requirements

Four sets of modeling activities are described in this section, as follows:

1. Four-step sequential procedure based on discrete choice models and solved with feedback;
2. Sequenced Activity Mobility Simulator proposal of Kitamura, Pas et al;
3. TRANSIMS, the micro-simulator under development at Los Alamos National Laboratory;
4. Combined network equilibrium models including discrete choice functions and trip chains.

The first set represents the trips-based approach as implemented in the traditional four-step procedure. The second and third sets are two distinct implementations of what I understand to be the activities-based approach. The fourth set represents the equilibrium-based approach. Each of these model sets is now briefly described in terms of the requirements listed above. An attempt to summarize these descriptions is presented in Table 1.

Four-Step Sequential Procedure with Discrete Choice Models and Solved with Feedback

Being the most familiar to practitioners, this trips-based model set perhaps requires the least description in terms of requirements 1-5 listed above.

1. Trips are categorized by purpose, and forecast in terms of the number of daily departures and daily arrivals in relation to zonal household and personal attributes and zonal activity levels. Auto availability is forecast from household attributes. Trip departures and arrivals are connected into daily trips, which are then allocated to modes and submodes. Auto trips are allocated to routes, resulting in 24-hour link flows and corresponding speeds. Each step of the procedure is defined as a separate model. Discrete choice functions are typically employed in connecting departures to arrivals (trip distribution) and in allocating these trips to mode and submode (mode choice). These two models may be linked by using an appropriate aggregation of the modal times and costs. Sometimes discrete choice models are also used to forecast auto availability.

2. Zone-to-zone travel choices by mode and route depend on the aggregate effects of those choices through the *feedback* mechanism. As generally implemented, there is no assurance that such feedback will lead to a stable or consistent result for the several travel choices.
3. The principal solution variables of the four-step procedure are roadway link flows, typically for a 24-hour day, which are sometimes subsequently allocated to time-of-day periods. Other solution variables include such summary measures as vehicle-miles traveled, trip length frequency distributions in terms of travel time and distance, public transit boardings and revenues, etc.
4. The parameters of each model (corresponding to a given step of the procedure) are generally calibrated from household survey and census data for a base year independently of the other models/steps. The entire procedure may be validated from roadway link counts (usually 24-hour averages), public transit boardings, independent estimates of vehicle-miles traveled, etc.
5. Generally, the four-step procedure is now solved with a commercially available transportation planning software system, using a Personal Computer with a large memory and disk drive. In some applications UNIX workstations are used.

Sequenced Activity Mobility Simulator (SAMS)

As one example of an activity-based approach, I consider the recent proposal of Kitamura et al (1996) for a sequenced activity mobility simulator (SAMS). Since no operational model yet exists, this description can only be based on the authors' conceptual thinking. The proposal is to implement a sequence of micro-simulators pertaining to land use and development, socio-economic/demographics, vehicle transactions, activity-mobility and dynamic networks. Only the activity-mobility simulator (AMOS), described as the heart of SAMS, has been implemented as a prototype in the Washington, D.C. metropolitan area (Pendyala, et al, 1997). The activity-mobility simulator begins with a set of activity-travel patterns, simulates each individual's adaptation process to external variables, yielding a forecast of how individuals and households adapt to a new environment, such as a change in the pricing of the transportation system.

The following comments represent my attempt to describe SAMS in terms of requirements 1-5.

1. At a conceptual level, SAMS seeks to describe and forecast all pertinent activity and travel variables for individuals and households. The forecast is performed by simulating the choices of each individual from behavioral rules derived from activity-based household surveys and stated responses to hypothetical changes in the transportation system. The forecasted choices are summarized in terms of trip purpose, mode, trip duration, time-of-day period, etc. Detailed route choice behavior is not included in the current implementation.
2. In the conceptual description of SAMS, elaborate feedbacks from a dynamic network simulator to other subsystem simulators are depicted. In this general description, no mechanisms for achieving a stable and consistent solution between the activity and

transportation simulators are depicted. The use of an existing (static) equilibrium assignment model is suggested as an interim measure.

- 3 The proposal for SAMS includes all the desired solution variables, as well as an emissions module. Since the emphasis of the present implementation is on activities and trip chains, as contrasted with route choice and network variables, however, there is no basis for expecting that dynamic forecasts of link flows and speeds will be available from this effort.
- 4 Pendyala, et al, 1997 describe in detail the estimation of the activity-mobility simulator (AMOS) with household surveys and stated intention surveys. Their model estimation is based on a sample of 98 commuters, together with standard zone-to-zone travel time estimates, using neural network methodology. Validation studies are not described.
- 5 Computational requirements and performance are not described by Kitamura et al (1996) and Pendyala et al (1997).

The above points are based on an early and relatively general description of the SAMS concept.

Transportation Analysis and Simulation System (TRANSIMS)

TRANSIMS is described as a set of integrated analytical and simulation models and supporting data bases for predicting travel of individual household, residents and vehicles (TRANSIMS Travelogue, 1996). This activity-based model consists of four components: household and commercial activity disaggregation, intermodal route planner, transportation microsimulation and environment (primarily air quality). In the first implementation and testing, roadway traffic microsimulation of a 25 square-mile area is emphasized.

The first component generates synthetic households from Census of Population and Housing files using the iterative proportional fitting technique. Drawing on activity analysis concepts, household activities, priorities, locations and times, and mode and travel preferences will be forecast. The second component, the intermodal route planner, seeks to find a travel plan for each individual that optimizes the traveler's goal function. The third component represents the transportation system. The implementation reviewed for this paper uses a cellular automata simulation. Finally, the emissions component translates travel behavior into emissions and energy consumption at various scales.

Technical papers concerning TRANSIMS were obtained from the TRANSIMS WWW site: <http://www-transims.tsasa.lanl.gov>. Papers by Nagel and Barrett (1997) and Rickert and Nagel (1997) are the basis for describing this model in terms of requirements 1-5.

1. In the papers reviewed, only route choice and roadway variables are discussed. Link and route travel times from a previous solution of the model are averaged over 15 minutes periods, and a time-dependent shortest route algorithm is applied to find best routes (plan sets) from trip origins to destinations within the 25 square-mile test area. The roadway system consists of 914 one-way arterial street links and 374 intersections, of which 93 are signalized. Each link is divided into cells of 7.5 meters in length. Each vehicle is represented separately by its cell location, its speed and the distance to the preceding vehicle

for every time step (second). Accelerations and decelerations of vehicles are not explicitly modeled. The solution variables of the simulation include each vehicle's travel time through the test area, vehicle locations at each time step, and various measures of delay and gridlock.

2. No mechanism is provided in the simulation model for adjusting the route plans in response to network delays. A heuristic averaging procedure is described for iteratively adjusting trip plans to congestion on the network. No measures of convergence are offered, but substantial oscillations caused by rerouting are noted.
3. Rather detailed solution variables are provided by this procedure with respect to emissions modeling. Still, they are not as detailed as requested in recent years by air quality modelers.
4. Although detailed maps of vehicle locations are included in the papers reviewed, no validations are offered. Since the behavioral model pertains only to route choice, and no route choice data are available for the test area, evidently no parameter estimation is possible.
5. A detailed discussion of the computational performance of the model is provided, which are summarized briefly as follows. The tests were performed using five Sun Sparc 5 computers connected in parallel. To perform a simulation of a 5-hour period, with 10 iterative adjustments of route plans, a total of 5 days of computing effort was required. Altogether, the results reported in the paper required 25 days of computing. Evidently, the results were not considered satisfactory because of oscillations.

Combined Network Equilibrium Models with Discrete Choice Functions and Trip Chains

This section describes an example of the equilibrium-based approach, also termed *network equilibrium models* or *combined models*. Network equilibrium models stem from the pathbreaking work of Beckmann et al (1956). In this approach one seeks to include all travel choices in one unified problem formulation, in contrast with other approaches in which each travel choice is the subject of a separate model. The formulation is solved by a single algorithm, insuring that the travel choice variables and travel times and costs are internally consistent.

Network equilibrium models have been applied in both academic research and professional practice. The fixed demand version of the model is the standard traffic assignment model in use today. The variable demand version, or Combined Model of travel choices, has been recently applied by Resource Systems Group (1997) to a tollway planning controversy in Northeastern Illinois. The description of the Combined Model in terms of requirements 1-5 is based on this application and ongoing applications at the University of Illinois at Chicago.

1. In recent applications of the Combined Model, choice of destination, mode, route and time period of travel for several trip purposes are represented in a unified framework. Generally, the destination, mode and time period choice models are formulated as logit functions. Route choice is presently represented as user-optimal assignment. Location of activities, trip frequencies and auto ownership are exogenous. However, Safwat has extended the basic formulation to include the frequency of travel (trip generation). Anderstig and Mattsson have implemented a model for Stockholm that combines residential location of households and

location of firms with travel choices of employees traveling from home to work. A version of the model based on trip chains is presently being implemented with software of PTV VISION, a German transportation planning software system (Fellendorf et al, 1995).

2. In the Combined Model, zone-to-zone travel choices depend on the travel times and costs resulting from those choices, which are fully internal (endogenous) to the model. The solution algorithm for the model insures that the choices and costs are stable and consistent, thereby automatically fulfilling the so-called *feedback* requirement.
3. The solution variables of implemented models are equivalent to those of the four-step procedure, except that the model is typically applied to time-of-day periods, rather than the 24-hour weekday. Thus, link flows, link speeds and link travel times and operating costs on the roadway network are constant over this period.
4. Combined Models are estimated from household travel surveys using maximum likelihood methods with separate parameters for each trip purpose. Validation studies based on independent estimates of vehicle-miles traveled and other data have been performed.
5. Combined Models may be solved as an EMME/2 macro on both large PCs and UNIX workstations (Metaxatos, et al, 1995). To my knowledge, the Combined Model cannot be solved with other software systems at present. Research-oriented computer codes are also being applied to investigate and evaluate alternative parameter estimation methods. A 650 zone model of the Chicago Region with four trip purposes has been solved (10 iterations) in an actual planning application with an EMME/2 macro in about three hours on a Sun Microsystems SPARC 20.

Conclusions

It is not the objective of this paper to recommend which model to apply in practice. Rather, the intention is to clarify the properties of various approaches in terms of requirements of professional practice.

Unrealistic claims are sometimes made for new approaches to travel forecasting, in an effort to secure funding for model development, implementation and testing. I observe that practitioners are sometimes confused by these claims, and may accept them as model properties, rather than unsatisfied requirements. This performance gap seems to accompany a top-down, or requirements approach, to model development in which *desired* model properties are stated in advance of research, implementation and testing. For this reason, academic researchers tend to prefer a bottom-up approach in which models are first formulated, then implemented, tested and evaluated, and finally described in terms of their performance.

Acknowledgement

The support of the National Science Foundation through the National Institute of Statistical Sciences is gratefully acknowledged.

References

- Beckmann, M., C. B. McGuire and C. B. Winsten (1956) **Studies in the Economics of Transportation**, Yale University Press, New Haven.
- Ben-Akiva, M. and S. Lerman (1985) **Discrete Choice Analysis**, MIT Press, Cambridge.
- Boyce, D. E. and M. S. Daskin, Urban Transportation, Ch. 7 in **Design and Operation of Civil and Environmental Engineering Systems**, C. ReVelle and A. McGarity (eds.), Wiley, New York.
- Fellendorf, M., T. Haupt, U. Heidl and W. Scherr (1995) PTV VISION: Activity based demand forecasting in daily practice, paper presented at the **Travel Model Improvement Program Conference**, Daytona Beach, PTV System Software and Consulting GmbH, D-76131 Karlsruhe, Germany.
- Kitamura, R. E. I. Pas, C. L. Lula, T. K. Lawton and P. E. Benson (1996) The sequenced activity mobility simulator (SAMS): an integrated approach to modeling transportation, land use and air quality, **Transportation**, **23**, 267-291.
- Metaxatos, P., D. E. Boyce, M. Florian and I. Constantin. Implementing combined model of origin-destination and route choice in EMME/2 System, **Transportation Research Record**, **1493**, 57-63.
- Nagel, K. and C. L. Barrett (1997) Using microsimulation feedback for trip adaptation for realistic traffic in Dallas, Las Alamos National Laboratory, Los Alamos, NM.
- Oppenheim, N. (1995) **Urban Travel Demand Modeling**, Wiley, New York.
- Patriksson, M. (1994) **The Traffic Assignment Problem**, VSP, Utrecht.
- Ortúzar, J. de D. and L. G. Willumsen (1990) **Modelling Transport**, Wiley, Chichester, UK.
- Pendyala, R. M., R. Kitamura, C. Chen and E. I. Pas (1997) An activity-based microsimulation analysis of transportation control measures, **Transport Policy**, **4**, 183-192.
- Resource Systems Group, Inc. (1997) **Route 53 Alternatives Study: Lake County Model Description; Analysis of 2020 Lake County Traffic Levels With and Without the Route 53 Tollway** reports to Environmental Law and Policy Center, Chicago.
- Rickert, M. and K. Nagel (1997) Experiences with a Simplified Microsimulation for the Dallas/Fort Worth Area, Las Alamos National Laboratory, Los Alamos, NM.
- Smith, L. L. (1996) Transims Travelogue, Las Alamos National Laboratory, Los Alamos, NM. <http://www-transims.tsasa.lanl.gov/>

Table 1: Summary Description of Travel Forecasting Models

Travel Forecasting Model	Model Characteristics					Computing Platform
	Choice Variables	Exogenous Variables	Endogenous Variables	Parameter Estimation		
Four-Step Sequential Procedure (trip-based)	trips: number by purpose, O/D, mode, route	networks: highway, transit activity locations	link flows; link times/costs given choices; solved with "feedback"	survey data; studies external to the model	mainly PCs; commercial software systems	
Activity Analysis (SAMS) (activity-based)	activities: type, location, duration trips: time-of-day, O/D, mode	networks: highway, transit activity locations	link flows; link times/costs given choices; solved with "feedback"	survey data; studies external to the model	not described	
TRANSIMS (routing module) (activity-based)	trips: routes by 15 minute time periods	networks: highway	link flows; link times solved with "feedback"	not described	networked UNIX workstations	
Combined Network Equilibrium Model (equilibrium-based)	trips: O/D, mode, route, time period (for given Os and Ds by purpose)	networks: highway, transit activity locations	link flows; link times/costs in equilibrium with choices	survey of travel choices; consistent with predicted network times/costs	PCs and UNIX workstations; EMME/2 macros and research codes	