Convergence of an Activity-Based Travel Model System to Equilibrium: Experimental Designs and Findings

John Gibb, DKS Associates, 8950 Cal Center Drive, Suite 340, Sacramento, CA 95826.
530-878-6991, jag@dkssacramento.com

John L. Bowman, Ph. D. Transportation Systems and Decision Sciences, 28 Beals Street, Brookline, MA 02446. 617-232-8189, John_L_Bowman@alum.mit.edu

Presented at the 11th TRB National Transportation Planning Applications Conference, Daytona Beach, Florida, May 6-10, 2007. This edition is slightly expanded from the submittal to the conference proceedings.

Abstract

This report describes some variations on the "method of successive averages" to efficiently converge to equilibrium a travel demand model system built around an activity-based model, and experimental results for the convergence and efficiency of a test case. (The activity-based travel model itself, developed by Bradley and Bowman, is not detailed in this report.) The application system includes the activity-based model plus auxiliary trip-based models for external, airport, and commercial vehicle trips, along with the processes that provide network level-of-service matrices to the activity and auxiliary models, and assign trips from those models onto networks. Experiments upon test cases indicate: (a) The system can be converged close to equilibrium by running the activity-based model upon small samples of the population during early iterations, progressing to larger samples in later iterations, so there are much fewer total activity simulations per person than the number of system iterations; (b) a constant step size of one-half converged more rapidly than the customary 1/(iteration number) step size; (c) application of the step size using a common "preloading" option of assignment is found to require significantly fewer assignment iterations than customary application upon the demand trips or the volumes. Various measures of system convergence are examined, including measures of travel time stability between system iterations.

Introduction

This paper reports on developments and experimental results of equilibration of demand and assignment in a regional activity-based travel forecasting system "SacSim" developed in 2005-2006 for the Sacramento Area Council of Governments (SACOG). The activity-based model at the heart of this system, "DaySim" is a disaggregate microsimulation of personal activities and travel (Bowman and Bradley 2006).

Following a brief overview of DaySim, this paper identifies the equilibrium problem of its application. Computational problems and opportunities are then discussed and some approaches for efficient equilibration are introduced. For two prototypical iteration schedules applied by this approach, computation speed and equilibrium convergence are compared.

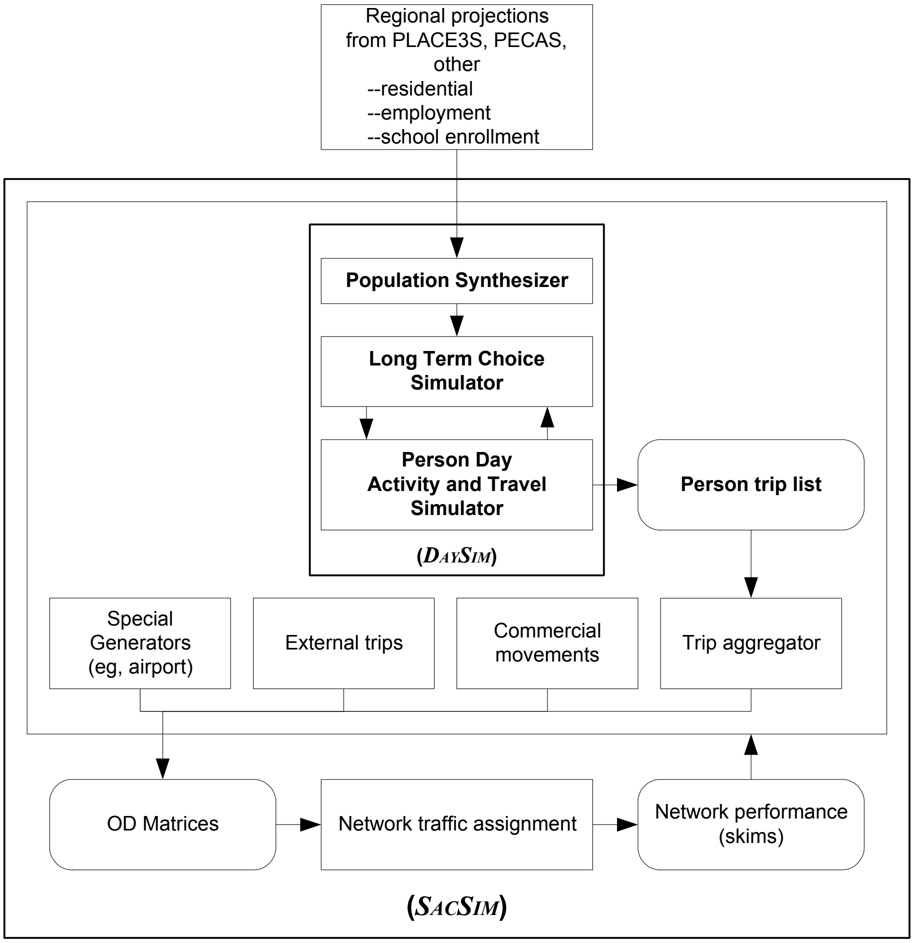
Activity-Based Model Overview

Figure 1 shows the major SacSim components. The activity-based model application, DaySim, consists of (1) a population synthesizer to create a synthetic population organized in households (normally only once at the beginning of a model process run, or in advance) from U.S. Census Public Use Microdata Sample, and zonal and parcel data, (2) a long-term choices model simulating work and school location and auto ownership for the synthetic population, and (3) the person day activity and travel simulator, simulating a one-day diary and travel schedule for each person in the population. These components are implemented in a single software program as an integrated hierarchy of multinomial and nested logit models.

SacSim consists of DaySim, plus network travel cost ("skim") measurements (needed as input to DaySim), trip-based models predicting special generator, external, and commercial vehicle travel demands, and network assignment to compile and load the trips onto the network. Network assignments and travel cost measurements are performed for four periods of the day: AM peak 3-hour period, mid-day, PM peak 3-hour period, and evening. Vehicle traffic assignment uses the Frank-Wolfe fixed-demand equilibrium assignment process, with its own internal iterations. All components of SacSim outside of DaySim are implemented in TP+/Cube Voyager software.

Due to the vast multidimensional choice sets in its underlying models, it would be impractical to enumerate and predict probabilities of each possible choice, as is normally done with trip-based models. DaySim works around this using Monte Carlo microsimulation to generate a single outcome of each modeled choice, expressed as an activity and travel diary of each (synthetic) person for one day. (In this report, "simulation" refers to DaySim's application of its choice models to synthesize these diaries, not anything like a second-by-second process model or virtual-reality presentation of each person and vehicle.)

Figure 1
SacSim Regional Travel Forecasting System



Equilibrium Problem

Figure 1 indicates the cyclical relationship between network performance and travel demand, that network performance both determines and results from travel demand. The model system is in equilibrium when the network performance used as input to DaySim and the other trip models matches the network performance resulting from assignment of the resulting trips. Trip-based travel demand models, in theory if not always in practice, have the same requirement.

Naïve direct feedback has proven to be inefficient and unreliable at converging them to a consistent solution. There is no reason to expect activity-based models like DaySim to equilibrate any better this way.

Almost all practical travel models that converge toward equilibrium use some form of the Method of Successive Averages (MSA), also called convex combinations. During any iteration *i*, MSA estimates an updated approximate solution vector of travel $\mathbf{x}_i$ (zone-to-zone demand and/or link volumes) as a weighted average of the previous estimation, $\mathbf{x}_{i-1}$, and the travel from the current iteration of the demand models, $\mathbf{y}_i$. Representing the weight for the new iteration by λ_i (which must satisfy $0 < \lambda \le 1$), MSA calculates $\mathbf{x}_i = (1-\lambda_i)\mathbf{x}_{i-1} + \lambda_i \mathbf{y}_i$. The weight λ_i is commonly called the iteration's "step size." (In the first iteration, there is no $\mathbf{x}_{i-1}$, so λ_1 must be 1, and $\mathbf{x}_1 = \mathbf{y}_1$.) The first iteration typically uses network performance skim matrices from free-flow link times.

Some trip-based model systems are known in which each iteration's step size can be chosen to optimize an objective function that is maximum at equilibrium (Evans 1976, Florian et al 1975). But in activity-based models like DaySim, as in most trip-based models in practice, an objective function is not available or economical to calculate, so the step sizes must be predetermined.

The classic predetermined step size schedule is $\lambda_i = 1/i$. Its effect is that after any iteration *i*, the solution approximation is the average of all the iteration-result vectors computed so far: $\mathbf{x}_i = \frac{\mathbf{y}_1 + \mathbf{y}_2 + \ldots + \mathbf{y}_i}{i}$. (Sometimes this step size formula in particular is called the Method of Successive Averages, but this paper uses that term for any step size.) This step size schedule is known to reliably converge a wide range of models, but experience has often shown it to converge slowly.

Some trip-based models converge reliably and more efficiently with a fixed step size (Bar Gera and Boyce, 2006). Care must be taken in the choice of that step size, which depends on the problem. A step size of one-half has been successful in many models in practice.

With any iterative algorithm to solve for equilibrium must be some measures of convergence progress, to either decide when to stop iterating, or at least assess how well it has converged after a given iteration, and to verify that the system is actually converging. Some measures are:

1) Trip-based, such as total misplaced trips (Bar-Gera and Boyce 2006). This is the sum, for all O-D pairs, of the absolute values of the differences in travel demands from one iteration to the next. Variations include root-mean-squared differences.
2) Travel-time-based, such as the maximum absolute or root-mean-squared change in zone-to-zone travel time.
3) Link-based, such as the maximum, among all links, of the absolute volume change from one iteration to the next.

Random "Noise" Problem

Further complicating equilibrium convergence is that Monte Carlo microsimulation creates random "noise" that does not occur in trip-based models (when calculated at sufficient precision).

To quantify the noise from the microsimulation, consider some selected aggregation of demand, say, trips from one certain zone to another. Let D be its "true" number of trips, due to the underlying discrete choice models (enumerating all choices and accumulating each one's probability). Running the Monte Carlo process an infinite number of times with different random seeds, this aggregation will average to D, but in a particular run it may deviate significantly from D. Assuming each trip is almost an independent event in a Poisson or binomial process, the variance of trips in our selected aggregation, across several runs, will approximate D as well, so its standard deviation would be $\sqrt{D}$. Certain aggregations exhibit greater or lesser measured variances due to interdependencies between trips.

Effects of random variation upon model forecasts have been studied with a model similar to DaySim (Castiglione, Freedman, Bradley 2003). A major consideration at the outset of this study is to what extent random noise frustrates equilibrium convergence. It can be seen that total misplaced trips, from one iteration to the next, is overwhelmed by random noise, which is inevitably a large portion of the total demand.

Computational Problems and Opportunities

The computational run times of DaySim are quite long. For this study's experiments DaySim required approximately 5 hours on a 3 GHz processor to perform a full simulation of day activities for 1.5 million persons. Run times are governed mainly by the number of persons and households simulated, rather than the number of zones and iterations as in trip-based models. So to permit quick test and approximation runs, an option was built into DaySim to simulate activities for a user-specified fractional sample $1/s$ of the households, instead of all of them. This study explores this option as an opportunity to run the model system itself more quickly at low sampling rates, at least in early iterations, despite the increased random noise.

If we have a population of N households, but run DaySim at a sampling rate $1/s$, so as to draw an arbitrary sample of N/s households for simulation, then we scale the resulting demands up to full size by a factor of s, so that the same selected aggregation of demand still has expected value D. But since D trips arise now from D/s independent random events, the standard deviation of D from such runs would be $s\sqrt{\frac{D}{s}}$, which is $\sqrt{Ds}$. Thus the noise of the sampled application increases by a factor of $\sqrt{s}$, that is, $\sqrt{\frac{\text{total population}}{\text{sampled population}}}$, compared to a "whole" model.

Conversely, one may reduce the noise of a forecast by running DaySim several times upon the whole population, each with different random-number seeds, and averaging the demands.

Coordination of Step Sizes and Sampling Rates

Since both the step size and the sampling rate for simulation for each iteration must be pre-scheduled, the relationship between the two should be considered.

The weighted average of trips D_1 at sampling rate s_1, and trips D_2 at sampling rate s_2, has theoretical variance $(1-\lambda)^2 s_1 D_1 + \lambda^2 s_2 D_2$. The implication is that the random noise of a weighted average can be less than the random noise of the individual samples. Predicted random noise is minimized by choosing $\lambda = \frac{\frac{D_2}{s_2}}{\frac{D_1}{s_1} + \frac{D_2}{s_2}} \approx \frac{\frac{1}{s_2}}{\frac{1}{s_1} + \frac{1}{s_2}}$. At any iteration *i*, this equates to:

$$\lambda_i = \frac{\text{Households sampled in Iteration } \boldsymbol{i}}{\text{Total households sampled in Iterations 1 through } \boldsymbol{i}} .$$

(Providing this relationship between sampling and step sizes neither assures or optimizes convergence to equilibrium, nor is this necessary to efficiently converge.)

Another consideration is whether to draw samples with or without replacement. In this study's experiments, household samples for simulation were drawn without replacement across a number of iterations, until the population was exhausted. This way, the households in simulations being "pooled" among iterations are all different households; none are different simulations of the same household, until all households have been simulated and included equally. Space permits only to explain that this arrangement puts the law of large numbers into effect upon the largest number of households available. (Further study may examine sampling strategies that draw with deliberate replacement, to take advantage of DaySim's variance reduction by random seed retention.)

Table 1 shows an example of a simple iteration schedule of the 1/*i* step size type (customary MSA). Each member of the population generates a simulation once sometime within its five iterations, and each one yet simulated is weighted equally at any iteration.

Table 1
Example Iteration Schedule

Iteration Num.	Sampling Rate	Offset	Households Sampled	Step Size	Scale
1	1 in 5	1	1, 6, 11, 16...	1	5
2	1 in 5	2	2, 7, 12, 17...	1/2	2.5
3	1 in 5	3	3, 8, 13, 18...	1/3	1.67
4	1 in 5	4	4, 9, 14, 19...	1/4	1.25
5	1 in 5	5	5, 10, 15, 20...	1/5	1

Experimental Demonstration Cases

Early test models of the SacSim model system used the customary declining step-size MSA, with various numbers of iterations and sampling rates. Most aggregate statistics (VHT, VMT, total trips) of a base-year model tended to settle to the "neighborhood" of equilibrium in 3 to 5 iterations, but further progress was slow, warranting 15 to 20 iterations to converge substantially better. Low simulation sampling rates did not appear to hinder reasonable convergence toward equilibrium.

Those early experiments suggested improvements that form the basis of the two prototype models examined in this report. These are:

Demonstration Case 1: Staged declining step size of 8 equal-sample iterations, starting not with free-flow link times but with link times resulting from 4 "warm-up" iterations at a low sampling rate. A final iteration simulates all households to produce a single consistent database, as shown in Table 2;

Demonstration Case 2: Constant step size, progressively increasing sampling rate, culminating in an iteration with a full simulation of all households, as shown in Table 3.

Table 2
Iteration Schedule for Demonstration Case 1, Staged Declining Step Size

Iteration Num.	Sampling Rate	Offset	Households Sampled	Step Size	Scale
1	1 in 30	1	1, 31, 61…	1	30
2	1 in 30	2	2, 32, 62…	1/2	15
3	1 in 30	3	3, 33, 63…	1/3	10
4	1 in 30	4	4, 34, 64…	1/4	7.5
5	1 in 8	1	1, 9, 17…	1	8
6	1 in 8	2	2, 10, 18…	1/2	4
7	1 in 8	3	3, 11, 19…	1/3	2.67
8	1 in 8	4	4, 12, 20…	1/4	2
9	1 in 8	5	5, 13, 21…	1/5	1.6
10	1 in 8	6	6, 14, 22…	1/6	1.33
11	1 in 8	7	7, 15, 23…	1/7	1.14
12	1 in 8	8	8, 16, 24…	1/8	1
13	1 in 1	1	1, 2, 3, 4…	1	1

Table 3
Iteration Schedule for Demonstration Case 2, Constant Step Size

Iteration Num.	Sampling Rate	Offset	Households Sampled	Step Size	Scale
1	1 in 128	128	128,256,384,512...	1	128
2	1 in 128	64	64, 192,320,448...	1/2	64
3	1 in 64	32	32, 96, 160, 224...	1/2	32
4	1 in 32	16	16, 48, 80, 112...	1/2	16
5	1 in 16	8	8, 24, 40, 56...	1/2	8
6	1 in 8	4	4, 12, 20, 28...	1/2	4
7	1 in 4	2	2, 6, 10, 14...	1/2	2
8	1 in 2	1	1, 3, 5, 7…	1/2	1
9	1 in 1	1	1, 2, 3, 4…	1/2	1/2

Convergence Progress

A first look at a relationship between noise and convergence examines 10 runs of a model system of the Demonstration Case 1 schedule, differing only by the random seed in the simulations. Figure 2 shows the progress of individual-iteration vehicle trips in the PM period (the most congested and volatile period) across iterations. “Error bars” show their average plus-or-minus two theoretical standard deviations at the iteration’s sampling rate. Following a predictable decrease from the first to the second iteration, individual runs tend to exhibit more random fluctuation than a specific trend toward equilibrium. Final totals, however, are within the range expected from full-population simulation.

Figure 2
Progress of PM Iteration Vehicle Trips in 10 Staged-MSA Runs

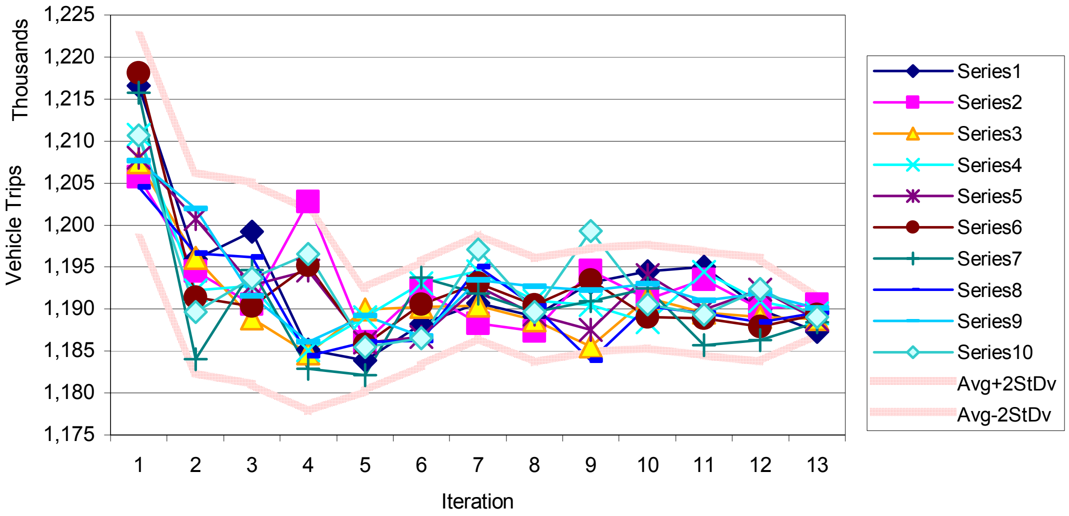
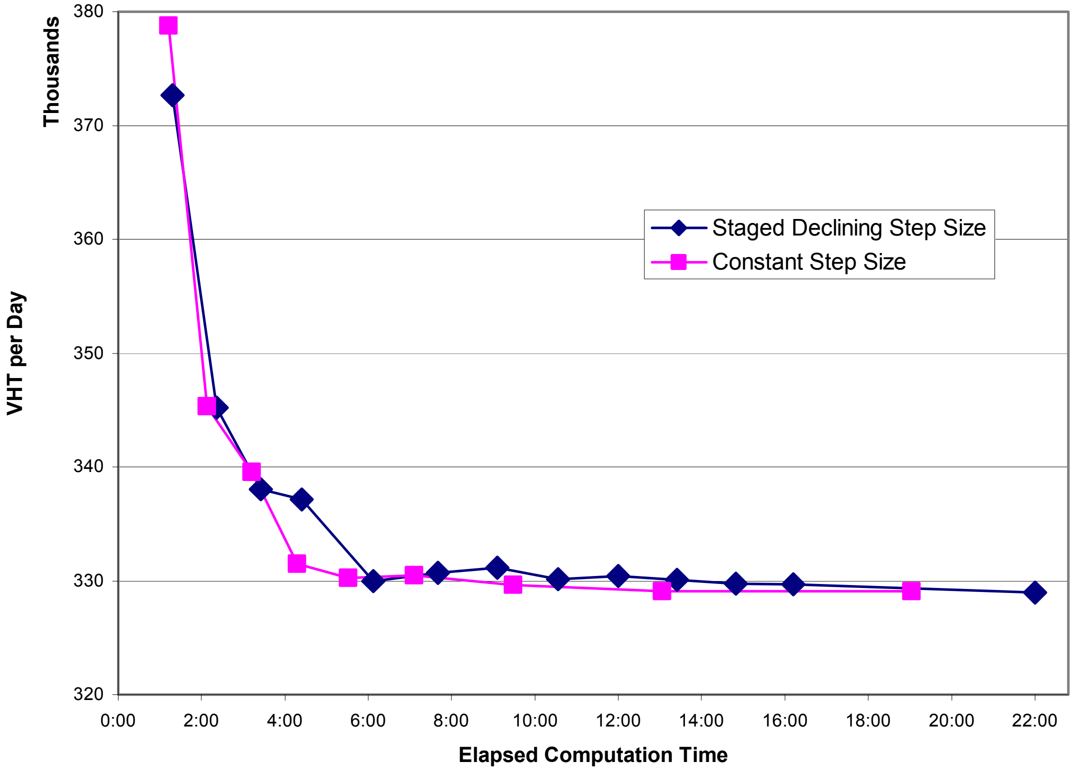


Figure 3 shows convergence of vehicle-hours traveled (VHT) for the PM peak period, for one run of each of the two demonstration cases. These result from successively averaged traffic assignments. After the predictable rapid drop from iteration 1 to iteration 2, vehicle-hours change more slowly afterwards, due to convergence in demand, and particularly in Demonstration Case 1, declining step sizes. (Its jump in start-over iteration 5 toward near-equilibrium after a slow-down is evidence that the step sizes get too small too soon, also seen in the early simple-MSA runs.) Both methods progress toward equilibrium values despite the random noise exacerbated by simulation sampling, especially in the constant step size case’s early iterations.

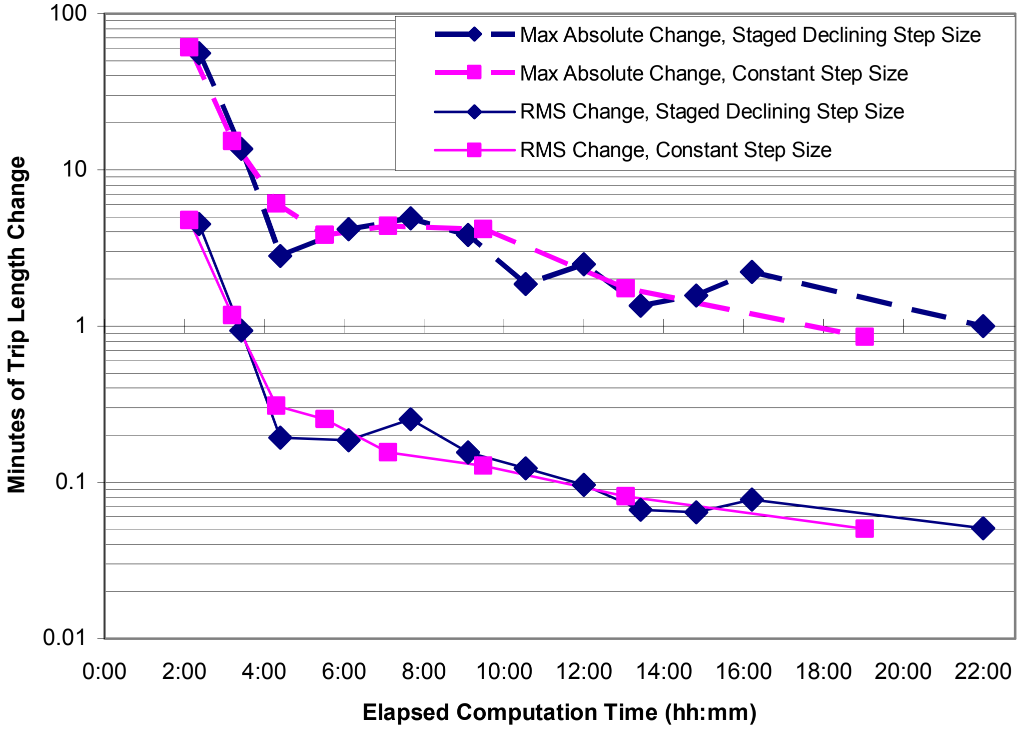
Convergence progress in Figure 3 is plotted against elapsed computation time rather than number of iterations. Both demonstration cases show (1) a roughly one-hour fixed time for non-DaySim computations per iteration, (2) a 5-hour DaySim runtime upon the full population, proportionately less at lower sampling rates, and (3) a faster overall runtime for the constant step size method among the two demonstration cases.

Figure 3
PM VHT Comparison



Convergence of zone-to-zone travel time was also examined. The first measure is the largest absolute change of skimmed travel time for O-D pairs having at least one trip, from the preceding iteration; the second is its trip-weighted root-mean-square (RMS). Figure 4 compares the progress of these measures for the two demonstration cases. Both methods achieved comparable precision of nearly 1 minute maximum and 3 seconds RMS, though the constant step size method did so with less computation time. This method's travel times also fluctuated less than the declining step size's in later iterations, probably because of its higher sampling rate in those iterations. This hints that random demand perturbations may still have a detectable effect upon travel times.

Figure 4
Zone-to-Zone Auto Trip Length Change Statistics,
all periods, relative to preceding iteration



An Efficient Successive Averaging Method

Whether assigning successively-averaged trip demands, or assigning current-iteration demands for later averaging, Frank-Wolfe equilibrium assignment to a good equilibrium “gap” tolerance is a major computational effort in practical model systems with “feedback,” since it is slow, must be repeated many times, and cannot be practically sped up with the help of previous iterations’ assignments.

An alternative successive-averaging method during assignment that we think is novel was found to reduce run-times considerably. This method uses a standard fixed-demand Frank-Wolfe equilibrium assignment program to assign the step-size fraction of the new demand, with the complementary proportion of the previous system-iteration’s final volumes treated as preloaded volume. Link time calculations always use the combined volume from preloading and assignment. Specifically:

Preload volume = $(1-\lambda_i)V_{i-1}$
Demand assigned = λ_i(iteration i demand trips)

where λ_i is the predetermined step size for system-iteration i (not of the internal iterations of assignment).

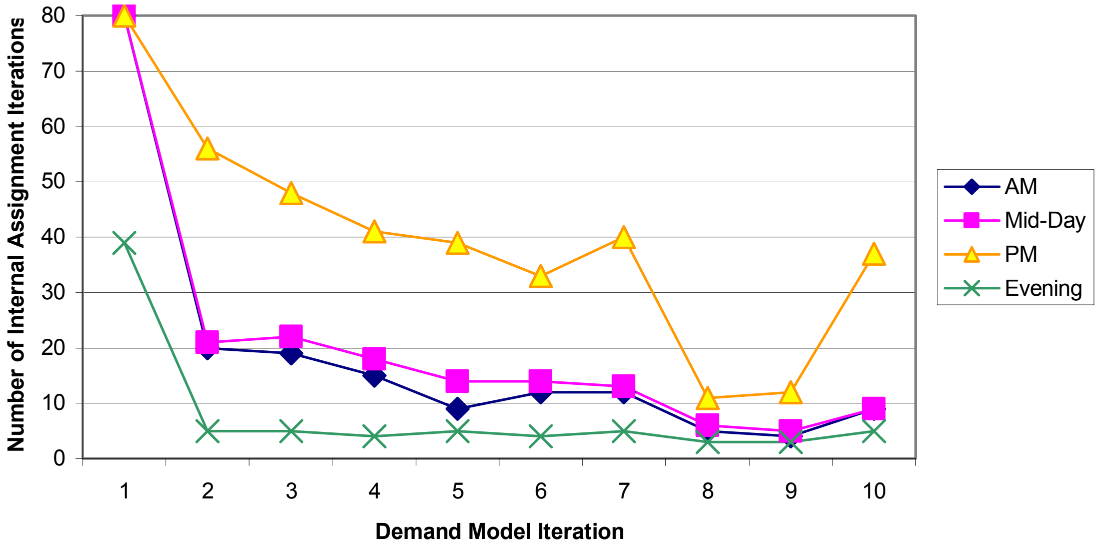
This method is comparable to the procedure of Evans (1976) and EMME/2 "Elastic Demand Assignment" (INRO Consultants, 1998). Their procedures assign each iteration's trips in an all-or-nothing assignment on the same paths used to derive the skims, and average these assigned volumes with the previous successively-averaged volumes. These models choose an optimal step size according to a combined objective function of trip distribution and assignment. The method studied here adds an "inner loop" of assignment iterations, allowing the trips being assigned to shift routes until acceptably equilibrated. This assignment chooses optimal step sizes for its "inner loop" iterations as normal for Frank-Wolfe assignment; these are distinct and separate from the predetermined step sizes of the "outer loop" iterations of the whole model system which determine how much traffic to preload and how much to assign.

Several tests indicate this method yields assignments that compare reasonably close to assignments of successively-averaged trip matrices, but converge with far fewer assignment iterations than needed for an ordinary "cold" assignment of demand. Comparisons between this method and assignments of successively averaged trip matrices were found to depend on the relative gap stopping criterion. When the gap criterion was small (tight) and properly defined, the assignments of the two methods compared quite close, but with lenient gap, the two methods differed more.

Figure 6 shows the number of internal iterations of assignment required to reach a fairly stringent relative gap for each of the four time periods in SacSim, in each iteration of a model similar to (but not exactly) demonstration case 2. The number of iterations required appears to be related to factors such as:

1) Congestion. In this model the PM period is the most congested, and the evening the least, and these periods consistently need the most and the least iterations, respectively.
2) Step size: The step size is constantly one-half, except for iterations 8 and 9 in which it is lower (and iteration 1 of course). The numbers of iterations increase or decrease with the step size.
3) Convergence of demands: As the demands become closer to constant from one iteration to the next, fewer iterations are usually required for the same step size.

Figure 6
Internal Iterations of Assignment Required
Stopping criteria are relative gap < 0.0002 or 80 iterations



It must be admitted that this assignment method is somewhat a stopgap. Emerging path-based and origin-based assignments (Bar-Gera 1999) in theory are simple to start "warm" in the modern feedback context of both trip-based and activity-based models.

Conclusions

An activity-based model system with Monte Carlo simulation was shown to converge successfully toward equilibrium much faster than would otherwise be possible, due to a method of sampling households for simulation, and with an apparently novel method of application of an otherwise conventional assignment procedure. The greater random noise due to sampled simulations did not hamper convergence. A strategy to minimize noise for the given sampling was identified. A constant step size of one-half was effective and efficient, coordinated with a growing sample for each succeeding iteration. Less efficient was a declining step size schedule, which remains a fallback for difficult-to-converge cases.

This study stopped short of creating a black box one would iterate indefinitely until a convergence criterion is met for the model system. The choice of stopping criterion depends on the problem, and must be considered due to the long running times. Studies of regional system performance might focus on travel times, which converge relatively quickly. Estimates of the number of trips in a specified location or O-D pattern may require numerous iterations with full-population simulations to overcome randomness. Link volume forecasts, while not examined in this report, should not require so many. In any case, actual planning analyses are expected to require additional iterations with full-population simulations after starting with either of the prototypical schedules.

Further experimental and theoretical study is needed to determine the interactions of equilibrium precision and trip demand precision. That is, what kind and amount of equilibrium tolerance must be achieved to ensure a trip demand quantity's random error is sufficiently close to what is expected at its sampling rate, and what amount of equilibrium tolerance can be achieved at a given sampling rate?

References

Bowman, J. L., Bradley, M. A. (2006) Activity-Based Travel Forecasting Model for SACOG: Technical Memos 1 through 11, available on http://jbowman.net.

Evans, S. (1976) Derivation and Analysis of Some Models Combining Trip Distribution and Assignment, *Transportation Research Part B* 10, 37-57.

Florian, M., S. Nguyen and J. Ferland. (1975). On the Combined Distribution-Assignment of Traffic, *Transportation Science* 9, 43-53.

Bar-Gera, H. and D. Boyce, (2006), Solving a non-convex combined travel forecasting model by the method of successive averages with constant step sizes. In *Transportation Research Part B,* 40 (2006) 351–367.

Castiglione, J., Freedman, J., Bradley, M. (2003) Systematic Investigation of Variability Due to Random Simulation Error in an Activity-Based Microsimulation Forecasting Model, *Transportation Research Record* 1831, 76-88, Transportation Research Board.

INRO Consultants, Inc., Montréal (1998), *EMME/2 Release 9,* pp. 6-17 to 6-21.

Bar-Gera, H. (1999) *Origin-Based Algorithms for Transportation Network Modeling*, National Institute of Statistical Sciences, Technical Report Number 103.