APPLICATION OF A COMBINED TRAVEL DEMAND AND MICROSIMULATION MODEL FOR A SMALL CITY

Daniel Morgan
Transportation Engineer
Caliper Corporation
1172 Beacon Street, Newton, MA 02461
Phone: (617) 527-4700
Fax: (617) 527-5113
daniel@caliper.com

Rick Mayberry
Transportation Engineer
California Department of Transportation
1656 Union Street, Eureka, CA 95502
Phone: (707) 445-6634
Fax: (707) 499-4594
rick_mayberry@dot.ca.gov

This paper describes the development and application of a multi-modal microsimulation model for the Greater Eureka Area (GEA), a small community of 25,000 in Northern California. The travel demand model for the GEA was used to develop estimates of the traffic demand for base and forecast years. A rigorous data collection and calibration effort was made to calibrate the simulation model for the base year. The simulation of pedestrian activity and bus routes is included in the model.

The application is unique in its methods and in its ultimate objective. First, the microsimulation model was developed on a geographic information system platform shared with the travel demand model, allowing the fusion of geographic information and the application of geographic analysis methods to assist in the refinement of peak period trip tables for simulation. The travel demand model was used to develop initial estimates of the traffic demand. Additional analysis was performed to develop a dynamic temporal profile in the demand. Simulation-based dynamic traffic assignment methods were used to calibrate route choices in the model.

Second, the microsimulation model was designed not for a specific and finite project in the common tradition of planning and engineering practice but for the purpose of becoming a living model to be adopted and maintained by local authorities for use in all manner of planning and traffic impact studies, both big and small, throughout the city. Thus, the microsimulation model will serve as a natural corollary and complement to the travel demand model. A variety of alternatives, including the additional lanes on a key corridor and traffic signal optimization, are analyzed to demonstrate the effectiveness of microsimulation for improving the estimates of project impacts in the planning context.

In addition to serving as an illustrative case study for the application of microsimulation in small and medium-sized communities, this paper demonstrates the advantages of GIS for making the development of a microsimulation model for small and medium-sized communities feasible and cost-effective. Lessons learned and guidance for similar applications elsewhere are provided.

Submittal is for a presentation and written paper.

Description of paper: A case study demonstrating the development, calibration, and application of a microsimulation model in a small community, including travel demand forecasting, trip matrix estimation, and dynamic traffic assignment techniques.

Key words: traffic simulation, microsimulation, travel demand model, dynamic traffic assignment

Student status: Neither of the authors is a student

APPLICATION OF A COMBINED TRAVEL DEMAND AND MICROSIMULATION MODEL FOR A SMALL CITY

INTRODUCTION

With the rising profile of mesoscopic and microscopic traffic simulation models and higher fidelity, notably dynamic, traffic assignment methods in the planning community, there is little in the way of guidance or standards of practice when it comes to estimating, calibrating, or applying these models. Even as the experts struggle to define dynamic traffic assignment (DTA), there is little experience for travel demand modelers to draw upon. This paper describes a methodology and case study for a successful application of a combined travel demand and microsimulation model for a small city. More specifically, this paper will focus on the microscopic simulation-based DTA methods that were used in the estimation and calibration of the model.

The microsimulation model spans the whole of Eureka, California, which represents the core of the Greater Eureka Area (GEA) Travel Model, the travel demand model that covers all of Humboldt County. The purpose of the project is to develop a traffic microsimulation model that extends and compliments regional planning activities dependent, for the most part, on the Greater Eureka Area Travel Model, which, like most travel demand models, lacks the detail and fidelity to be used to evaluate projects or to estimate the operational impacts of growth, changes in land use, or roadway and traffic signal improvements. To a lesser extent, modeling activities involve other traffic simulation and traffic signal timing analysis software, but none that bear any formal relationship with the travel demand model or with one another. Given that the Humboldt County Association of Governments (HCAOG) has invested over a period of years, and will continue to invest, the energies and resources of its member governments to develop and maintain the travel demand model, a microsimulation model that builds on and complements the travel demand model was also a motivating factor.

The design of the microsimulation model is such that it will rely on the travel demand model to produce base estimates of the traffic demand for peak periods, but use field data to improve those estimates by way of industry-standard calibration and validation techniques. The geographic scope of the microsimulation model is far greater than that of a typical traffic simulation project, including every street in a relatively dense grid network covering more than 16 square miles. To contend with the trip table estimation and route choice challenges that are pivotal to the successful calibration and validation of such a model, local geographic information, including information about the predominant land uses in each traffic analysis zone, is used.

Once estimates of the trip tables for the AM and PM peak periods are produced using the travel demand model, field data, and geographic information, the most formidable technical barrier to simulating areas as wide as even a small city is the question of route choice. The simulation-based DTA model that is the subject of this paper is the principle tool used to decide the routes that drivers will take in the GEA microsimulation model. The DTA model is based on principles that are familiar to planners, such as User Equilibrium (UE), but uses route choice methods to determine the

paths to which vehicles are assigned and microscopic traffic simulation models (e.g., car following, lane changing) to determine link performance (i.e., loaded travel times and delays).

PROBLEM STATEMENT

The methodology in this project is designed to answer two questions, the answers to which are inextricably linked and interdependent:

(1) What are the volumes of vehicles travelling between origin and destination zones in the network?
(2) What are the likely routes drivers take between those origin and destination zones?

The challenges in answering these questions effectively and accurately stem mostly from limitations in existing methodologies and in the data that is typically used to answer them. To truly understand the trip pattern and route choices of drivers in a region, it is imperative to directly observe the origins, destinations, and routes. This can be achieved with license plate surveys, for example. Other innovative methods have been used that track or match the identities of vehicles observed at different locations at different times, such as video recorded from airplanes circulating above a site.

TRAFFIC DATA

Traffic data used to estimate trip tables are commonly limited to what can be inexpensively collected or has already been collected, namely traffic counts. This dependence almost exclusively on counts has a number of drawbacks:

(1) Counts reveal neither the origins and destinations of vehicles nor their routes.
(2) Poor coverage of the study area may leave links on key routes between origin-destination pairs countless, which degrades the quality of the origin-destination matrix estimation (ODME) solution.
(3) Analysts are tempted to combine counts from different days, or even years, to increase coverage, or to average counts together to reflect an "average" day.

These problems are sources of error and uncertainty in the trip table solution, and tend to be, at worst, poorly understood or overlooked and, at the very least, underappreciated.

ODME METHODS

Furthermore, ODME methods are imperfect. ODME methods use traditional, typically static (i.e., one time period) traffic assignments to load trips from a matrix onto a network. Loaded flows are compared with counted volumes in order to calculate an adjustment to the matrix that, when loaded again, will improve the match between assigned flows and counts. This procedure continues iteratively until the match between the flows loaded from the estimated matrix and the counts cannot be improved further. Two shortcomings with these methods are, like the shortcomings in the traffic count data described above, also not well or widely understood:

(1) The solution is heavily influenced by the matrix used in the initial loading (i.e., the *seed* matrix).

(2) Volumes in a cell in the matrix are adjusted based on the flows and counts on links on the used path(s) between the corresponding origin-destination (OD) pair; crude heuristics must be used to estimate volumes between OD pairs between which no counts are available on the path(s) between them.

(3) No unique solution can be proven to exist, meaning that any number of estimated matrices might match the counts equally well when assigned to the network. In other words, a good match with the counts does not in and of itself prove a good estimate of the trip pattern.

The seed matrix that is of such critical importance to the quality of the ODME solution is usually produced by a subarea analysis in a travel demand model. Thus, a poorly calibrated travel demand model can be, alongside the traffic count data, yet another source of error in the ODME solution.

In summary, effective use of state-of-the-practice ODME methods requires a thoughtful consideration of these limitations. Experience applying the methods and the know-how to identify defects in the solution are almost a precondition for success.

ROUTE CHOICE CONSISTENCY

Lastly, once a matrix is estimated, assuming that it can be trusted to be representative of the real pattern of travel in the study area, traffic simulation models use route choice models to determine the paths vehicles choose to take. Routes determined by traditional static UE assignments in travel demand models, a tempting source for the routes to be used a microsimulation model, are neither unique nor reliable, and are in all likelihood probably infeasible. More recent developments in traffic assignment methods in the literature (Slavin *et al.*, 2009; Caliper Corporation, 2010) might one day change this fact. However, there is not yet any proof or reason to believe that the routes reported by a static UE assignment of any kind will yield a sensible set of paths for microsimulation.

The more useful route choice models in the microsimulation context are behavioral, meaning that drivers choose a route that minimizes some perceived total travel time or generalized cost or that achieves some other objective. Each driver makes his or her decision out of self interest, independently of the decisions of others. The loading mechanism (e.g., user equilibrium) used in the ODME methods are thus inherently inconsistent with the route choice loading. In other words, even if the match between the loaded flows and counted volumes resulting from the ODME is very good, there is no guarantee that the route choice model will produce the same loading. The match between the simulated flows and the counts might possibly be "worse" than that achieved by the ODME. Only a simulation-based ODME, using a route choice loading consistent with the simulation, can overcome this inherent inconsistency.

The methodology presented in this paper recognizes these shortcomings and proposes alternative methods to help improve the estimation and calibration.

General Methodology

The methodology used to estimate and calibrate a dynamic trip table for microsimulation of the GEA is summarized in Figure 1.
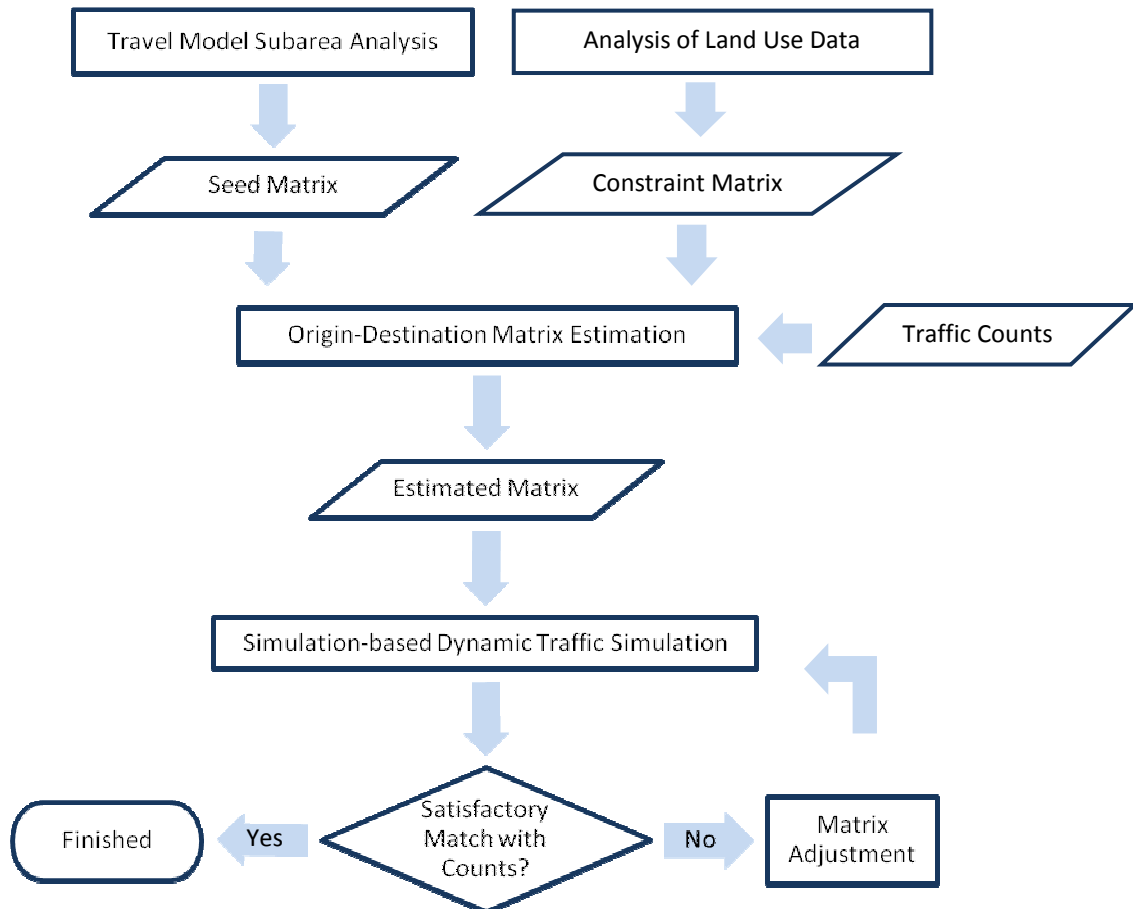
Figure 1. A flow chart illustrating the trip table and route choice estimation methodology

Subarea Analysis

The first step in the methodology is to perform a subarea analysis in the travel demand model. The core of the GEA study area is approximatly 3.5 miles by 3.5 miles, spans the city limits, and covers parts of surrounding Humboldt County. Stretches of Route 101 several miles beyond that 3.5 x 3.5-mile core are included in the subarea. The subarea is performed in the most recently calibrated base year 2005.

The subarea analysis is a traditional static UE traffic assignment for the AM and PM peak hours. It is run to a relative gap of 10^{-6}, beyond which only very minute changes in the flow vector occur over continued iterations, where the relative gap is computed as:

$$Gap \; = \frac{\displaystyle\sum_{i \in I} \sum_{k \in K_i} f_k \, t_k - \sum_{i \in I} d_i \, t_{\min,i}}{\displaystyle\sum_{i \in I} d_i \, t_{\min,i}}$$

where:

Gap	=	Relative gap;
I	=	Set of all O-D pairs i;
K_i	=	Set of paths used by trips traveling between O-D pair i;
f_k	=	Number of trips taking path k (i.e. path flow);
t_k	=	Travel time on path k;
d_i	=	Demand departing;
$t_{\min,i}$	=	Travel time on shortest path between O-D pair i;

The output of this procedure is the matrix that is used as the seed for the ODME.

ORIGIN-DESTINATION MATRIX (TRIP TABLE) ESTIMATION

Traffic count data was collected throughout the city in the Spring of 2009, the base year for the microsimulation model calibration. So that all counts could be observed during the same hours of the same days for purposes of consistency, cameras were used to record intersections so that turning movement volumes could be counted during back-office data reduction. Elsewhere, manual turning movement counts and pneumatic tubes were used to collect traffic counts. These count data were processed and compiled into a geographic count database in TransModeler, the microscopic traffic simulation platform used in the project.

The minimum requirements to perform the ODME are a seed matrix and traffic counts. However, it was found that, due to some of the issues with ODME methods described earlier, the ODME solutions based on these inputs alone are inadequate. Eureka has a fairly dense, mostly grid layout, which allows for numerous feasible paths between a significant majority of origin-destination pairs. Furthermore, the traffic analysis zones are fairly small, resulting in a rather dense geographic distribution of centroids (i.e, around 50 per square mile). The dense zonal structure and grid street network present stiff challenges for the ODME effort. For instance, the ODME solution may achieve excellent results in terms of matching counts by producing unreasonable volumes of trips traveling short distances between agricultural and low-density residential areas. Thus, to steer the ODME toward a more probable solution, geographic analysis was used to produce a matrix of constraints that would limit the volume of trips that can be produced in the ODME between zones a very short distance a part and having predominantly agricultural, low-density residential and other land uses not expected to be significant producers or attractors of trips.

Using the seed matrix from the subarea analysis, the traffic counts, and the constraint matrix, an ODME solution was produced for the AM and PM peak periods that matched counts with a root mean square error of less than 10% and that satisfied general a priori expectations about trip pattern in the city.

SIMULATION-BASED DYNAMIC TRAFFIC ASSIGNMENT

The network loading (i.e., the link flows resulting from the trip assignment) in the ODME method used in this application is based on a traditional static (i.e., single time period) UE traffic assignment. Thus, the ODME loading does not account for the capacity and delay effects of traffic signals, stop signs, and other causes of interrupted flow that are prevalent in Eureka, a city whose street network is predominantly, if not almost entirely, made up of urban surface arterials and local street. The ODME loading also allows volumes to exceed capacity. The simulation-based DTA model, on the other hand, is sensitive to all of the aforementioned effects and does not permit volumes to exceed capacity. In fact, the loading resulting from the simulation-based DTA ought to be different from, and will be inconsistent with, that resulting from the ODME, the basis on which the matrix of trips is estimated. This means that the step-wise approach that begins with static ODME and is followed by DTA, though it represents the best tool set the state of the practice has to offer, has deficiencies that should be recognized up front. That said, additional steps, which will be described later, can be taken to rectify some discrepancies between the two techniques.

In order for reasonable route choices to be simulated, congested (i.e., "loaded") travel times on which the route choices are based must be estimated. This is the primary function of the simulation-based DTA method in TransModeler. The simulation is run to completion for the entire time period iteratively, with the method of successive averages applied to output travel times every iteration. The route choices of each run are thus a function of the travel times simulated and averaged over prior runs. In Eureka, a 15-minute temporal profile in the demand was estimated based on 15-minute count data. Thus, dynamic, 15-minute travel times were used. Through the simulation-based DTA, those dynamic travel times (and the dynamic route choices) are expected to stabilize. The assignment runs until it has converged to a target relative gap measure for UE or until a maximum number of iterations is reached.

The relative gap calculation is similar to that of the subarea analysis, but a gap is computed both for the entire simulation period as well as for each time interval τ as follows:

$$Gap^{\tau} = \frac{\displaystyle\sum_{i \in I} \sum_{k \in K_i} f_k^{\tau} t_k^{\tau} - \sum_{i \in I} d_i^{\tau} t_{min,i}^{\tau}}{\displaystyle\sum_{i \in I} d_i^{\tau} t_{min,i}^{\tau}}$$

where:

Gap^{τ} = Relative gap in time interval τ;

I = Set of all O-D pairs i;

K_i = Set of paths used by trips traveling between O-D pair i;

f_k^{τ} = Number of trips taking path k (i.e. path flow) in time interval τ;

t_k^τ = Travel time on path k in time interval τ;

d_i^τ = Demand departing in time interval τ;

$t_{\min,i}^\tau$ = Travel time on shortest path between O-D pair i in time interval τ;

Very good are achieved using the DTA to estimate the route choices of trips generated from the estimated trip tables. Routes observed visually between OD pairs and passing through critical links all satisfy basic a priori expectations about the feasible routes drivers take. Unreasonable routes, such as those along corridors interrupted by stop signs every block, are effectively filtered out of the set of route choices. Furthermore, multiple route choice alternatives are frequently used between many OD pairs.

The ODME and DTA together produce very good results in Eureka given the complexity of the route choice problem in the city's grid street network and the imperfections inherent in their step-wise appliction. However, as expected, the quality of the goodness of fit with the counts, the ultimate target of the calibration effort, degrades as the simulated (DTA) loading diverges from the ODME loading. In Eureka, the root mean square error in the volumes increased from under 10% as determined by the ODME loading to as much as 35% (see Figures 2 and 3). However, for any given application, this will vary depending on the scale and complexity of the model. This necessitates the final step of the methodology, which is to systematically refine the trip table based on gaps that are observed in the model between simulated volumes and counts.

TRIP TABLE REFINEMENT

Refinement of the trip table is the last step in the methodology, and generally requires a number of iterations before a desirable degree of match between the simulated and observed volumes is reached. Using a critical link analysis based on the path flows resulting from the simulated route choices, the dominant OD pairs producing trips passing links where the match is poor are identified. The critical OD pairs in the trip table are thus identified by the software and scaled by some appropriate factor provided by the user to improve the match. If the changes are modest, another simulation-based dynamic assignment can be avoided. Rather, a single simulation can be run to determine the improvement in agreement with the counts. This can be repeated as many times as is necessary to improve the trip table and to reduce the error to acceptable calibration targets. This process effectively mirrors the automated ODME routine of alternating matrix adjustments and assignments, but is subjectively prioritized and driven by the modeler.

The manual refinement process has two important merits. First, the model volumes are a reflection of the route choices made in the simulation model, so there are no issues of consistency. Second, the process involves and indeed requires the interaction of the user, bringing into the process human judgement, intuition, and knowledge of the study area, all things the automated static ODME method lacks. However, the adjustment of the matrix should be done with thoughtful consideration of the effects scaling all the OD pairs identified via the "critical" link will have on other links in the network. Judgment must be used in order to leverage both the systematic and objective nature of the computerized ODME process and the injection of local knowledge about the general traffic pattern in the study area that is a virtue of the manual refinement.

The refinement of the trip table resulted in a substantial improvement in all goodness-of-fit measures that were used to compare the simulated link volumes with ground counts. The goodness-of-fit measures that were considered included the relative (i.e., percent) root mean square error (RMSE), absolute error, relative error, and the GEH statistic. The RMSE is a poor measure for goodness-of-fit in the wide area simulation context because it can exaggerate large relative errors on links with low counts – errors that are typically more acceptable to the practitioner than the same relative error on a link with high counted volumes, artifacts for which measures like the GEH statistic are designed to account. However, the relative RMSE gives a helpful single-value picture of the goodness-of-fit, at least when used to compare with other solutions. As shown in Figure 2, the relative RMSE deteriorated from 9% to 33% from the ODME to the DTA, then improved to 17% after the trip table refinement. Figure 3 shows the same for the PM case, where the relative RMSE increased from 10% to 35% after the DTA, then improved to 18% after the trip table refinement.
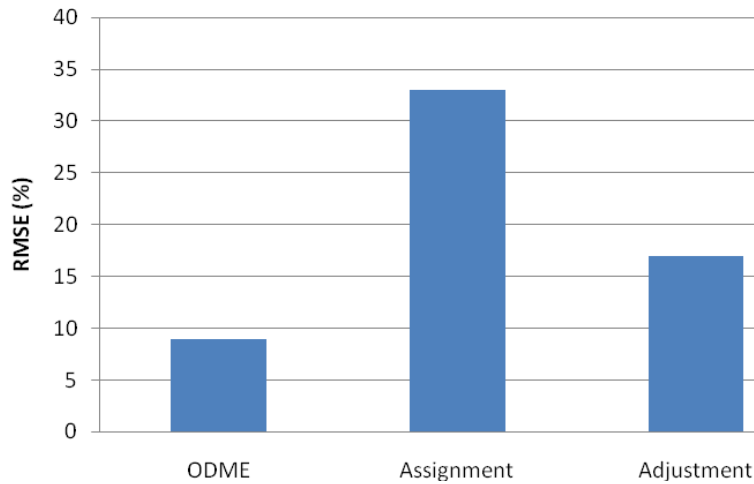
Figure 2. Relative RMSE in the AM period counts at various stages of the calibration process
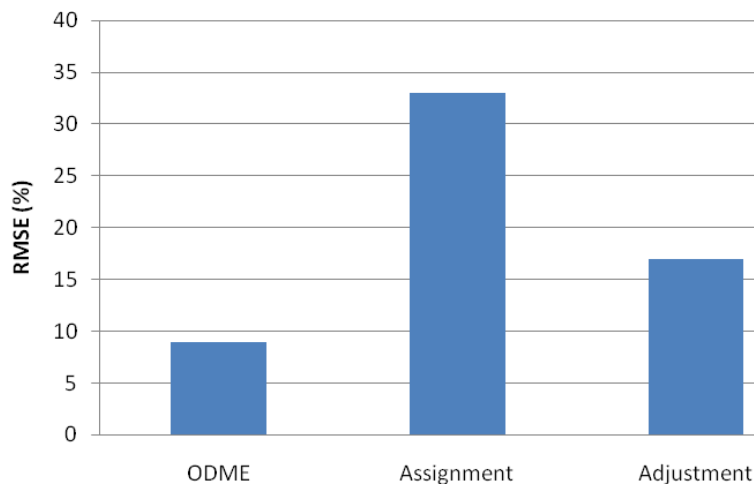
Figure 3. Relative RMSE in the PM period counts at various stages of the calibration process

# APPLICATION OF THE SIMULATION-BASED DTA

The simulation-based DTA requires no more data preparation or model development time than that required to develop a model for simulation. The DTA analysis is clearly more meaningful when used with a dynamic trip table, which is not the standard in the present state of simulation practice, though it is the opinion of the authors that it ought to be, and that the impacts of time-varying demand on network performance, both simulated and real, are grossly underestimated or overlooked.

To apply the DTA, however, demands modern computing resources, by which we mean desktop computers with 4 GB of RAM or more and multiple cores, or processors, operating at clock speeds of 2.00 GHz or more. Computers with these specifications are readily commercially available at reasonable prices. Modern computing power is needed because the DTA must simulate the full study period many times over for a given scenario or time period. Each execution of the simulation produces new reported travel times and turning movement delays that will be used to reevaluate route choices prior to the next simulation. For large models such as the GEA that have many route choices, this can mean running the simulation dozens of times such that all potentially competing route choice alternatives between all OD can be "explored," and such that the best paths (i.e., those that minimize travel costs) can be allowed to distinguish themselves from inefficient ones (i.e., those involving a lot of turns or suffering excessive turning delays).

Having no blueprint from the research or practice, nor any practical justification, for choosing a particular convergence criterion or maximum number of iterations, many experiments were tried. As many as 150 iterations of the simulation were executed in a single application of the DTA. To run many simulations quickly, TransModeler takes full advantage of all the computer's processors. On a computer with 8 GB of RAM and a 2.0 GHz Intel® Core i7 processor, which has four cores hyper-threaded, thus making eight threads available for simulating vehicles, TransModeler is able to simulate the base year PM scenario in the GEA model more than 40 times faster than real time. To simulate four hours of traffic between 2:00 and 6:00 PM takes a mere 5.7 minutes on average.

Additional processing is performed between iterations to average the travel times or path flows and to compute the relative gap for the convergence threshold, but this means that large numbers of DTA iterations can be executed in well under 24 hours. Larger models with greater and longer-lasting congestion than that found in the GEA will require longer running times, but are entirely within the means of today's mid-to-high-performance desktops.

The application of simulation-based DTA in the GEA model held a number of lessons for practice. Perhaps most interestingly, because of the stochastic nature of route choice and microscopic traffic simulation, and the "noise" generated in these models, the UE relative gap alone is not a sufficient stopping criterion. The relative gap alone masks the effects of subtle shifts in route choice that can have important implications for calibration and validation locally (i.e., at an intersection or along a corridor). For these reasons, the GEA model is allowed to run for many more iterations (i.e., on the order of 75-100) beyond the point at which the relative gap can visibly be seen to level off. As shown in Figure 4, this occurs as early as 30 iterations into the DTA.
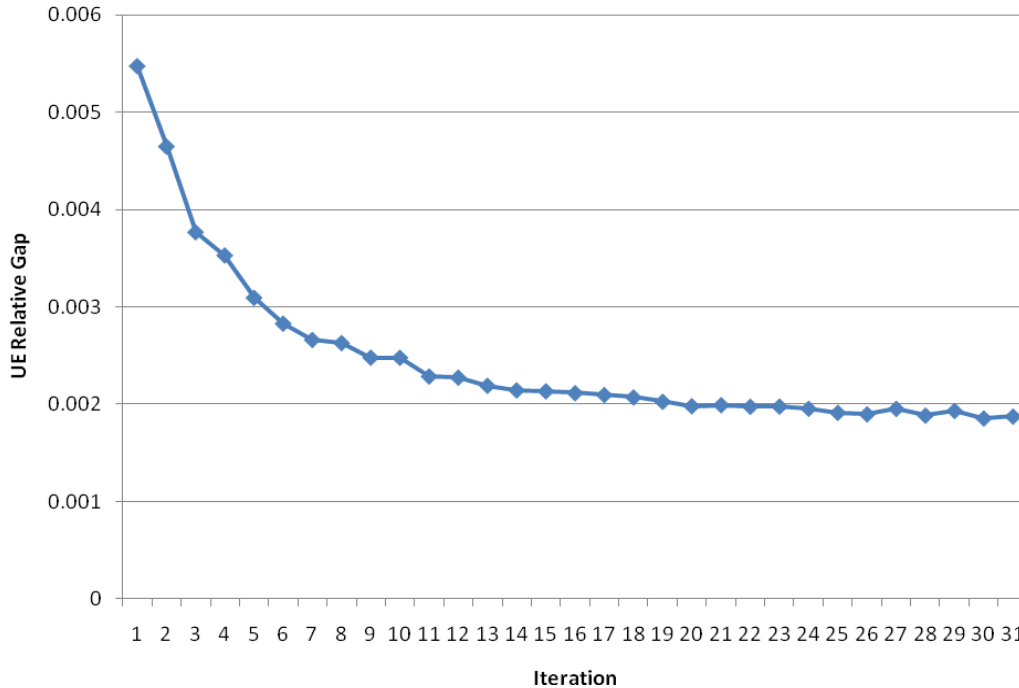
Figure 4. Relative RMSE in the PM period counts at various stages of the calibration process

Additional convergence information, such as the maximum flow change on any link in the network, are being considered for future applications.

SUMMARY

This paper gives a methodology for estimating and calibrating a microsimulation model for an entire city using a combination of travel demand methods, geographic analysis, origin-destination matrix estimation, and dynamic traffic simulation. The Eureka case study demonstrates that such an approach can be successful in overcoming some of the more difficult challenges in the application of microscopic traffic simulation to wide areas, particularly in the planning context. Trip table estimation and calibration methods in the state of the practice can be significantly improved by exploiting geographic information and analysis and by using emerging dynamic traffic assignment methods.

REFERENCES

Slavin, H., Brandon, J., Rabinowicz, A., and Sundaram, S. Application of accelerated user equilibrium traffic assignments to regional planning models, Presented at the 12th TRB National Transportation Planning Applications Conference, May 17-21, 2009, Houston, TX

Caliper Corporation (2010) What TransCAD Users Should Know about New Static Traffic Assignment Methods, Newton, MA