METHODOLOGY FOR THE CALIBRATION OF MICROSCOPIC TRAFFIC SIMULATION MODELS WITH AGGREGATE DATA

Tomer Toledo[a], Haris N. Koutsopoulos[b], and Moshe E. Ben-Akiva[c]
[a]Center for Transportation and Logistics, Massachusetts Institute of Technology, 77 Massachusetts Ave., NE20-208, Cambridge MA 02139 Tel: 617 252-1116 Fax: 617 252-1130
[b]Department of Civil and Environmental Engineering, Northeastern University, 437 Snell Engineering Center, Boston MA 02115 Tel: 617 373-4635 Fax: 617 373-4419
[c]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave., 1-181, Cambridge MA 02139 Tel: 617 253-5324 Fax: 617 253-0082
toledo@mit.edu, haris@coe.neu.edu, mba@mit.edu

Abstract

This paper presents a framework for the calibration of microscopic traffic simulation models using aggregate data. The framework takes into account the interactions between the various model parameters and the OD flows by estimating OD flows jointly with the model parameters. An optimization-based approach has been used for the joint calibration. A systematic search approach based on the Box algorithm is adopted for the solution of the resulting minimization problem. OD estimation is based on the generalized least squares (GLS) estimator. Since the calibration of the parameters depends on the estimated OD flows and vice versa, the proposed framework is iterative. The applicability of the approach is demonstrated through its application to case studies using MITSIMLab, a microscopic traffic simulation model.

Keywords: Traffic simulation; Calibration; OD estimation; Aggregate data
Topic area: D3 Integrated Supply / Demand Modeling

1. Introduction

Microscopic traffic simulation models have drawn significant attention in evaluating the impact of changes in network infrastructure (e.g. adding a lane to an existing roadway or adding a road to a network), traffic control devices (e.g. retiming of traffic signal settings or installing a ramp metering scheme) and application of Intelligent Transportation Systems (ITS). However, the reliability of such models hinges on how field conditions are captured by the parameters in the simulation model. Calibration of the simulation model is required in order to achieve the best reproducibility of field conditions. Calibration is the process by which the parameters of various components of the simulation model are set so that the model will accurately replicate observed traffic conditions.

Two groups of parameters require calibration in traffic simulation models: driving behavior parameters and travel behavior parameters. Driving behavior includes acceleration, lane-changing and intersections models. Travel behavior is represented by route choice models. In addition, OD flows are an important input to the simulation model. However, because of the spatial extend of the applications, OD matrices, let alone accurate, dynamic ones, are not readily available, and so, input OD flows need to be estimated.

Calibration of traffic simulation tools, especially microscopic ones, is not a trivial task. The source of the difficulty is that the data usually available is aggregate measurements of traffic characteristics (e.g. flows, speeds and occupancies at sensor locations, travel times, queue lengths), which are the emergent results of the interactions between various behaviors of individual vehicles. Therefore, this type of data does not support independent calibration of the various models the traffic simulator consists of. A number of papers have been published on the subject of calibration of microscopic simulation models. Overall, the treatment of the problem is at a very early stage and rather incomplete and limited. Most published studies focus on one component of the simulation model, while assuming the others are given. For example, Daigle et al. (1998), Abdulhai et al. (1999), Lee et al. (2001), Gardes et al. (2002), Kim and Rilett (2003) and Park and Schneeberger (2003) calibrate only driving behavior parameters. These studies all apply the simulation model to traffic corridors that do not involve route choice. They also assume that OD flows are either given or estimated, independent of the simulation model. Ma and Abdulhai (2002) use genetic algorithms to calibrate various parameters, including route choice parameters, but still assume given OD flows.

The calibration is in many cases an ad-hoc, sequential procedure, using algorithms that may not be appropriate for the problem. Some parameters are calibrated, often through trial and error. Their values are then fixed for the calibration of a second set and so on. Such procedures do not include feedback loops to capture interactions between the parameters of interest. Hourdakis et al. (2003) propose one such procedure. They first seek to match observed traffic flows by calibrating global parameters, such as vehicle characteristics. Next they calibrate local link-specific parameters, such as speed limits, to match observed speeds. A quasi-Newton algorithm is used for the solution of the various sub-problems.

In contrast, significant amount of research has dealt with the estimation of OD flows. Dynamic OD estimation techniques were proposed, among others, by Cremer and Keller (1987), Nihan and Davis (1987), Cascetta et al. (1993) and Ashok and Ben-Akiva (1993). These estimators rely on the availability of an assignment matrix, which captures the effects of route choice and traffic dynamics. Cascetta and Postorino (2001) suggest an OD estimation approach that explicitly includes a route choice model, but assume that the parameters of that model are given.

The literature on joint OD estimation and parameter calibration is very limited. Liu and Fricker (1996) presented a two-step heuristic search method to sequentially estimate OD flows and route choice parameters on uncongested networks. In the first step, the route choice parameters are fixed and OD flows are estimated by minimizing the difference between observed and modeled link flows. In the second step, the link flows that were obtained from the first step are used to calibrate the route choice parameters using a maximum likelihood method. Yang et al. (2001) proposed an optimization model for simultaneous estimation of OD flows and coefficient of the travel cost in a logit-based stochastic user equilibrium model. The use of an analytic model allowed the problem to be formulated as a differentiable, nonlinear optimization problem.

The objective of this paper is to present a systematic procedure for joint estimation of OD flows and calibration of behavior parameters using aggregate data. The rest of this paper is organized as follows: Section 2 describes the overall framework for the calibration of traffic simulation models and the role of aggregate calibration with in its scope. In section 3 we formulate the aggregate calibration problem. And in section 4 propose a solution approach and describe its details. The proposed approach is demonstrated through case studies in section 5. Finally, section 6 summarizes our findings and proposes directions for future research.

2. Calibration methodology

An overall framework for calibration and validation of traffic simulation models is shown in Figure 1 (Toledo et al. 2003).

The framework consists of two steps: initially, the individual models the simulation consists of (e.g. driving behavior and route choice models) are estimated using disaggregate data, independent of the overall simulation model. Disaggregate data includes detailed driver behavior information such as vehicle trajectories. The individual models may be tested independently, for example, using a holdout sample.

In the second step, the simulation model as a whole is calibrated and validated using aggregate data. The role of aggregate calibration is twofold: (i) to ensure that the interactions between the individual models within the simulator are captured correctly, and (ii) to refine previously estimated parameter values for the specific site being studied.

While this two-step approach is desirable, data availability often dictates what steps are feasible. Most often, only aggregate data collected through loop detectors is available and therefore only aggregate calibration and validation are possible.

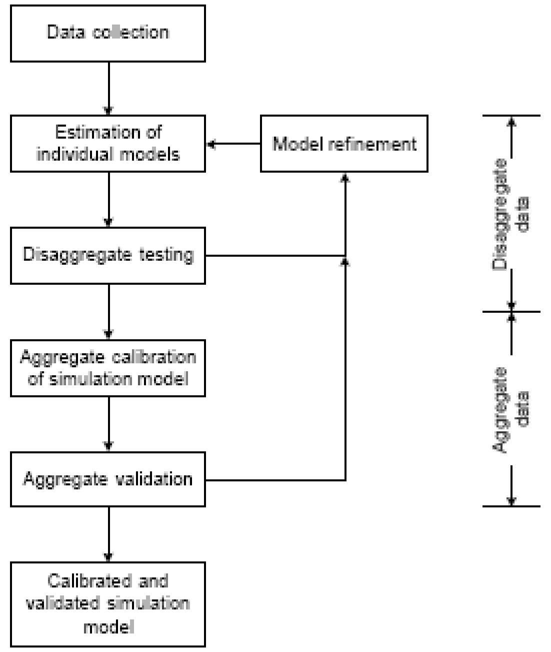

Figure 1 Overall calibration and validation framework

3. Aggregate calibration formulation

Aggregate calibration can be formulated as an optimization problem, which seeks to minimize a measure of the deviation between observed and corresponding simulated measurements and between the calibrated parameter values and their a-priori expected values. We will formulate the problem and develop solution algorithms under the assumption of stationary steady state conditions. The assumption is that the observation days are drawn during a period in which steady state traffic conditions prevail. That is, while OD flows and experienced travel times may vary for various observation days, these differences are due to random effects and do not represent a change in the underlying distributions of these variables. Furthermore, driving behavior and route choice parameters are assumed stable over the period of observation. It is important to note that the steady state assumption concerns the variability between observation days. We do not assume a steady state within each observation day.

Under the assumption of steady state traffic conditions the experienced travel times produced by the simulation model using the estimated parameters and OD flows should be consistent with the habitual travel times used as inputs to the simulation model. The resulting problem formulation is given by:

$$
\begin{aligned}
&\min_{\beta,OD} Z = \sum_{i=1}^{N}\left[f_1\left(M_i^{obs}, M^{sim}\right) + f_2\left(OD, OD^0\right) + f_3\left(\beta, \beta^0\right)\right] \\
&s.t.\quad M^{sim} = S_M(\beta, OD, TT^{hab}) \\
&\qquad TT^{hab} = TT^{exp} \\
&\qquad TT^{exp} = S_{TT}(\beta, OD, TT^{hab})
\end{aligned}
\tag{1}
$$

where,

β, OD : parameters to be calibrated: driving and route choice behavior and OD flows.

M^{obs}, M^{sim} : vectors of average (over days and replications) observed and simulated traffic measurements at various time-space points, respectively.

β^o, OD^o : a-priori behavioral parameters and OD flows.

$f_1(\cdot)$: measure of discrepancy between and .

$f_2(\cdot)$: measure of discrepancy between the calibrated and a-priori OD flows.

$f_3(\cdot)$: measure of discrepancy between the calibrated and a-priori values of behavior parameters.

TT^{hab}, TT^{exp}: time dependent expected habitual and experienced link travel times, respectively.

$S_M(\cdot)$, $S_{TT}(\cdot)$: the simulation model as the function which generates simulated traffic measurements and link travel times , respectively.

The exact form of the objective function in the formulation above depends on the assumption regarding the distribution of the modeling error. For example, under the assumption that these errors are normally distributed the generalized least squares formulation with the following objective function may be used:

$$
\begin{aligned}
Z = &\left(M^{sim} - M^{obs}\right)^T W_1^{-1}\left(M^{sim} - M^{obs}\right) + \\
&+\left(OD - OD^o\right)^T W_2^{-1}\left(OD - OD^o\right) + \left(\beta - \beta^o\right)^T W_3^{-1}\left(\beta - \beta^o\right)
\end{aligned}
\tag{2}
$$

where, *W1* , *W2* , *W3* : variance-covariance matrices of traffic measurements, OD flows and behavior parameters.

4. Solution approach

The problem formulation in equation (1) is difficult to solve. Evaluation of the objective function involves running the simulation model and is therefore computationally expensive. Furthermore, the dimensionality of the calibration parameters, in particularly the OD flows, can be very high even for networks of modest size. In order to overcome these difficulties, we propose an iterative solution approach, which is based on decomposition of the problem by parameter group (i.e. OD flows, behavior parameters). The OD estimation sub-problem The proposed approach is shown in Figure 2. This strategy creates two sub-problems: an OD estimation problem, for which existing efficient solution methods may be used, and a parameter calibration, which typically has a much lower dimensionality.

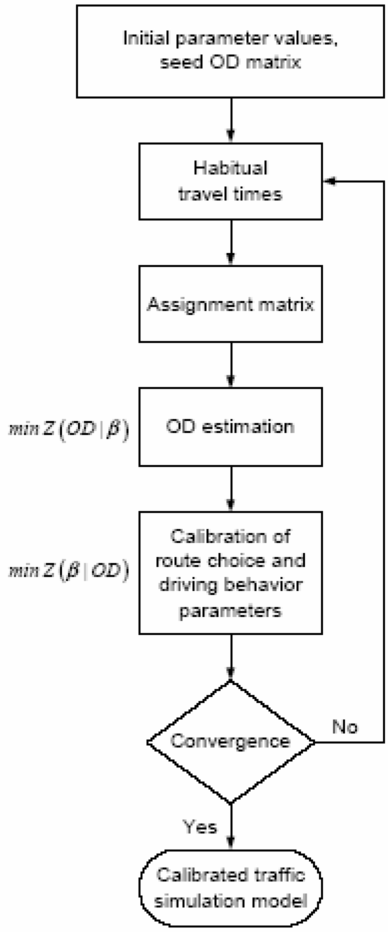

Figure 2 Solution approach to the steady state calibration problem

Every iteration consists of several steps. At each step a set of parameters are calibrated, while the remaining parameters are fixed to their previous values. The OD estimation step requires the generation of an assignment matrix, which itself depends on the route choice behavior and experienced travel times. Habitual travel times are important explanatory variables in route choice models. Hence, based on the existing OD flows and simulation parameters habitual travel times are calculated. These travel times along with the current route choice parameters are used to generate an assignment matrix. Using this assignment matrix, OD estimation can be performed (using GLS or other methods). The new OD flows are then used to re-calibrate route choice and driving behavior parameters and so on.

Various variations of the basic solution approach are possible. For example, habitual travel times may be re-calculated following the updating of each set of parameters and inputs (i.e. OD flows, route choice parameters and driving behavior parameters) or only once a full iteration in which the estimates of all the parameters have been modified is completed. Moreover, the order in which the three sets of parameters are calibrated may be modified. Another variation, considering the closer inter-dependency between OD flows and route choice parameters is to perform several iterations of these two steps before updating the driving behavior parameters. In this case, the calibration of route choice and driving behavior parameters will be done in two separate steps, and using similar mathematical formulations.

The next sections describe the various components of the solution approach in detail.

4.1. Habitual travel times

Route choices depend on habitual path travel times as explanatory variables. Calibration of the model parameters requires knowledge of these travel times. However, field measurements of travel times are most often unavailable, and planning studies can only provide initial values, which may be inconsistent with the simulation output obtained using the estimated OD flows and the calibrated parameters of the route choice model. Furthermore, travel times from planning studies are static and inconsistent with the dynamic nature of most simulation applications. For a given candidate solution for the OD matrix, route choice and driving behavior parameters, an iterative day-to-day learning model, in which habitual travel times are calculated as the weighted average of the experienced travel time and the expected travel time of the previous iteration (Cascetta and Cantarella 1991) is used. At each iteration of this process, representing a day, habitual travel times are updated as follows:

$$TT_{it}^{hab,j+1} = \lambda^j TT_{it}^{exp,j} + \left(1-\lambda^j\right) TT_{it}^{hab,j} \tag{3}$$

$TT_{it}^{hab,j}$, $TT_{it}^{exp,j}$: expected habitual and experienced travel times on link *i*, time period *t* on iteration *j*, respectively.

λ^j: weight parameter (0 <λ<1).

This process is repeated over several iterations until travel times converge to a steady state equilibrium in which habitual travel times and experienced travel times are consistent.

4.2. Assignment matrix

OD estimation, which will be described in the next section, requires an assignment matrix as input. Usually the assignment matrix, which is a function of the route choice and travel times, is not readily available and needs to be generated from the model. Assuming a path-based route choice model, the assignment matrix may be calculated analytically as:

$$a_{sh}^{rp} = \sum_{k \in K_r} \alpha_{sh}^{kp} f_{kp} \tag{4}$$

where,

a_{sh}^{rp}: fraction of vehicles from OD pair *r* departing at time interval *p*, and traveling through sensor location *s* at time interval *h*.

K_r: the set of paths connecting OD pair *r*.

α_{sh}^{kp}: fraction of vehicles using path *k* departing at time interval *p*, and traveling through sensor location *s* at time interval *h* (sensor-path fractions).

f_{kp}: fraction of demand for travel (between OD pair *r*) departing at time interval *p*, which uses path *k*.

The route choice fractions, f_{kp} , may be calculated using the route choice model implemented in the simulation model. The sensor-path fractions can be calculated using the experienced link travel times and appropriate assumptions about their distributions and about the distribution of the departure within time intervals. For example, assuming deterministic travel times and a uniform distribution of departures within time intervals, sensor-path fractions are calculated as follows:

$$\alpha_{sh}^{kp} = \begin{cases} \dfrac{pD + TT_s^{exp,kp} - (h-1)D}{D} & if\ (p-1)D + TT_s^{exp,kp} \le (h-1)D \le pD + TT_s^{exp,kp} \\ \dfrac{hD - TT_s^{exp,kp} - (p-1)D}{D} & if\ (p-1)D + TT_s^{exp,kp} \le hD \le pD + TT_s^{exp,kp} \\ 0 & otherwise \end{cases} \qquad (5)$$

where
D: the duration of a time interval.
$TT_s^{exp,kp}$: experienced travel time from the origin to sensor location *s* for vehicles departing at time interval *p* and using path *k*.

Alternatively, in cases where the analytical calculation of the assignment matrix is prohibitively expensive (e.g. when a link-based route choice model is used or in the presence of traveler information) the assignment matrix may be inferred directly from simulation realizations. Multiple runs of the simulation should be used in order to obtain the *expected* value of the assignment matrix, rather than the one corresponding to a specific realization.

Regardless of the way the assignment matrix is calculated it may be beneficial to smooth it with previous estimates. In particular, smoothing may be useful if the assignment matrix is derived from a small number of simulation realizations. In this case the stochastic nature of the simulation model may cause the values of various entries in the matrix, especially those corresponding to OD pairs with very low demand, to fluctuate erratically from iteration to iteration. Typically, the smoothing function will take the form:

$$A^{j+1} = \pi^j A^j + \left(1 - \pi^j\right) \hat{A}^{j+1} \qquad (6)$$

where,
A^j, A^{j+1} : assignment matrix used in iterations *j* and *j+1*, respectively.
$\hat{A}^{j+1}$: assignment matrix estimated for iteration *j+1*, either analytically or from simulation realizations.

π^j : smoothing factor, for example using MSA: $\pi^j = \dfrac{j}{j+1}$.

Although the smoothing formulation above only uses the assignment matrix from the previous iteration, the effects of earlier values are implicitly captured since they are encapsulated in this matrix. It is also straightforward to extend the above formulation to explicitly include additional previous estimates (at the cost of having to maintain these estimates).

4.3. OD Estimation

The OD estimation problem requires three sets of inputs: traffic measurements, a seed OD matrix (which includes a-priori estimates of OD flows) and an assignment matrix. OD flows estimated in previous studies may be used as the seed OD flows. Seed OD flows may also be extracted from planning models. Although this matrix may not be up to date, it still contains valuable information regarding the structural relationships among OD pairs (Ashok and Ben-Akiva 2000) and therefore can improve the quality of the solution. The assignment matrix, as discussed above, is usually estimated from the simulation model itself. Assuming a known assignment matrix *A*, a seed OD matrix and that the available measurements are traffic counts, the OD estimation sub-problem is formulated as a constrained optimization problem, which seeks to minimize the deviations between estimated and observed traffic counts while also minimizing the deviation between the estimated OD flows and seed OD flows. The constraint being imposed is that OD flows are non-negative. In this paper, we adopt the GLS formulation proposed in Cascetta and Nguyen (1988) although other methods may also be used. The GLS formulation is given by:

$$\min_{X\geq 0}\left(AX-Y^{obs}\right)^T W_1^{-1}\left(AX-Y^{obs}\right)+\left(X-OD^o\right)^T W_2^{-1}\left(X-OD^o\right) \tag{7}$$

where,

Y^{obs}: observed traffic counts at sensor locations.

A : assignment matrix that maps OD flows to counts at sensor locations.

As with the assignment matrix and for similar reasons, smoothing of OD matrices with those estimated in previous iterations may be useful. For example, using the OD matrix from the previous iteration:

$$OD^{j+1}=\rho^j OD^j+\left(1-\rho^j\right)\widehat{OD}^{j+1} \tag{8}$$

where,

OD^j, OD^{j+1}: OD matrix used in iteration j and $j+1$, respectively.

$\widehat{OD}^{j+1}$: OD matrix estimated for iteration j using equation (7).

ρ^j: smoothing factor, for example using MSA: $\rho^j=\dfrac{j}{j+1}$.

$j+1$

4.4. Parameter calibration

Using the GLS formulation and given OD flows and habitual travel times, the parameters of the driving behavior models and the route choice model are obtained as the solution to the following optimization problem:

$$\begin{aligned}&\min_{\beta} Z=\left(M^{sim}-M^{obs}\right)^T W_1^{-1}\left(M^{sim}-M^{obs}\right)+\left(\beta-\beta^o\right)^T W_3^{-1}\left(\beta-\beta^o\right)\\&s.t.\\&M^{sim}=S(\beta,\theta,OD,TT)\end{aligned} \tag{9}$$

The above formulation assumes that driving behavior and route choice parameters are calibrated jointly. However, for a specific application, there may be reason to believe that either travel behavior or driving behavior is the main source of the simulation error. In this case, variations of the calibration procedure may be utilized. For example, if it is assumed that travel behavior is an important source of error, a procedure in which several iterations of OD estimation and route choice calibration are performed before a driving behavior calibration step may yield better results. Toledo et al. (2003) apply this variation to a network in Stockholm, Sweden. The formulations of the route choice and driving behavior parameter calibration steps are similar to the one presented in equation (9), but with only the relevant set of parameters in each case.

The selection of a solution algorithm for the parameter calibration problem must recognize the simulation stochasticity. For the case studies reported in this paper, a sequential search technique based on Box's Complex algorithm (Box 1965) was used. The advantage of the Complex algorithm for this application is that, it only requires calculation of the objective function value and does not use any gradient information, which is difficult to calculate accurately in a stochastic model. The Complex algorithm is a sequential search technique designed for nonlinear optimization problems with nonlinear constraints. The algorithm is initiated by randomly generating a set of m feasible starting points in the n-dimensional space of the decision variables ($mn\geq+1$, Box recommends $m=2n$). The objective function at each point is evaluated and at each iteration, point *w*, the one with the worst (highest) objective value, is replaced by a new point which is $\alpha \geq 1$ times as far from the centroid of the remaining points as the reflection of point *w* in the centroid:

$$X_w^{j+1} = \alpha\left(X_{cen}^j - X_w^j\right) + X_{cen}^j \tag{10}$$

X_w^j and X_w^{j+1} are the values of point *w*, at iterations *j* and *j+1*, respectively. X^j is the centroid of the remaining points. α is a parameter. Box recommends the value 1.3.

If a point repeats as the worst one in consecutive iterations, it is moved half the distance to the centroid of the remaining points:

$$X_w^{j+1} = \frac{1}{2}\left(X_{cen}^j + X_w^j\right) \tag{11}$$

If the generated point violates any of the constraints it is moved a small distance δ into the feasible area for constraints involving a single variable or towards the centroid of the remaining points according to equation (11) for more complex constraints. The same procedure is also used to generate the feasible initial set of points. These steps are repeated until convergence is reached.

The initial points are spread over the entire feasible region, and therefore, the algorithm tends to find a global optimum solution. This is an important advantage of the Complex algorithm in the context of traffic simulation models, given the highly nonlinear nature of these models.

Every objective function evaluation requires running several replications of the simulation model. Therefore, the computational time increases significantly with the number of parameters to calibrate. However, the number of parameters in a microscopic traffic simulation model is typically very large. Thus, it is crucial to identify the set of parameters that have the largest impact on the simulation outputs through sensitivity analysis and focus the calibration effort on a these parameters.

5. Case studies

In this section we present the application of the proposed calibration methodology to calibrate the microscopic traffic simulation model MITSIMLab (Yang and Koutsopolous 1996) for two different applications. The first is a small application with known OD flows and no route choice that allows testing the behavior of the Complex algorithm for parameter calibration. The second is a complex, medium scale network that demonstrates the complete calibration methodology.

5.1. Case study 1: the HCQS network

The Highway Capacity and Quality of Service (HCQS) committee of Transportation Research Board developed an artificial case study to test various traffic simulation models (Bloomberg et al. 2003). The network, shown schematically in Figure 3, includes a freeway and two intersecting arterials. In addition to the detailed geometric layout, vehicles mix and signal settings, average link speeds and densities during the third 15-minute interval within the peak hour were available for calibration. This case study assumes perfect knowledge of OD flows. Moreover, no route choice is present in the network. Therefore, driving behavior is the only source of simulation error and the only component to be calibrated.

A sensitivity analysis was performed to find the most important parameters for calibration. The network is relatively uncongested, and so, simulation results were more sensitive to the parameters of the acceleration behavior compared to lane-changing parameters. The acceleration model implemented in MITSIMLab includes three driving regimes: car-following acceleration, car-following deceleration and free-flow (see Ahmed

1999 for details). The scale parameters of these three models were selected for calibration. The formulation set forth in equation (9) was used for the parameter calibration. Based on the available data, the following from of the objective function was used:

$$f(M^{obs}, M^{sim}) = \sum_i \left(v_i^{obs} - v_i^{sim}\right)^2 + \omega \sum_i \left(k_i^{obs} - k_i^{sim}\right)^2 \qquad (12)$$

v_i^{obs} and v_i^{sim} are the observed and simulated speeds on link *i*, respectively.

k_i^{obs} and k_i^{sim} are the corresponding density measurements. ω is a weight parameter for the density measurements (the weights for speed measurements are normalized to 1).

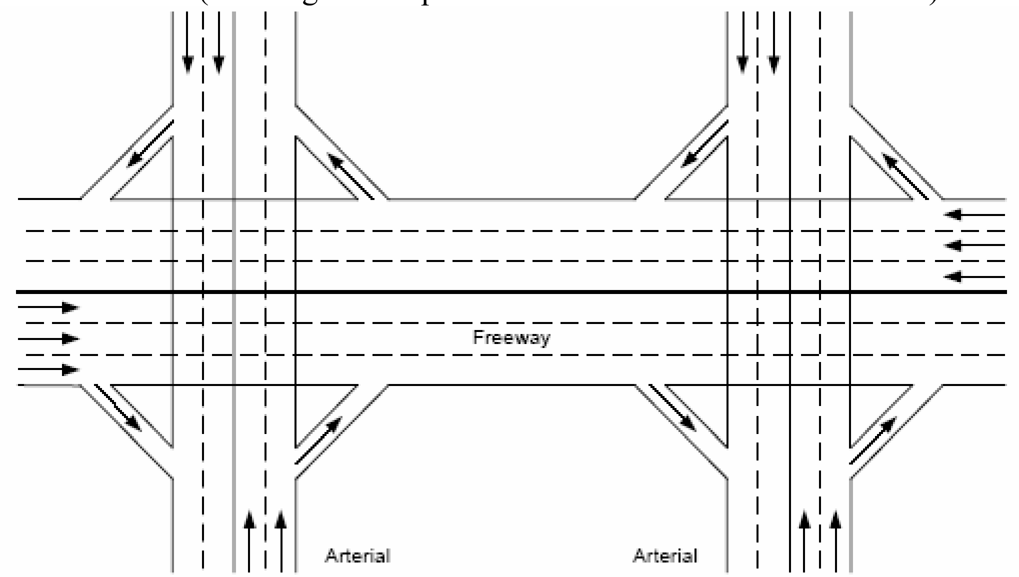

Figure 3 The HCQS network

A set of six candidate points was used with the Complex algorithm. The convergence of the algorithm in terms of the lowest and highest objective values is shown in Figure 4. The difference between the highest and the lowest objective values after 20 iterations was 4.4%. Convergence is also achieved in the parameter value. The values of the three parameters for all six points are within 1.9%, 6.2% and 0.2% of the point with the lowest objective value. Since OD flows were perfectly known, simulated counts exactly matched the observed ones. Comparisons of the observed and simulated speed measurements on the freeway and arterial segments are shown in Figure 5. The results show a good fit between observed and simulated data. The root mean squared error (RMSE) of the model is 1.6 mph and the root mean squared percent error is 7.1%. The mean absolute error (MAE) is 1.3 mph and the mean absolute percent error (MAPE) is 4.6%.

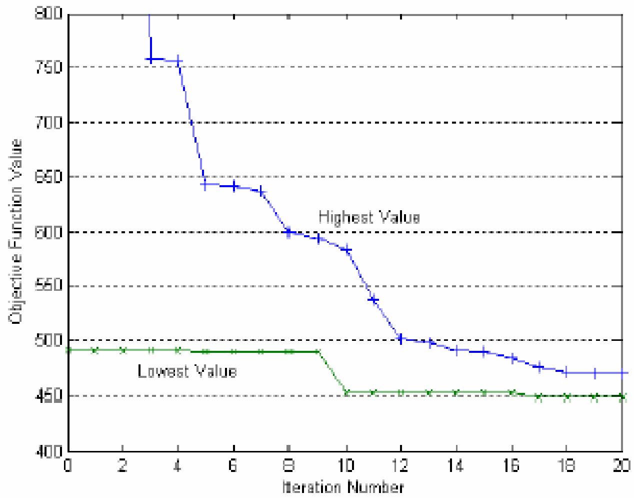

Figure 4 Convergence of the Complex algorithm

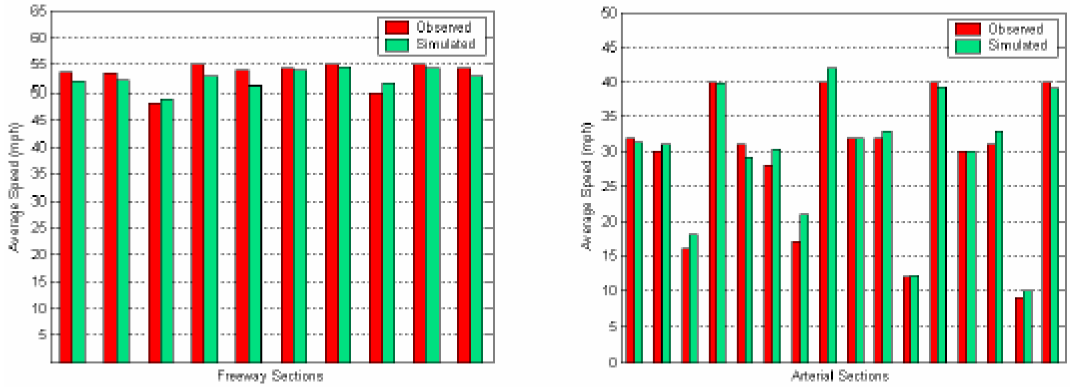

Figure 5 Comparison of observed and simulated speeds in the HCQS network

5.2. Case study 2: Irvine, CA

The study network, shown in Figure 6, is comprised of three major freeways (I-5, I-405 and Route 133) and a dense network of urban arterials around them. The simulation representation consists of 298 nodes and 618 links. There are 80 signalized intersections within the study area. Data for calibration included time-dependent loop detector data collected from 68 sensor stations (30 on freeways, 38 on arterials) during 5 weekdays, and a static seed OD matrix from a planning study.

Figure 6 The Irvine CA network

The availability of multiple days of data raises a question whether the assumption that stationary steady state conditions exist is realistic or not. The data shows little day-to-day variability, as Figure 7, which plots of time-dependent traffic counts for different days at two sensors locations, indicates. This result supports the steady state assumption and so, all observations were used jointly in the calibration process.

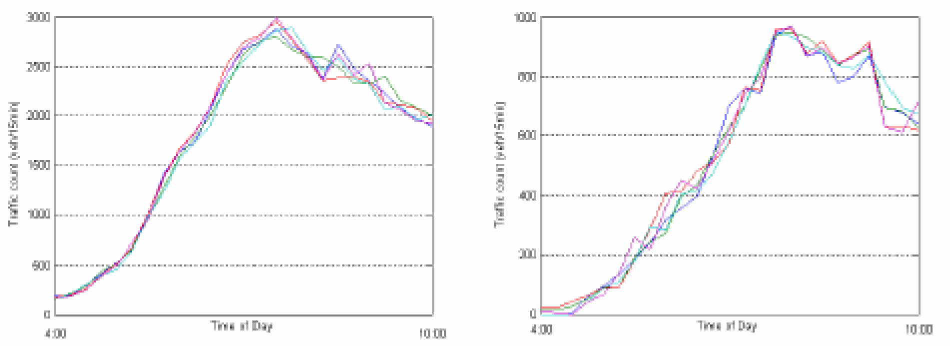

Figure 7 Variability of traffic counts across days

The simulation runs were performed for the AM peak period. Measurements were aggregated in intervals of 15 minutes. The calibration included estimation of OD flows and calibration of the travel time coefficient of the route choice model, the acceleration scale parameters described in the HCQS case study and two constants, one in each of the critical gap functions used to determine gap acceptance of the lead and lag gap in lane changing. The objective function in the parameter calibration step [equation (9)] was defined by:

$$f(M^{obs}, M^{sim}) = \sum_i \left(y_i^{obs} - y_i^{sim} \right)^2 \qquad (13)$$

y_i and y_i^{sim} are the average observed and simulated traffic counts at time space point i, respectively.

Calibration results, comparing observed and simulated sensor counts are shown in Figure 8 for two of the time intervals. The RMSE and MAE of the model are 15.9 and 9.2 vehicles per interval, respectively.

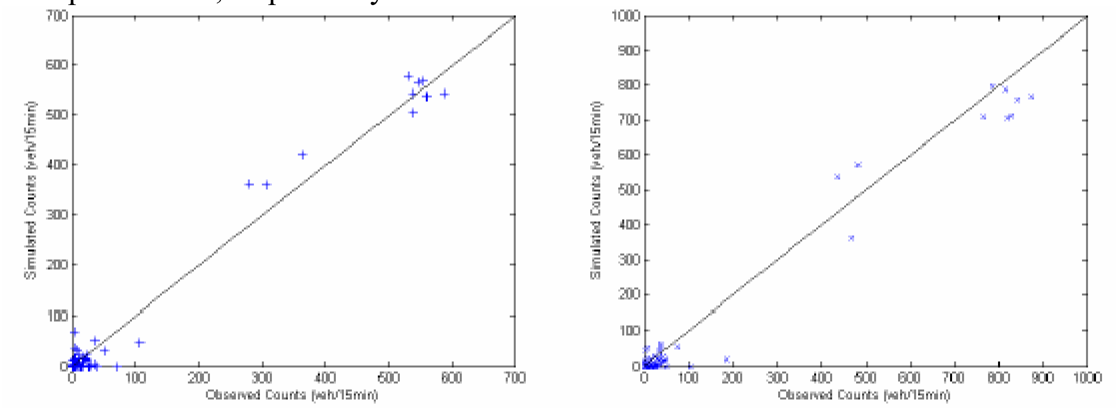

Figure 8 Comparison of observed and simulated counts in the Irvine CA network

6. Conclusion

A framework for the calibration of microscopic traffic simulation models using aggregate data was presented. The framework takes into account the interactions between the various model parameters and the OD flows by estimating OD flows jointly with the model parameters. An optimization-based approach has been used for the joint calibration. Since the calibration of the parameters depends on the estimated OD flows and vice versa, the proposed framework is iterative. OD estimation is based on the well known GLS estimator. A systematic search approach based on Box's Complex algorithm is adopted for calibration of the parameters. This algorithm is particularly useful for the problem at hand since it does not require calculations of derivatives of the objective function, which would have been prohibitively expensive in the context of a simulation model. Moreover, the algorithm uses a set of initial starting points randomly spread over the search space and so, tends to find the global rather than local optimum points. Nevertheless, further research is required to identify efficient algorithms to perform the parameter calibration step.

The applicability of the approach was demonstrated through its application to case studies using MITSIMLab, a microscopic traffic simulation model. While the results are promising, the case studies also demonstrated that further research is needed to improve computational performance. Research directions that may lead to improvements include development of more efficient optimization algorithms for calibration, and OD estimation techniques, and extensions to non-stationary cases.

Acknowledgement

This material is based upon work supported by Jacobs Civil Inc. and by the National Science Foundation under Grant No. CMS-0339005 and CMS-0339108. The authors would like to thank Markos Papageorgiou who suggested the use of the Complex algorithm.

References

Abdulhai B., Sheu J.B. and Recker W., 1999. Simulation of ITS on the Irvine FOT Area Using 'Paramics 1.5' Scalable Microscopic Traffic Simulator: Phase I: Model Calibration and Validation, California PATH Research Report UCB-ITS-PRR-99-12.

Ahmed, K.I., 1999. Modeling Drivers' Acceleration and Lane Changing Behavior, PhD Dissertation, Department of Civil and Environmental Engineering, MIT.

Ashok, K., and Ben-Akiva, M.E., 1993. Dynamic Origin-Destination Matrix Estimation and Prediction for Real-Time Traffic Management Systems, in International Symposium on Transportation and Traffic Flow Theory, C.F. Daganzo, ed., pp. 165-181, Elsevier Science, New-York.

Ashok, K., and Ben-Akiva, M.E., 2000. Alternative Approaches for Real-Time Estimation and Prediction of Time-Dependent Origin-Destination Flows, Transportation Science, 34 21-36.

Bloomberg, L., Swenson, M., and Haldors, B., 2003. Comparison of Simulation Models and the HCM, in Preprints of the 82nd Transportation Research Board Annual Meeting, Washington DC, January 2003 (CD-ROM).

Box, E.P., 1965. A New Method for Constrained Optimization and a Comparison with Other Methods, Computer Journal, 8, 42-52.

Cascetta, E., and Cantarella, C., 1991. A Day-to-day and within Day Dynamic Stochastic Assignment Model, Transportation Research Part A, 25 277-291.

Cascetta, E., Inaudi, D., and Marquis, G., 1993. Dynamic Estimators of Origin-Destination Matrices Using Traffic Counts, Transportation Science, 27 363-373.

Cascetta, E., and Nguyen, S., 1988. A Unified Framework for Estimating or Updating Origin/Destination Matrices from Traffic Counts, Transportation Research Part B, 22 437-455.

Cascetta, E., and Postorino, M.N., 2001. Fixed Point Approaches to the Estimation of O/D Matrices Using Traffic Counts on Congested Networks, Transportation Science, 35 134 –147.

Cremer, M., and Keller, H., 1987. A New Class of Dynamic Methods for the Identification of Origin-Destination Flows, Transportation Research Part B, 21 117-132.

Daigle, G., Thomas, M., and Vasudevan, M., 1998. Field Applications of CORSIM: I-40 Freeway Design Evaluation, Oklahoma City, OK, in Proceedings of the 1998 Winter Simulation Conference, D.J. Medeiros, E.F. Watson, J.S. Carson and M.S. Manivannan, eds., pp. 1161-1168.

Gardes, Y., May, A.D., Dahlgren, J., and Skabardonis, A., 2002. Freeway Calibration and Application of the Paramics Model, in Preprints of the 81st Transportation Research Board Annual Meeting, Washington DC (CD-ROM).

Hourdakis, J., Michalopoulos, P.G., and Kottommannil, J., 2003. A Practical Procedure for Calibrating Microscopic Traffic Simulation Models, in Preprints of the 82nd Transportation Research Board Annual Meeting, Washington DC (CD-ROM).

Kim, K., and Rilett, L.R., 2003. Simplex Based Calibration of Traffic Micro-Simulation Models Using ITS data, in Preprints of the 82nd Transportation Research Board Annual Meeting, Washington DC (CD-ROM).

Lee, D.H., Yang, X., and Chandrasekhar, P., 2001. Parameter Calibration for PARAMICS Using Genetic Algorithm, in Preprints of the 80th Transportation Research Board Annual Meeting, Washington DC (CD-ROM).

Liu, S.S., and Fricker, J.D., 1996. Estimation of a Trip Table and the Theta Parameter in a Stochastic Network, Transportation Research Part A, 30, 287-305.

Ma, T., and Abdulhai, B., 2002. Genetic Algorithm-Based Optimization Approach and Generic Tool for Calibrating Traffic Microscopic Simulation Parameters, Transportation Research Record, 1800, 6-15.

Nihan, N.L., and Davis, G.A., 1987. Recursive Estimation of Origin-Destination Matrices from Input Output Counts. Transportation Research Part B, 21 149-163.

Park, B., and Schneeberger, J.D., 2003. Microscopic Simulation Model Calibration and Validation: A Case Study of VISSIM for a Coordinated Actuated Signal System, in Preprints of the 82nd Transportation Research Board Annual Meeting, Washington DC (CD-ROM).

Toledo, T., Koutsopoulos, H.N., Davol, A., Ben-Akiva, M.E., Burghout, W., Andréasson, I., Johansson, T. and Lundin, C., 2003. Calibration and Validation of Microscopic Traffic Simulation Tools: Stockholm Case Study, in Preprints of the 82nd Transportation

Research Board Annual Meeting, Washington DC.

Yang, Q., and Koutsopoulos, H.N., 1996. A Microscopic Traffic Simulator for Evaluation of Dynamic Traffic Management Systems, Transportation Research Part C, 4, pp. 113-129.

Yang, H., Meng, Q., and Bell, M.G.H., 2001. Simultaneous Estimation of the Origin-Destination Matrices and Travel-Cost Coefficient for Congested Networks in a Stochastic User Equilibrium, Transportation Science, 35 107-123.