

# Calibration of Microscopic Traffic Simulation Models with Aggregate Data

Tomer Toledo, Moshe E. Ben-Akiva, Deepak Darda, Mithilesh Jha, and Haris N. Koutsopoulos

**A framework for the calibration of microscopic traffic simulation models using aggregate data is presented. The framework takes into account the interactions between the various inputs and parameters of the simulator by estimating origin–destination (O-D) flows jointly with the behavioral parameters. An optimization-based approach is used for the joint calibration. Since the calibration of the parameters depends on the estimated O-D flows and vice versa, the proposed framework is iterative. O-D estimation is based on the well-known generalized least squares estimator. A systematic search approach based on the complex algorithm is adopted for calibration of the behavioral parameters. This algorithm is particularly useful for the problem at hand since it does not require calculations of derivatives of the objective function. The applicability of the approach is demonstrated through its application to case studies using MITSIMLab, a microscopic traffic simulation model.**

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 intelligent transportation systems (ITS) applications. 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, origin–destination (O-D) flows are an important input to the simulation model. However, because of the spatial extent of the applications, O-D matrices, especially accurate, dynamic

ones, are not readily available, and so input O-D 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 are aggregate measurements of traffic characteristics (e.g., flows, speeds and occupancies at sensor locations, travel times, and 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 that comprise the traffic simulator.

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. (1), Abdulhai et al. (2), Lee et al. (3), Gardes et al. (4), Kim and Rilett (5), and Park and Schneeberger (6) 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 O-D flows are either known or estimated independently of the simulation model. Ma and Abdulhai (7) use genetic algorithms to calibrate various parameters, including route choice parameters, but still assume known O-D 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. (8) 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 subproblems.

In contrast, a significant amount of research has dealt with the estimation of O-D flows. Dynamic O-D estimation techniques were proposed, among others, by Cremer and Keller (9), Nihan and Davis (10), Cascetta et al. (11), and Ashok and Ben-Akiva (12). These estimators rely on the availability of an assignment matrix, which captures the effects of route choice and traffic dynamics. Cascetta and Postorino (13) suggest an O-D estimation approach that explicitly includes a route choice model, but they assume that the parameters of that model are given.

The literature on joint O-D estimation and parameter calibration is very limited. Liu and Fricker (14) presented a two-step heuristic

---

T. Toledo, Center for Transportation and Logistics, NE20-208, and M. E. Ben-Akiva, Department of Civil and Environmental Engineering, 1-181, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139. D. Darda, IBI Group, 3 Copley Place, 3rd Floor, Boston, MA 02116. M. Jha, Jacobs Civil, Inc., 222 South Riverside Plaza, Chicago, IL 60606. H. N. Koutsopoulos, Department of Civil and Environmental Engineering, Northeastern University, 437 Snell Engineering Center, Boston, MA 02115.

*Transportation Research Record: Journal of the Transportation Research Board*, No. 1876, TRB, National Research Council, Washington, D.C., 2004, pp. 10–19.

search method to sequentially estimate O-D flows and route choice parameters on uncongested networks. In the first step, the route choice parameters are fixed and O-D 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. (15) proposed an optimization model for simultaneous estimation of O-D flows and the 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 O-D flows and calibration of behavior parameters using aggregate data. The paper is organized as follows: the first section describes the overall framework for the calibration of traffic simulation models and the role of aggregate calibration within this framework. The aggregate calibration problem is then formulated. Next a solution approach is proposed and its details are described. The application of the proposed approach is demonstrated through case studies in the following section. Finally, the findings are summarized and directions for future research are proposed.

## CALIBRATION METHODOLOGY

An overall framework for calibration and validation of traffic simulation models is shown in Figure 1 (16). The framework consists of two steps. Initially, the individual models that comprise the simulation (e.g., driving behavior and route choice models) are estimated using disaggregate data, independent of the overall simulation model. Disaggregate data include detailed driver behavior information such as vehicle trajectories. The individual models may be tested indepen-

dently (e.g., by 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:

1. To ensure that the interactions between the individual models within the simulator are captured correctly and
2. To refine previously estimated parameter values for the specific site being studied.

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

## 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 O-D flows and the corresponding a priori expected values. The problem will be formulated and solution algorithms will be developed 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 O-D 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. A steady state is not assumed 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 O-D flows should be consistent with the habitual travel times used as inputs to the simulation model. The resulting problem formulation is given by

$$\begin{aligned} \min_{\beta, OD} Z &= f_1(M^{obs}, M^{sim}) + f_2(OD, OD^o) + f_3(\beta, \beta^o) \\ \text{s.t.} \\ TT^{hab} &= TT^{exp} \\ M^{sim} &= S_M(\beta, OD, TT^{hab}) \\ TT^{exp} &= S_{TT}(\beta, OD, TT^{hab}) \end{aligned} \quad (1)$$

where

$\beta, OD$  = parameters to be calibrated: driving behavior and route choice parameters and O-D flows, respectively;

$M^{obs}, M^{sim}$  = vectors of the average (over days) observed traffic measurements at various time-space points and the corresponding (average over replications) simulated measurements;

$\beta^o, OD^o$  = a priori values of behavioral parameters and O-D flows;

$f_1(\cdot)$  = measure of discrepancy between observed and simulated traffic measurements;

$f_2(\cdot)$  = measure of discrepancy between the calibrated and a priori O-D flows;

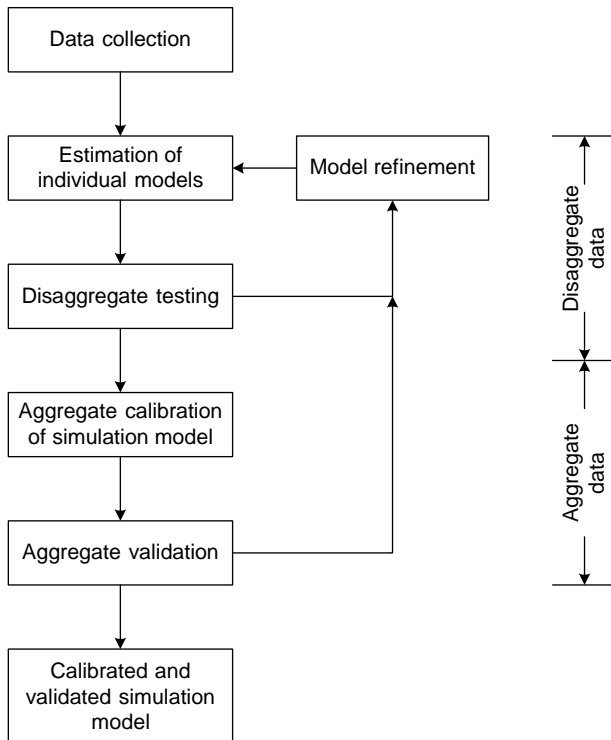


FIGURE 1 Overall calibration and validation framework.

$f_3(\cdot)$  = measure of discrepancy between the calibrated and a priori values of behavior parameters;

$TT^{\text{hab}}, TT^{\text{exp}}$  = time-dependent expected habitual and experienced link travel times, respectively; and

$S_M, S_{TT}$  = simulator as function that generates average simulated traffic measurements and experienced link travel times, respectively.

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

$$\begin{aligned} Z = & (M^{\text{sim}} - M^{\text{obs}})' W_1^{-1} (M^{\text{sim}} - M^{\text{obs}}) \\ & + (OD - OD^o)' W_2^{-1} (OD - OD^o) \\ & + (\beta - \beta^o)' W_3^{-1} (\beta - \beta^o) \end{aligned} \quad (2)$$

where  $W_1$ ,  $W_2$ , and  $W_3$  are variance–covariance matrices of the traffic measurements, O-D flows, and behavior parameters, respectively.

## 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 parameters to be calibrated, in particular the O-D flows, can be very high even for networks of modest size. In order to overcome these difficulties, an iterative solution approach is proposed, which is based on decomposition of the problem by parameter group (i.e., O-D flows, behavior parameters). The proposed approach is shown in Figure 2. This strategy creates two subproblems: an O-D estimation problem, for which existing efficient solution methods may be used, and a parameter calibration, which typically has a much lower dimensionality.

An iteration of the solution algorithm consists of several steps. In each step a set of parameters is calibrated, while the remaining parameters are fixed to their previous values. The O-D 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 O-D 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. With this assignment matrix, O-D estimation can be performed (using GLS or other methods). The new O-D flows are then used to recalibrate 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 recalculated following the updating of each set of parameters and inputs (i.e., O-D 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, which may be useful considering the closer interdependency between O-D flows and route choice parameters, is to perform several iterations of these two steps before updating the driving behavior

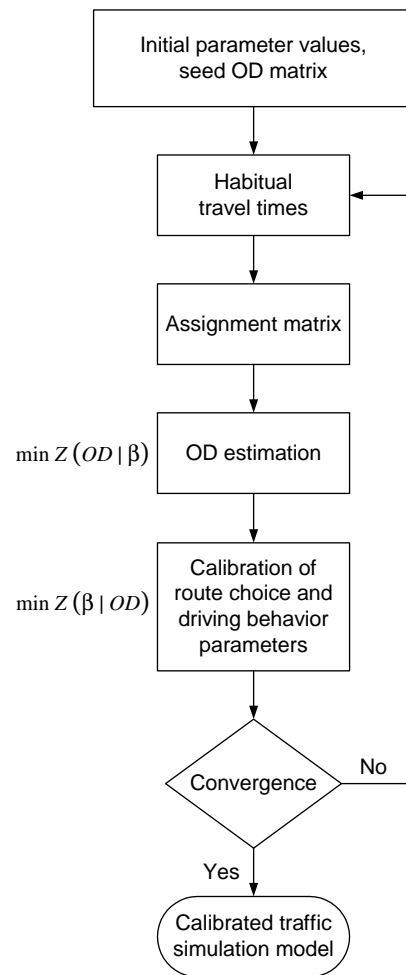


FIGURE 2 Solution approach to steady-state calibration problem.

parameters. In this case, the calibration of route choice and driving behavior parameters will be done in two separate steps. The various components of the solution approach are described in further detail in the following text.

## 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, direct field measurements of travel times are most often unavailable. Planning studies can often provide initial values. However, travel times from planning studies are static and do not capture the dynamics of traffic flow. Furthermore, habitual travel times that are input to the simulation model must be consistent with the experienced travel times obtained as output from the simulation. Obtaining consistent travel times essentially requires the solution of a fixed-point problem, which is represented by the first constraint in the problem formulation in Equation 1.

For a given candidate solution for the O-D matrix, route choice, and driving behavior parameters, an iterative averaging method, 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 (17), is used. In every iteration of this process, habitual travel times are updated as follows:

$$TT_{it}^{\text{hab},j+1} = \lambda TT_{it}^{\text{exp},j} + (1 - \lambda) TT_{it}^{\text{hab},j} \quad (3)$$

where  $TT_{it}^{\text{hab},j}$  and  $TT_{it}^{\text{exp},j}$  are expected habitual and experienced travel times on link  $i$ , time period  $t$  in iteration  $j$ , respectively, and  $\lambda$  is the weight parameter ( $0 < \lambda < 1$ ). 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.

## Assignment Matrix

O-D estimation, which is described in the next section, requires an assignment matrix as input. The assignment matrix depends on the route choice and travel times. It is usually 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

$$a_{sh}^r = \sum_{k \in K_r} \alpha_{sh}^{kp} f_{kp} \quad (4)$$

where

$\alpha_{sh}^r$  = fraction of vehicles that departed during time interval  $p$  with O-D pair  $r$ , and passed through sensor location  $s$  during time interval  $h$ ;

$K_r$  = set of paths connecting O-D pair  $r$ ;

$\alpha_{sh}^{kp}$  = fraction of vehicles that departed during time interval  $p$  using path  $k$  (between O-D pair  $r$ ), and passed through sensor location  $s$  during time interval  $h$  (sensor-path fractions); and

$f_{kp}$  = fraction of demand for travel (between O-D pair  $r$ ) that departed during 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 among drivers and about the distribution of the departures 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:

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

where  $D$  is the duration of a time interval and  $TT_s^{\text{exp},kp}$  is the experienced travel time from the origin to sensor location  $s$  for vehicles that depart during time interval  $p$  and use path  $k$ .

Alternatively, in cases in which 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 in which 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 O-D pairs with very low demand, to fluctuate erratically from iteration to iteration. Typically, the smoothing function will take the form

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

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; and

$\pi^j$  = smoothing factor, for example using method of successive averages (MSA):  $\pi^j = [j/(j+1)]$ .

Although the smoothing preceding formulation 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).

## O-D Estimation

The O-D estimation problem requires three sets of inputs: traffic measurements, a seed O-D matrix (which includes a priori estimates of O-D flows), and an assignment matrix. O-D flows estimated in previous studies may be used as the seed O-D flows. Seed O-D 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 O-D pairs (18) and therefore can improve the quality of the solution. The assignment matrix, as previously discussed, is usually estimated from the simulation model itself. Assuming a known assignment matrix  $A$ , a seed O-D matrix, and that the available measurements are traffic counts, the O-D estimation subproblem 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 O-D flows and seed O-D flows. The constraint being imposed is that O-D flows are nonnegative. This paper adopts the GLS formulation proposed by Cascetta and Nguyen (19), although other methods may also be used. The GLS formulation is given by

$$\min_{X \geq 0} (AX - Y^{\text{obs}})' W_1^{-1} (AX - Y^{\text{obs}}) + (X - OD^o)' W_2^{-1} (X - OD^o) \quad (7)$$

where  $Y^{\text{obs}}$  is observed traffic counts at sensor locations and  $A$  is the assignment matrix that maps O-D flows to counts at sensor locations.

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

$$OD^{j+1} = \rho^j OD^j + (1 - \rho^j) \hat{OD}^{j+1} \quad (8)$$

where

$$\begin{aligned} OD^j, OD^{j+1} &= \text{O-D matrix used in iteration } j \text{ and } j+1, \text{ respectively;} \\ \hat{OD}^{j+1} &= \text{O-D matrix estimated for iteration } j \text{ using Equation 7; and} \\ \rho^j &= \text{smoothing factor, for example using MSA:} \\ &\rho^j = [j/(j+1)]. \end{aligned}$$

## Parameter Calibration

Assuming the GLS formulation in Equation 2, and given O-D flows and habitual travel times, the parameters of the driving behavior and the route choice models are obtained as the solution to the following optimization problem:

$$\begin{aligned} \min_{\beta} Z &= (M^{\text{sim}} - M^{\text{obs}})' W_1^{-1} (M^{\text{sim}} - M^{\text{obs}}) \\ &+ (\beta - \beta^o)' W_3^{-1} (\beta - \beta^o) \end{aligned}$$

s.t.

$$M^{\text{sim}} = S(\beta, OD, TT^{\text{hab}}) \quad (9)$$

The preceding 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 O-D estimation and route choice calibration are performed before a driving behavior calibration step may yield better results. Toledo et al. (16) applied 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 subset 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, Box's complex algorithm (20) was used. The complex algorithm is a sequential search technique designed for nonlinear optimization problems with nonlinear constraints. The advantage of this algorithm for this application is that it only requires calculation of the objective function value and does not use any derivative information, which is difficult to calculate accurately in a stochastic model. Initially, a set of  $m$  feasible solution points in the  $n$ -dimensional space of the decision variables ( $m \geq n + 1$ , Box recommends  $m = 2n$ ) are randomly generated. The objective function is evaluated at each one of these starting points. The iterative search then proceeds by replacing point  $w$ , the one with the worst (highest) objective value in the current solution set, by a new point. The coordinates of the new solution point are calculated to be  $\alpha \geq 1$  times as far from the centroid of the remaining points as the reflection of point  $w$  in the centroid

$$X_w^{j+1} = \alpha(X_{\text{cen}}^j - X_w^j) + X_w^j \quad (10)$$

where  $X_w^j$  and  $X_w^{j+1}$  are the values of point  $w$  in iterations  $j$  and  $j + 1$ , respectively;  $X_{\text{cen}}^j$  is the centroid of the remaining points; and  $\alpha$  is a parameter. Box recommends the value 1.3.

If the new point repeats as the worst in the solution set in consecutive iterations, it is moved half the distance to the centroid of the remaining points

$$X_w^{j+1} = \frac{1}{2}(X_{\text{cen}}^j - X_w^j) \quad (11)$$

If the generated point violates any of the constraints, it is moved a small distance  $\delta$  into the feasible area for constraints involving a single variable or toward 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 should be generated such that they 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 these parameters.

## CASE STUDIES

In this section the application of the proposed calibration methodology to calibrate the microscopic traffic simulation model MITSIMLab (21) is presented for two different applications. The first is a small application with known O-D 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.

### Case Study 1: HCQS Network

The Highway Capacity and Quality of Service (HCQS) committee of the Transportation Research Board developed a case study to test various traffic simulation models (22). 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-min interval within the peak hour were available for calibration. This case study assumes perfect knowledge of O-D 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 (23) 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

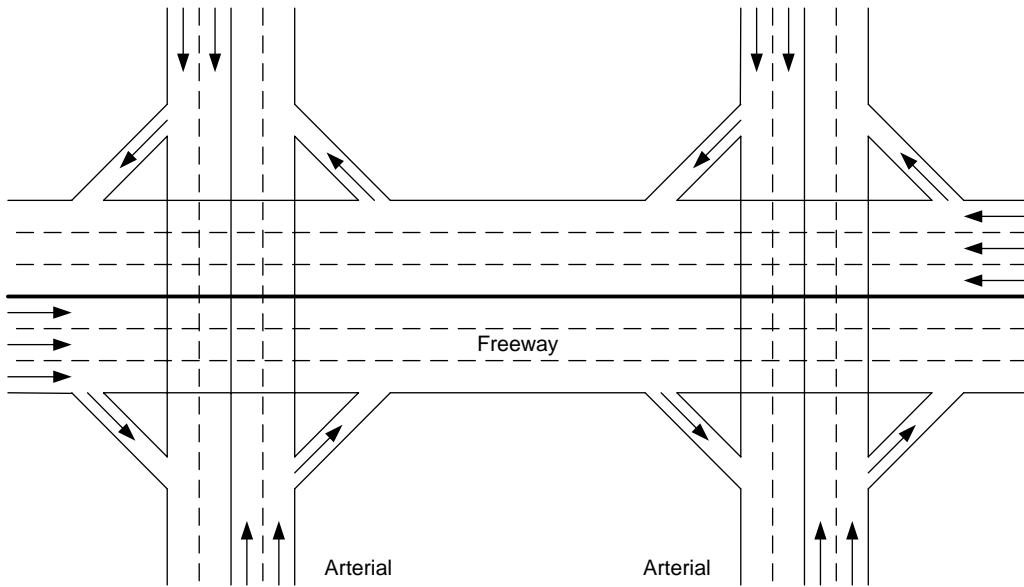


FIGURE 3 HCQS network.

on the available data, the following form of the objective function was used:

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

where  $v_i^{obs}$  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; and  $\omega$  is a weight parameter for the density measurements (the weights for speed measurements are normalized to 1).

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. For all the six points in the solution set, the values of the three parameters are within 1.9%, 6.2%, and 0.2% of the point with the lowest objective value.

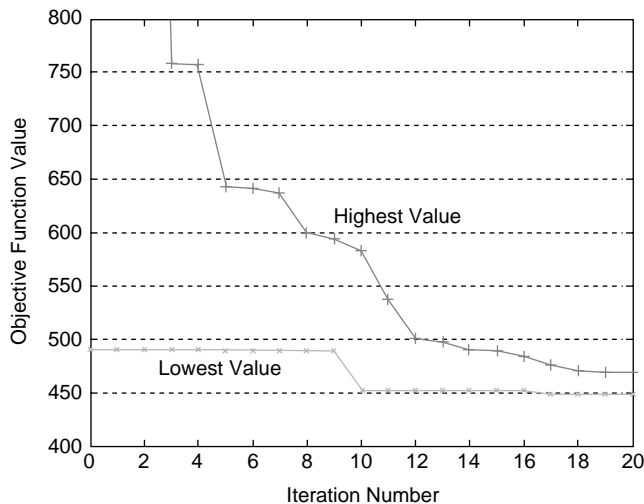


FIGURE 4 Convergence of complex algorithm.

Since O-D 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-square error (*RMSE*) of the model is 1.6 mph and the root-mean-square percent error is 7.1%. The mean absolute error (*MAE*) is 1.3 mph and the mean absolute percent error (*MAPE*) is 4.6%.

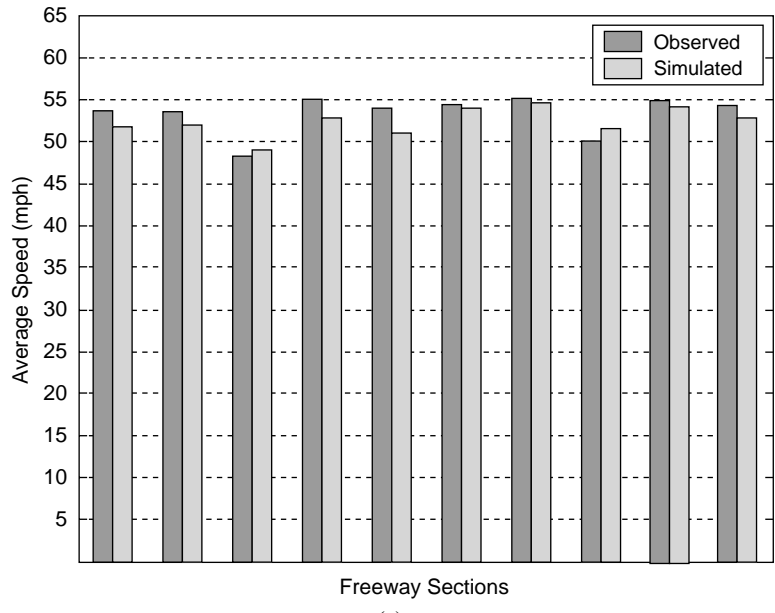
### Case Study 2: Irvine, California, Network

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 measurements collected from 68 sensor stations (30 on freeways, 38 on arterials) during 5 weekdays, and a static seed O-D matrix from a planning study.

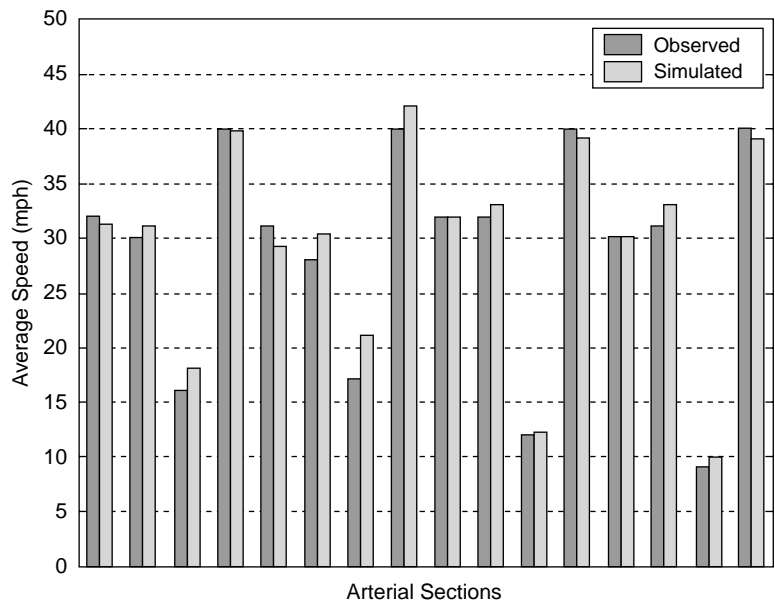
The availability of multiple days of data raises a question of whether or not the assumption that stationary steady-state conditions exist is realistic. The data show little day-to-day variability, as indicated in Figure 7, which plots the time-dependent traffic counts in the different days at two sensor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

The simulation runs were performed for the a.m. peak period. Measurements were aggregated in intervals of 15 min. The calibration included estimation of O-D 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

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



(a)



(b)

FIGURE 5 Comparison of observed and simulated speeds in HCQS network: (a) freeway sections and (b) arterial sections.



FIGURE 6 Irvine network.

where  $y_i^{obs}$  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.

**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 O-D flows by estimating O-D flows jointly with the model parameters. An optimization-based approach was used for the joint calibration. Since the calibration of the parameters depends on the estimated O-D flows and vice versa, the proposed framework is iterative. O-D estimation is based on the well-known GLS estimator. A systematic search approach based on the complex algorithm is adopted for calibration of the behavioral 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 stochastic 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 demonstrate 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 O-D estimation techniques. Further research is also required to extend the problem formulations to nonstationary cases, such as simulation of incident scenarios.

**ACKNOWLEDGMENTS**

This material is based on work supported by Jacobs Civil, Inc. and by the National Science Foundation. The authors thank Markos Papageorgiou, who suggested the use of the complex algorithm.

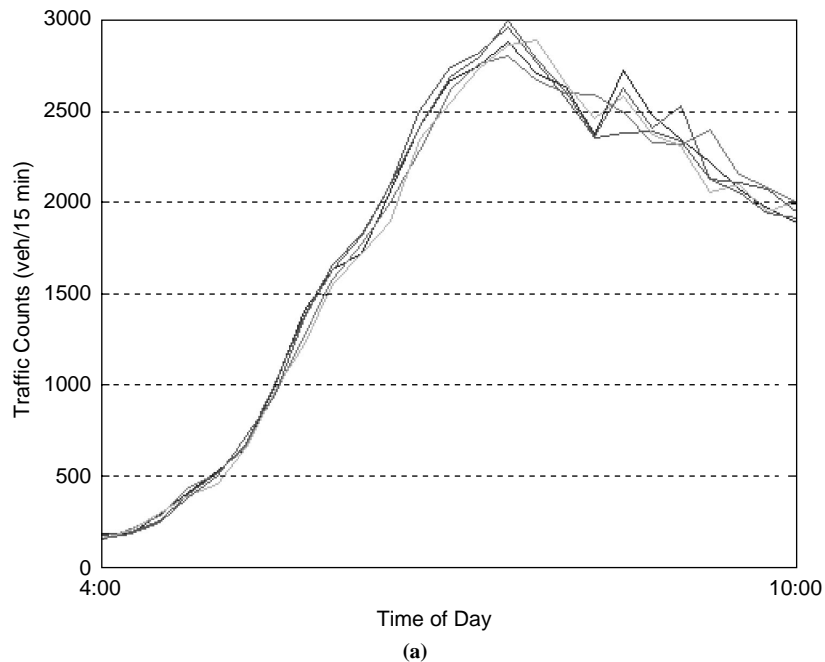


FIGURE 7 Variability of traffic counts across days: (a) Location 1. (continued on next page)



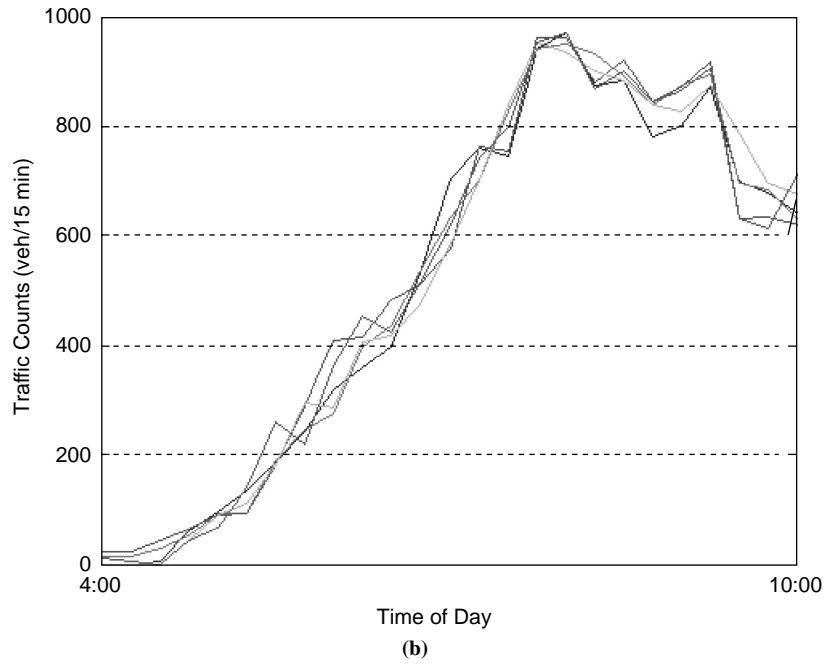


FIGURE 7 (continued) Variability of traffic counts across days: (b) Location 2.

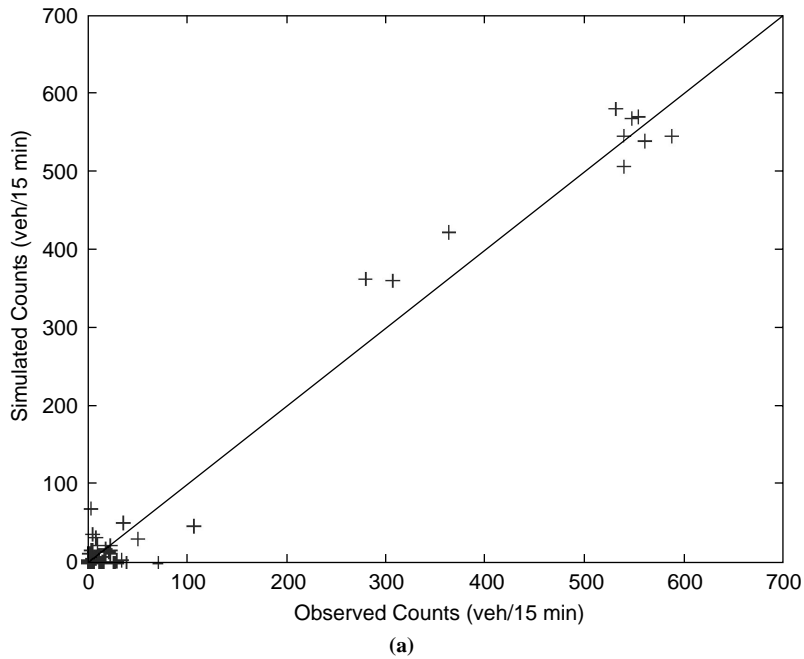


FIGURE 8 Comparison of observed and simulated counts in the Irvine network: (a) Time Interval 1.

(continued)

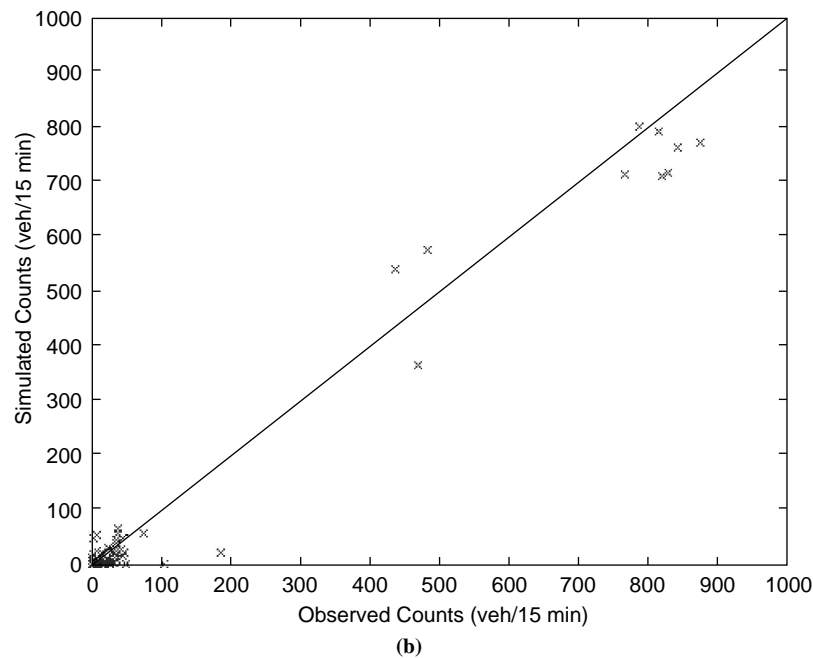


FIGURE 8 (continued) Comparison of observed and simulated counts in the Irvine network: (b) Time Interval 2.

## REFERENCES

- Daigle, G., M. Thomas, and M. Vasudevan. Field Applications of CORSIM: I-40 Freeway Design Evaluation, *Proc., 1998 Winter Simulation Conference* (D. J. Medeiros, E. F. Watson, J. S. Carson, and M. S. Manivannan, eds.), Oklahoma City, Okla., 1998, pp. 1161–1168.
- Abdulhai, B., J. B. Sheu, and W. Recker. *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. University of California, Berkeley, 1999.
- Lee, D. H., X. Yang, and P. Chandrasekar. Parameter Calibration for PARAMICS Using Genetic Algorithm. Presented at 80th Annual Meeting of the Transportation Research Board, Washington, D.C., 2001.
- Gardes, Y., A. D. May, Jr., J. W. Dahlgren, and A. Skabardonis. Freeway Calibration and Applications of PARAMICS Model. Presented at 81st Annual Meeting of the Transportation Research Board, Washington, D.C., 2002.
- Kim, K.-O., and L. R. Rilett. Simplex-Based Calibration of Traffic Microsimulation Models with Intelligent Transportation Systems Data. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1855, TRB, National Research Council, Washington, D.C., 2003, pp. 80–89.
- Park, B., and J. D. Schneeberger. Microscopic Simulation Model Calibration and Validation: Case Study of VISSIM Simulation Model for a Coordinated Actuated Signal System. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1856, TRB, National Research Council, Washington, D.C., 2003, pp. 185–192.
- Ma, T., and B. Abdulhai. Genetic Algorithm-Based Optimization Approach and Generic Tool for Calibrating Traffic Microscopic Simulation Parameters. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1800, TRB, National Research Council, Washington, D.C., 2002, pp. 6–15.
- Hourdakis, J., P. G. Michalopoulos, and J. Kottommannil. Practical Procedure for Calibrating Microscopic Traffic Simulation Models. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1852, TRB, National Research Council, Washington, D.C., 2003, pp. 130–139.
- Cremer, M., and H. Keller. A New Class of Dynamic Methods for the Identification of Origin-Destination Flows. *Transportation Research B*, Vol. 21, 1987, pp. 117–132.
- Nihan, N. L., and G. A. Davis. Recursive Estimation of Origin-Destination Matrices from Input Output Counts. *Transportation Research B*, Vol. 21, 1987, pp. 149–163.
- Cascetta, E., D. Inaudi, and G. Marquis. Dynamic Estimators of Origin-Destination Matrices Using Traffic Counts. *Transportation Science*, Vol. 27, 1993, pp. 363–373.
- Ashok, K., and M. E. Ben-Akiva. 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.), Elsevier Science, New York, 1993, pp. 165–181.
- Cascetta, E., and M. N. Postorino. Fixed Point Approaches to the Estimation of O/D Matrices Using Traffic Counts on Congested Networks. *Transportation Science*, Vol. 35, 2001, pp. 134–147.
- Liu, S. S., and J. D. Fricker. Estimation of a Trip Table and the Theta Parameter in a Stochastic Network. *Transportation Research A*, Vol. 30, 1996, pp. 287–305.
- Yang, H., Q. Meng, and M. G. H. Bell. Simultaneous Estimation of the Origin-Destination Matrices and Travel-Cost Coefficient for Congested Networks in a Stochastic User Equilibrium. *Transportation Science*, Vol. 35, 2001, pp. 107–123.
- Toledo, T., H. N. Koutsopoulos, A. Davol, M. E. Ben-Akiva, W. Burghout, I. Andréasson, T. Johansson, and C. Lundin. Calibration and Validation of Microscopic Traffic Simulation Tools: Stockholm Case Study. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1831, TRB, National Research Council, Washington, D.C., 2003, pp. 65–75.
- Bottom, J., M. E. Ben-Akiva, M. Bierlaire, I. Chabini, H. Koutsopoulos, and Q. Yang. Investigation of Route Guidance Generation Issues by Simulation with DynaMIT. In *Transportation and Traffic Theory* (A. Ceder, ed.), Pergamon, New York, 1999, pp. 577–600.
- Ashok, K., and M. E. Ben-Akiva. Alternative Approaches for Real-Time Estimation and Prediction of Time-Dependent Origin-Destination Flows. *Transportation Science*, Vol. 34, 2000, pp. 21–36.
- Cascetta, E., and S. Nguyen. A Unified Framework for Estimating or Updating Origin/Destination Matrices from Traffic Counts. *Transportation Research Part B: Methodology*, Vol. 22, 1988, pp. 437–455.
- Box, E. P. A New Method for Constrained Optimization and a Comparison with Other Methods. *Computer Journal*, Vol. 8, 1965, pp. 42–52.
- Yang, Q., and H. N. Koutsopoulos. A Microscopic Traffic Simulator for Evaluation of Dynamic Traffic Management Systems. *Transportation Research C*, Vol. 4, 1996, pp. 113–129.
- Bloomberg, L., M. Swenson, and B. Haldors. Comparison of Simulation Models and the *Highway Capacity Manual*. Presented at 82nd Annual Meeting of the Transportation Research Board, Washington, D.C., 2003.
- Ahmed, K. I. *Modeling Drivers' Acceleration and Lane Changing Behavior*. Ph.D. dissertation. Massachusetts Institute of Technology, Cambridge, 1999.