Guidelines for Calibration of Microsimulation Models

Framework and Applications

Richard Dowling, Alexander Skabardonis, John Halkias, Gene McHale, and Grant Zammit

The past few years have seen a rapid evolution in the sophistication of traffic microsimulation models and a consequent major expansion of their use in transportation engineering and planning practice. Researchers and practitioners have employed an extensive array of approaches to calibrate these models and have selected a wide range of parameters to calibrate and a broad range of acceptance criteria. A methodical, top-down approach to model calibration is outlined; it focuses the initial effort on a few key parameters that have the greatest impact on model performance and then proceeds to less critical parameters to finalize the calibration. A three-step calibration/validation process is recommended. First, the model is calibrated for capacity at the key bottlenecks in the system (the capacity calibration step). Second, the model is calibrated for traffic flows at nonbottleneck locations in the system (the route choice calibration step). Finally, the overall model performance is calibrated against field-measured system performance measures such as travel time and delay (the system performance calibration step). This three-step process is illustrated in an example application for a freeway/arterial corridor.

Microscopic simulation models simulate individual vehicle movements and their interactions on a second or subsecond basis for the purpose of assessing the traffic performance of highway and street systems. Over the past few years there has been a rapid evolution in the sophistication of microsimulation models resulting in a major expansion of their use in transportation engineering and planning practice.

Microsimulation can provide the analyst with a wealth of valuable information on the performance of the existing transportation system and potential improvements to it. However, microsimulation can also be a time-consuming and resource-intensive activity. The key to obtaining a cost-effective as well as accurate microsimulation analysis is to observe certain guiding principles for this type of analysis.

To assist users in performing successful simulation analyses, the FHWA has sponsored the development of a guide with step-by-step procedures on each of the tasks (*1*). Figure 1 illustrates the overall process for developing and applying a microsimulation model to a specific traffic analysis problem.

This paper focuses on the calibration task. The purpose is to provide a systematic framework for calibrating microsimulation models. Although there are several contributions in the literature on calibration/validation in simulation models (*2–9*), most of them focus on specific models and do not provide a systematic framework that is not model or software specific. Furthermore, this paper's approach and procedures place emphasis on the practical application of models (i.e., the task for calibrating models as part of practical operational studies rather than modeling idealized networks in a research environment).

OBJECTIVES OF CALIBRATION

Calibration is the adjustment of model parameters to improve the model's ability to reproduce local driver behavior and traffic performance characteristics. The calibration is performed after all the input data and model coding have been thoroughly checked. Calibration is performed on various components of the overall model.

The importance of calibration cannot be overemphasized. Recent tests of six different software packages found that calibration differences of 13% in the predicted freeway speeds for existing conditions ballooned to differences of 69% in the forecasted freeway speeds for future conditions (*10*).

Calibration is necessary because no single model can be expected to be equally accurate for all possible traffic conditions. Even the most detailed microsimulation model still contains only a portion of all of the variables that affect real-world traffic conditions. Because no single model can include the whole universe of variables, every model must be adapted to local conditions.

Every microsimulation software package comes complete with a set of user-adjustable parameters for the purpose of calibrating the model to local conditions. The software developers provide suggested default values for these model parameters. However, only under very rare circumstances will the model be able to produce accurate results for a specific area using only these default parameter values. The analyst should always perform some calibration tests to ensure that the coded model accurately reproduces local traffic conditions and behavior. The objective of calibration therefore is to find the set of parameter values for the model that best reproduces local traffic conditions.

The fundamental assumption of calibration is that the travel behavior models in the simulation model are essentially sound.

R. Dowling, Dowling Associates, Inc., 180 Grand Avenue, Suite 250, Oakland, CA 94612. A. Skabardonis, Institute of Transportation Studies, University of California, Berkeley, 109 McLaughlin Hall, Berkeley, CA 94720-1720. J. Halkias, FHWA, Office of Operations, Transportation Management, 400 7th Street, SW, HOTM-1, Washington, D.C. 20590. G. McHale, FHWA, Office of Operations, Research and Development, Turner-Fairbank Highway Research Center, 6300 Georgetown Pike, HRDO-4, McLean, VA 22101. G. Zammit, FHWA, 61 Forsyth Street, SW, Suite 17T26, Atlanta, GA 30303.

Transportation Research Record: Journal of the Transportation Research Board, No. 1876, TRB, National Research Council, Washington, D.C., 2004, pp. 1–9.

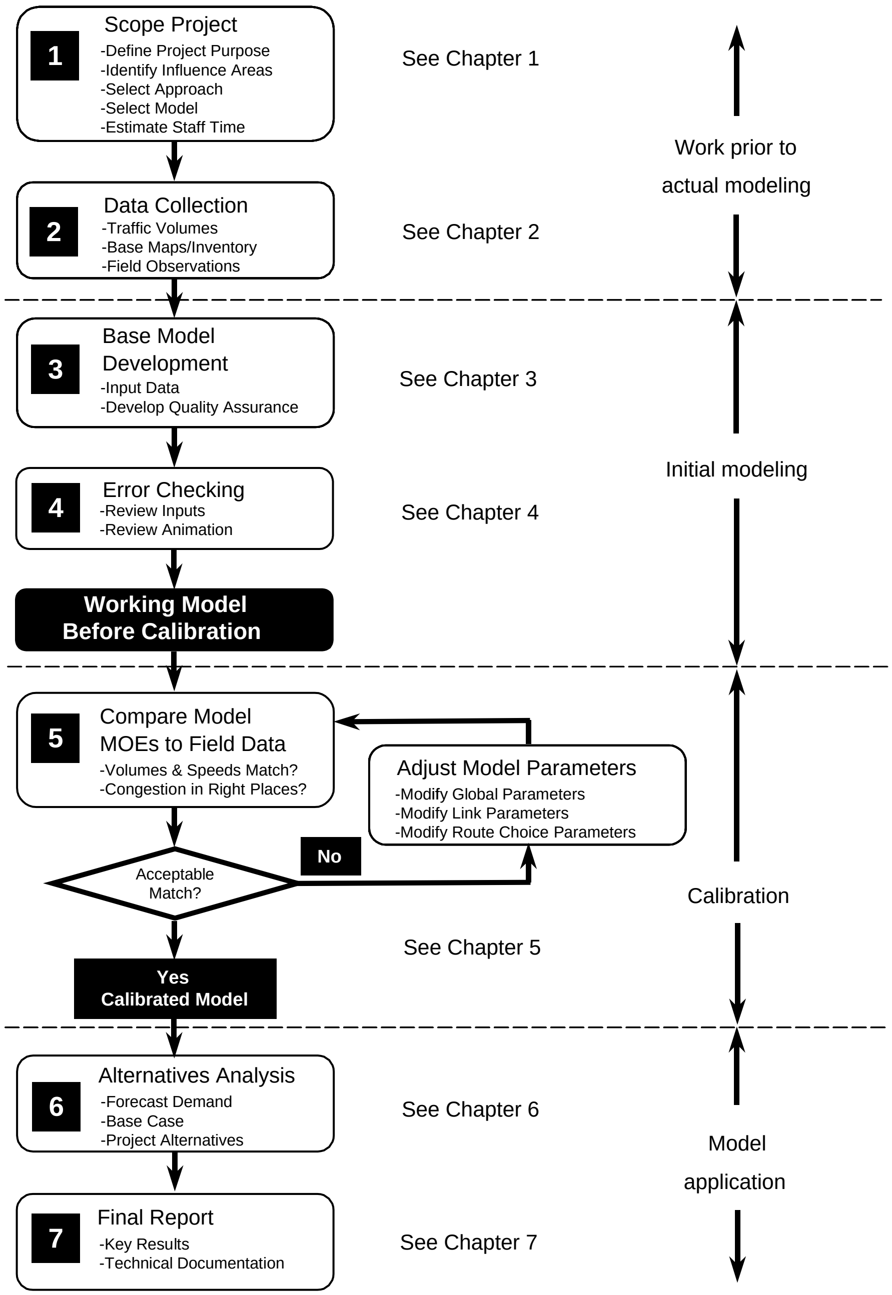

FIGURE 1 Prototypical microsimulation analysis task sequence. Adapted from Short, Elliott, Hendrickson, Inc., *Advanced CORSIM Training Manual*, Minnesota Department of Transportation.

There is no need to verify that they produce the correct delay, travel time, and density, when they are given the correct input free-flow speed and capacity for a link. The only remaining task for the analyst is to adjust these models slightly so that they correctly predict local traffic conditions.

PROPOSED APPROACH FOR CALIBRATION

Calibration involves the review and adjustment of numerous model parameters, each of which impacts the simulation results in a manner that is often highly correlated with that of the others. The analyst can easily get trapped in a never-ending circular process, fixing one problem only to find a new one pops up somewhere else. Therefore it is essential to break the calibration process into a series of logical, sequential steps, a "strategy" for calibration.

The model parameters must be divided into categories and each category must be dealt with separately. The available calibration parameters should be divided into the following two basic categories: parameters about which the analyst is reasonably certain and does not wish to adjust, and parameters about which the analyst is less certain and is willing to adjust. The analyst should avoid adjusting parameters about which the analyst has no information as to their appropriate meaning or value.

The analyst should attempt to keep the set of adjustable parameters to as small a set as possible in order to minimize the effort required to calibrate them. However, the trade-off is that more parameters allow the analyst more degrees of freedom to better fit the calibrated model to the specific location.

The set of "adjustable" parameters is then further subdivided into those that directly impact capacity (such as mean headway) and those that directly impact route choice. The capacity adjustment parameters are calibrated first; the route choice adjustment parameters are then calibrated second.

Each set of adjustable parameters can be further subdivided into those that affect the simulation on a global basis and those that affect the simulation on a more localized basis. The global parameters are calibrated first. The "local" link-specific parameters are used after the global calibration to fine-tune the results.

The following three-step strategy is recommended for calibration:

Step 1. Capacity calibration—An initial calibration is performed to identify the values for the capacity adjustment parameters that cause the model to best reproduce observed traffic capacities in the field. A global calibration is first performed, followed up by link-specific fine-tuning. The *Highway Capacity Manual* (HCM) (*11*) can be used as an alternative source of capacity target values, if field measurements are infeasible.

Step 2. Route choice calibration—If the microsimulation model includes parallel streets, then route choice will be important. In this case a second calibration process is performed, but this time with the route choice parameters. A global calibration is first performed, followed by link-specific fine-tuning.

Step 3. System performance calibration—Finally, the overall model estimates of system performance (travel times and queues) are compared to field measurements of travel times and queues. Fine-tuned adjustments are made to enable the model to better match the field measurements.

CALIBRATION DATA

Calibration data consist of measures of capacity and traffic counts and measures of system performance such as travel times, speeds, delays, and queues. Capacities can be gathered almost any time, but travel times, speeds, delays, and queues must be gathered simultaneously with the traffic counts to be useful in validating the model.

Field Measurements of Capacity

The identification of locations for field measurements of capacity will depend on the existing traffic conditions within the study area. For unsignalized facilities (freeways, rural highways, and rural roads), the analyst should identify locations where queues persist for at least 15 min and measure the flow rate at the point at which the queue discharges. This observed flow rate is measured only while an upstream queue is present. It is summed across all lanes and converted to an equivalent hourly flow rate. This is the field-measured capacity of the facility at this point.

For signalized intersections, the analyst should identify the approach legs that frequently have queues of at least 10 vehicles per lane and measure the saturation flow rate per hour per lane using the HCM procedures. Several measurements of maximum flow rates should be made in the field and averaged. Statistical procedures should be used to estimate how many measurements are required to estimate capacity within a desired confidence interval. If capacity cannot be measured in the field, then the HCM methods can be used to estimate capacity. However, these methods should not be considered as accurate as direct field measurements.

Travel Time, Delay, and Queue Data

The best source of point-to-point travel time data is "floating car runs." In this method, one or more vehicles are driven the length of the facility several times during the analysis period and the mean travel time is computed. The number of vehicle runs required to establish a mean travel time within a 95% confidence level range depends on the variability of the travel times measured in the field. Free-flow conditions may require as few as three runs to establish a reliable mean travel time. Congested conditions may require 10 or more runs.

Delay can be computed from floating car runs or from delay studies at individual intersections. Floating car runs can provide satisfactory estimates of delay along the freeway main line, but are usually too expensive to make all of the necessary additional runs to measure all of the ramp delays. Floating cars are somewhat biased estimators of intersection delay on surface streets because they reflect only those vehicles traveling a particular path through the network. For an arterial street with coordinated signal timing, the floating cars running the length of the arterial will measure delay only for the through movement with favorable progression.

Other vehicles on the arterial will experience much greater delays. This problem can be overcome by running the floating cars on different paths, but the cost is generally prohibitive. Comprehensive measures of intersection delay can be obtained from surveys of stopped delay on the approaches to an intersection. The number of stopped cars on an approach is counted at regular intervals, such as every 30 s. The number of stopped cars multiplied by the counting interval (30 s) gives the total stopped delay. Dividing the total stopped delay by the total number of vehicles that crossed the stop line (a separate count) during the survey period gives the mean stopped delay per vehicle. The stopped delay can be converted to control delay using HCM procedures.

Data from Surveillance Systems

Traffic management centers (TMCs) are a good source of simultaneous speed and flow data for urban freeways. However, the loop detectors may be subject to failures so the data must be reviewed carefully to avoid extraneous data points. Loop detectors are typically spaced ⅓ to ½ mile apart and their detection range is limited to a dozen feet. Under congested conditions, much can happen in between detectors, so the mean speeds produced by the loop detectors cannot be relied on to give system travel times under congested conditions. The loop-measured free-flow speeds may be reliable for computing facility travel times under uncongested conditions, but care should be taken when using these data. Many locations have only single loop detectors in each lane, so the free-flow speed must be estimated from an assumed mean vehicle length. The assumed mean vehicle length may be automatically calibrated by the TMC, but this calibration requires some means of identifying which data points represent free-flow speed, which data points do not, and which ones are aberrations. The decision process involves some uncertainty.

In addition, the mix of trucks and cars in the traffic stream varies by time of day; thus the same mean vehicle length cannot be used throughout the day. Therefore, loop-estimated/measured free-flow speeds should be treated with a certain amount of caution. They are precise enough for identifying the onset of congestion, but may not be reliable to the nearest 1 mph.

Calibration Targets

The objective of model calibration is to obtain the best match possible between model performance estimates and field measurements of performance. However, there is a limit to the amount of time and effort anyone can put into eliminating error in the model. There comes a point of diminishing returns where large investments in effort yield small improvements in accuracy. The analyst needs to know when to stop. This is the purpose of adopting calibration targets for the model.

Calibration targets are developed based on the minimum performance requirements for the microsimulation model, taking into consideration the available resources. The targets will vary according to the purpose for which the microsimulation model is being developed and the resources available to the analyst.

Table 1 provides an example of calibration targets that were developed by the Wisconsin Department of Transportation for their Milwaukee freeway system simulation model (*12*). These targets are based on guidelines developed in the United Kingdom (*13*).

TABLE 1 Wisconsin Department of Transportation Freeway Model Calibration Criteria

Criteria and Measures	Calibration Acceptance Targets
Hourly Flows, Model Versus Observed	
Individual link flows	
Within 15%, for 700 vph < flow < 2700 vph	>85% of cases
Within 100 vph, for flow < 700 vph	>85% of cases
Within 400 vph, for flow > 2700 vph	>85% of cases
Sum of all link flows	within 5% of sum of all link counts
GEH statistic < 5 for individual link flows[a]	>85% of cases
GEH statistic for sum of all link flows	GEH < 4 for sum of all link counts
Travel Times, Model Versus Observed	
Journey times network within 15% (or one minute, if higher)	>85% of cases
Visual Audits	
Individual link speeds	
Visually acceptable speed-flow relationship	to analyst's satisfaction
Bottlenecks	
Visually acceptable queuing	to analyst's satisfaction

[a] The GEH statistic is computed as follows:

$$\text{GEH} = \sqrt{\frac{(V - E)^2}{(E + V)/2}}$$

where E is the model estimated volume and V is the count.

PROCEDURES FOR APPLICATION

Step 1. Capacity Calibration

The capacity calibration step adjusts the global and link-specific "capacity-related" parameters in the simulation model to best replicate local field measurements of capacity. This is an important step because capacity has a significant effect on predicted system performance (delay and queues). Historically, it has been the practice to calibrate microsimulation models to all the traffic counts in the field. The majority of these counts will be at noncritical locations. The recommended strategy is to focus (at this point of the calibration process) only on the critical counts at the bottlenecks and get the model to reproduce these counts correctly. Once this is done, the rest of the counts are used later to check the route choice aspects of the model. All of this presupposes that the demands have been entered correctly and already have been checked against counts at the entry gates as part of the error-checking step of the model application (Figure 1).

The capacity calibration step consists of two phases: (a) a global calibration phase and (b) a fine-tuning phase. Global calibration is first performed to identify the appropriate network-wide value of the capacity parameter(s) that best reproduces conditions in the field. Link-specific capacity parameters are then adjusted to fine-tune the model so that it more precisely matches the field-measured capacities at each and every bottleneck. The capacity calibration procedure is as follows:

- Collect field measurements of capacity;
- Obtain model estimates of capacity;
- Select parameters to be calibrated;
- Set calibration objective function;
- Perform search for optimal parameter value; and
- Fine-tune the calibration.

Obtain Model Estimates of Capacity

Microsimulation models do not output a number called "capacity." Instead they output the number of vehicles that pass a given point. So the analyst must manipulate the input demand as necessary to create a queue upstream of the target section to be calibrated so that the model will report the maximum possible flow rate through the bottleneck.

If the model does not initially show congestion at the same bottleneck locations as exist in the field, then the demands coded in the model are temporarily increased to force the creation of congestion at those bottlenecks. These temporary increases must be removed later after the capacity calibration is complete but before the route choice calibration step.

If the model initially shows congested bottlenecks at locations that do NOT exist in the field, it will be necessary to temporarily increase the capacity at those false bottlenecks (using temporary link-specific headway adjustments). These temporary adjustments are then removed during the fine-tuning phase.

For unsignalized facilities (freeways, rural highways, and rural roads), the simulated queue should persist for at least 15 min of simulated time, across all lanes and links feeding the target section. The simulated capacity is then the mean flow rate at the target section (measured at a detector placed in the target section and summed across all lanes) averaged over the 15-min or longer period in which the queue is present. The result is then divided by the number of lanes and converted to an hourly flow rate.

For signalized intersections, the coded demand should be increased as necessary to ensure the necessary queues of at least 10 vehicles at the start of green. A detector is placed in the model at the stop line to measure the discharge headways (on a per lane basis) of the first 10 vehicles crossing the detector in each lane (the headways for the first three vehicles are discarded). The per lane headways are averaged for each lane and then averaged across lanes. The result is then converted to an hourly flow rate per lane.

Just as the field measurements of capacity were repeated several times and the results averaged, the model runs should be repeated several times and the maximum flow rate at each location averaged across the runs. The minimum required number of runs to obtain a value of capacity within a desired confidence interval can be calculated using statistical procedures.

Select Calibration Parameters

Only the model parameters that directly affect capacity are calibrated at this time. Each microsimulation software package has its own set of parameters that affect capacity, depending on the specific car-following and lane-changing logic implemented in the software. The analyst must review the software documentation and select one or two of these parameters for calibration. For example, a typical parameter is the mean headway (the minimum headway that vehicles are willing to accept in car-following situations).

Set Calibration Objective Function

It is recommended that the analyst seek to minimize the mean squared error (*MSE*) between the model estimates of maximum achievable flow rates and the field measurements of capacity. The *MSE* is the sum of the squared errors averaged over several model run repetitions. If it is difficult to perform numerous repetitive model runs with the selected software package, the analyst can use a single long (at least 1 h) simulation run for calibration. It is recommended though that once calibration is completed, several repetitive runs be made to establish a confidence range for the calibrated model.

Each set of repetitions has a single set of model parameter values (p) with different random number seeds for each repetition within the set. Some researchers have calibrated models using the percent mean squared error in order to avoid the unintended weighting effect when combining different measures of performance (such as volumes and travel time) into one measure of error. The percent *MSE* divides each square error by the field-measured value. The effect of using percent *MSE* is to place greater weight on large percentage errors rather than on large numerical errors. The simple *MSE* is recommended for calibration because it is most sensitive to large volume errors.

Select a set of model parameters p so as to minimize

$$MSE = \frac{1}{R}\sum_r (M_{ltpr} - F_l)^2$$

subject to $p_m^{\min} \le p_m \le p_m^{\max}$ for all user-adjustable model parameters p_m

where

MSE = mean squared error;
M_{ltpr} = model estimate of queue discharge flow rate (capacity) at location l and time t using parameter set p, for repetition r;
F_l = field measurement of queue discharge flow rate (capacity) at location l;
R = number of repetitive model runs with fixed parameter values p_m and different random number seeds [Because the objective is to minimize the error, dividing by a constant R (the number of repetitions) would have no effect on the result. However, R is included in the objective function to emphasize to the analyst the necessity of running the model several times with each parameter set.];
p_m = value of model parameter number m; and
$p_m^{\min}, p_m^{\max}$ = user-specified limits to the allowable range of parameter values (p_m). (Limits are necessary to avoid solutions that violate laws of physics, vehicle performance limits, and driver behavior extremes.)

Search for Optimal Parameter Value

The analyst must now find the capacity adjustment factor(s), p, that minimizes the *MSE* between the model and the field measurements of capacity. The calibration problem is a nonlinear, least squares optimization problem.

Since simulation models are complex models, it is not usually possible to formulate the models as a closed form equation for which traditional calculus techniques can be applied to find a minimum value solution. It is necessary to use some sort of search algorithm that relies on multiple operations of the simulation model, plotting of the output results as points, and searching between these points for the optimal solution. Search algorithms are required to find the optimal solution to the calibration problem.

Fine-Tune the Calibration

Once the optimal global capacity calibration parameter values have been identified, there will be still some locations where model performance deviates a great deal from field conditions. The next phase is, therefore, to fine-tune the predicted capacity to match the location-specific measurements of capacity as closely as possible.

Link-specific capacity adjustments account for the roadside factors that affect capacity but are not typically coded in the microsimulation network input data set (such as presence of on-street parking, numerous driveways, or narrow shoulders). Most simulation software packages have link-specific capacity (or headway) adjustment factors that apply only to the subject link. These capacity adjustment factors are used to fine-tune the model calibration. Link-specific adjustment factors should be used sparingly because they are not behavior based. They are fixed adjustments that will be carried through all future runs of the simulation model.

Step 2. Route Choice Calibration

Once the analyst is satisfied that the model reproduces as closely as possible field-measured capacities, the next step is then to calibrate the route choice parameters in the model to better match observed flows. The temporary demand adjustments used in the previous capacity calibration step are reversed. The model predicted volumes are then compared to field counts and the analyst adjusts the route choice algorithm parameters until the best volume fit is achieved. The specific route choice algorithm parameters will vary by software package. They generally relate to the driver's awareness of and perception and sensitivity to travel time, delay, and cost of alternate routes.

If the model network consists of only a single facility, then no route choice calibration is possible or needed. This step is skipped. This step is also skipped for microsimulation software that does not have route choice capabilities. Route choice calibration proceeds in two phases: global calibration and link-specific fine-tuning.

Global Calibration

Global calibration of route choice consists of the selection of a route choice algorithm and associated parameters. The specific parameters vary by algorithm and software package, but usually involve weights placed on the actual cost and travel time of each route. Additional parameters may relate to the familiarity of the driver with each route and the amount of error in the driver's perception of the cost and time for each route. The analyst must review the software documentation and select one or two of these parameters for calibration. Global calibration then proceeds through the same process as was used to calibrate capacity. The *MSE* between the field counts and the model volume estimates is minimized using one of the available nonlinear optimization techniques.

Fine-Tuning

Once the global calibration is completed, link-specific adjustments to cost or speed are made during the fine-tuning phase. The fine-tuning is complete when the calibration targets for volumes are met (see later section on calibration targets).

Step 3. System Performance Calibration

In this last step of the calibration, the overall traffic performance predicted by the fully functioning model is compared with field measurements of travel time, queue lengths, and duration of queuing. The analyst refines link free-flow speeds and link capacities to better match field conditions. Since changes made at this step may compromise the prior two steps of calibration, these changes should be made very sparingly.

EXAMPLE APPLICATION

Figure 2 shows the test network used to illustrate the calibration process. It consists of a freeway section with two diamond interchanges and a parallel arterial with six signalized intersections. The example problem is based in part on a sample problem used by Bloomberg et al. to compare various simulation models and HCM (*10*). Also, members of the TRB Committee on Highway Capacity and Quality of Service participated in the development of the original sample problem.

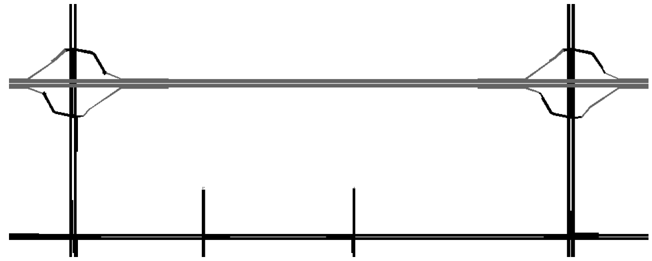

FIGURE 2 Example problem study network. For illustrative purposes this figure does not show portion of study areas lying 1.5 mi east and west of the two interchanges.

Capacity Calibration—Phase 1. Global Calibration

Collect Field Data on Capacity

The network under study is not congested. Therefore, field measurements of the capacity values on the freeway links and saturation flows at the traffic signals on the arterials cannot be obtained.

Two potential future bottleneck locations were selected on the study area network. The capacities for these potential bottleneck locations were estimated using the HCM procedures. The estimated value for the saturation flow for protected movements at signalized intersections was 1,800 vphgpl (vehicles per hour green per lane), and the capacity of freeway links was found to be 2,100 vphpl (vehicles per hour per lane).

Obtain Model Estimates of Capacity

The model estimates of capacity can be obtained from detector measurements or from the throughput values reported at the end of the simulation run. However, it is necessary to have upstream queues in order for the throughput values to represent capacity. Because of the lack of existing congestion, the existing volumes on the network links had to be artificially increased to trigger congestion upstream of the bottleneck locations (the model throughput volumes at these bottleneck locations would then be the model estimated capacities).

Select Calibration Parameters

The global parameter calibration was performed using the following parameters:

- Mean queue discharge headway at traffic signals. The default value for this parameter in the model is 1.8 s/vehicle (which is equivalent to a saturation flow of 2,000 vphgpl under ideal conditions). The queue discharge headway per vehicle type is randomly assigned by the model depending on the driver's aggressiveness factor.
- Mean headway (or car-following sensitivity factor) for the freeway links. This represents the minimum distance at which a driver is willing to follow vehicles. The values of the headway (or sensitivity factors) depend again on driver aggressiveness. Lower values would cause higher throughput.

Set Calibration Objective Function

The criterion chosen to determine the optimal parameter value was the minimization of the *MSE* between the model estimates of satu-

TABLE 2 Impact of Queue Discharge Headway on Predicted Saturation Flow

Param Value	Simulated Values of Saturation Flow (vphgpl) 1	2	3	4	5	6	7	8	9	10	Sample Mean	Sample SD
1.8	1925	1821	1812	1821	1861	1848	1861	1875	1763	1801	1839	45
1.9	1816	1857	1777	1809	1855	1815	1807	1786	1752	1784	1806	33
2.0	1772	1743	1720	1728	1799	1711	1693	1764	1673	1690	1730	40

ration flow/capacity and the "field" estimates. Table 2 shows the values of the mean queue discharge headway tested and the resulting simulated saturation flows. Ten replications for each assumed value of the parameter were performed to estimate the model and generate saturation flow within 5% of its true value. The mean values of saturation flows were used in the calculation of *MSE*. Figure 3 shows the values of *MSE* for each parameter value. The results indicate that the value of 1.9 s/vehicle produces the minimum *MSE* for this data set.

The same procedure was followed for the calibration of the "mean headway" values for freeway links. Model runs were performed assuming higher through volumes on the network links. The results indicated that the default minimum headway (car-following sensitivity factor) must be reduced by 5% (i.e., 95% of its default value in order for the model throughput to match the "observed" value of 2,100 vphpl). This value minimizes the *MSE*. Figure 4 illustrates the effect of this parameter on the freeway performance. Before calibration (Figure 4*a*) the average speed on a series of freeway links is much lower than the speeds after calibration of the headways to the HCM (Figure 4*b*).

Capacity Calibration—Phase 2. Fine-Tuning

The global calibration was performed on two links (a freeway link and a surface street link) using two parameters (mean queue discharge headway for signals and mean headway for freeways). Because just as many parameters were calibrated as links with capacity measurements, the calibration is a perfect fit for the two links. No further fine-tuning is required.

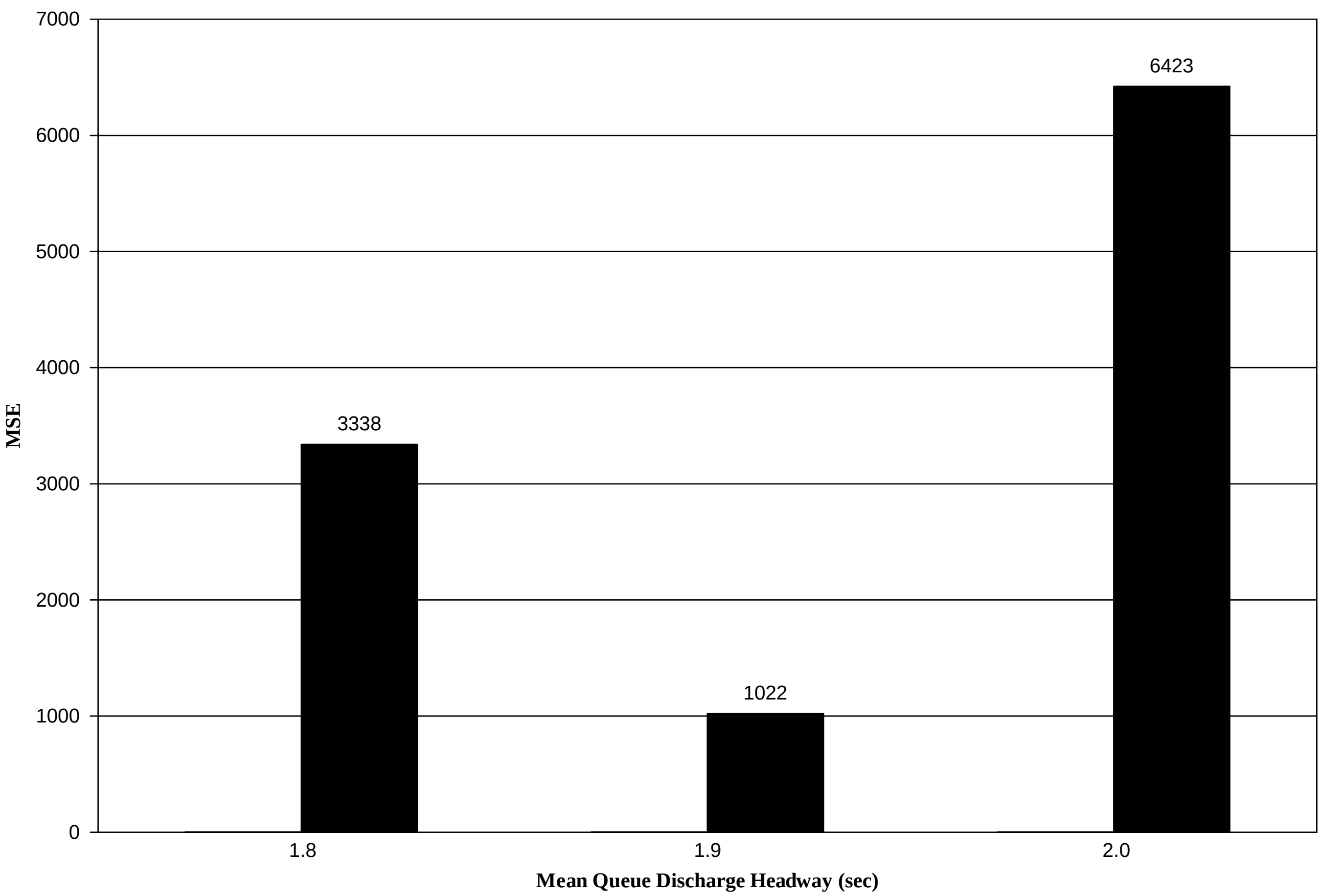

FIGURE 3 Impact of mean queue discharge headway on mean squared error of modeled saturation flow (*MSE*).

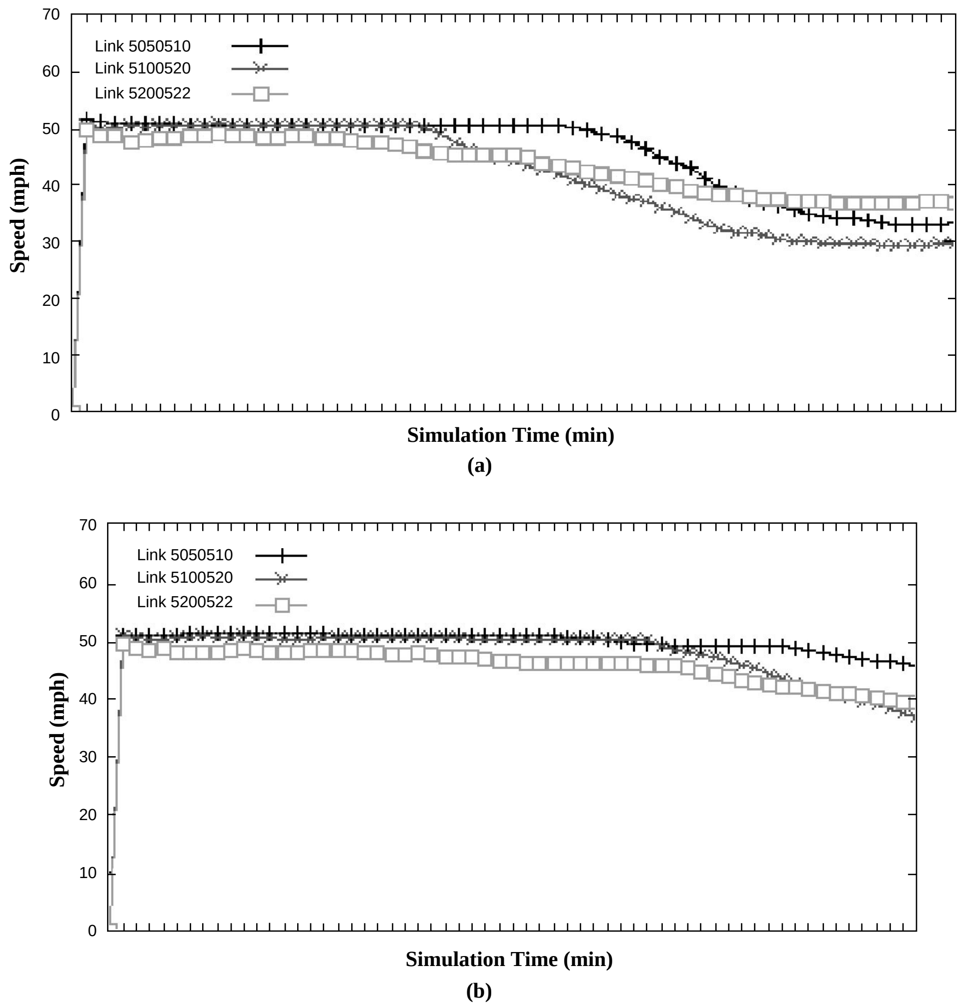

FIGURE 4 Test network: average speed by minute of simulation (three links) (*a*) before calibration and (*b*) after calibration.

System Performance Calibration

The model predictions were compared with the field data on speeds, travel times, and delays on the freeways and arterials. On the basis of this comparison, the free-flow speed distribution for the freeway was adjusted, so the model-generated free-flow speeds are within a more narrow range than the default distribution. This reflects commute traffic behavior. Furthermore, the free-flow speed on the arterial links was increased. Ten repetitions of the calibrated data set were performed and the link outputs were processed to produce performance measures to be compared with the field data. The comparison is shown in Table 3. The simulated values represent the mean value of the measures of effectiveness (MOEs) based on the 10 model runs for the third time slice of the simulated peak hour. Not shown here are comparisons with field measurement of delay because it was unclear how the field data on delay were collected. In general, users must be cautious when they compare field versus simulated delays, because the delay is typically defined differently in most software packages than the approaches commonly used for field measurements. The results indicate that the model satisfies the criteria for calibration.

CONCLUSIONS

The example application has illustrated a practical, top-down approach to microsimulation model calibration. The approach starts with the key global capacity parameters that have the greatest effect on model performance and then proceeds to parameters affecting route choice. Finally the overall model performance is compared against field measurements of system performance, such as travel time and delay.

ACKNOWLEDGMENTS

This paper is based on work supported by the FHWA. The California Department of Transportation (Caltrans) Microsimulation Applications Guidelines provided the starting point and initial testing for some of the materials contained in this paper. The authors thank Mary Rose Repine, Steven Hague, and Leo Gallagher, all of Caltrans, for their support, comments, and assistance. Adolf D. May, of the University of California, Berkeley provided an extensive review of the first draft of the background materials to this paper.

TABLE 3 Validation Results for Example Application

	Travel Time (minutes)				Speed (mph)				Density (vehicle/mile/lane)			
Freeway Segment	**Field**	**Sim.**	**Diff.**	**Abs. % Diff**	**Field**	**Sim.**	**Diff.**	**Abs. % Diff**	**Field**	**Sim.**	**Diff.**	**Abs. % Diff**
EB west of Wisconsin off	1.44	1.47	0.03	2.3%	53.5	52.42	-1.04	1.95%	38.1	36.4	-1.64	4.3%
EB Wisconsin off to on	0.22	0.22	-0.01	2.4%	53.4	51.84	-1.56	2.93%	38.9	33.0	-5.94	15.3%
EB Wisconsin on to Milwaukee off	1.00	0.95	-0.05	5.0%	48.1	51.29	3.18	6.61%	38.6	34.4	-4.23	11.0%
EB Milwaukee off to on	0.22	0.22	0.00	0.1%	55.0	52.24	-2.76	5.01%	30.3	31.4	1.13	3.7%
EB east of Milwaukee on	1.12	1.15	0.04	3.4%	53.8	51.96	-1.83	3.40%	34.0	33.9	-0.14	0.4%
WB east of Milwaukee off	1.11	1.13	0.02	2.1%	54.3	53.44	-0.85	1.57%	29.1	28.8	-0.25	0.9%
WB Milwaukee off to on	0.22	0.22	0.00	1.0%	55.0	52.65	-2.35	4.26%	26.5	27.1	0.59	2.2%
WB Milwaukee on to Wisconsin off	0.96	0.94	-0.02	2.4%	50.1	51.94	1.87	3.73%	30.4	31.0	0.61	2.0%
WB Wisconsin off to on	0.22	0.22	0.00	0.8%	55.0	52.46	-2.54	4.61%	28.5	29.6	1.09	3.8%
WB west of Wisconsin on	1.42	1.48	0.06	4.0%	54.2	52.10	-2.09	3.85%	32.1	32.9	0.78	2.4%
AVERAGE				**2.4%**				**3.79%**				**4.6%**

	Travel Time (minutes)			
Arterial Segment	**Field**	**Sim.**	**Diff.**	**Abs. % Diff**
SB Milwaukee between ramps	0.44	0.43	-0.01	3.0%
NB Milwaukee between ramps	0.47	0.48	0.01	3.1%
SB Wisconsin between ramps	0.57	0.58	0.01	2.4%
NB Wisconsin between ramps	0.74	0.66	-0.08	11.0%
AVERAGE				**4.9%**

REFERENCES

1. FHWA. *Guidelines for Applying Traffic Microsimulation Modeling Software.* Prepared by Dowling Associates, Aug. 2003.
2. Abdullai, 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 Program Research Report UCB-ITS-PRR-99-12. University of California, Berkeley, 1999.
3. Fellendorf, M., and P. Vortisch. Validation of Microscopic Traffic Flow Model VISSIM in Different Real-World Situations. Presented at 80th Annual Meeting of the Transportation Research Board, Washington, D.C., 2001.
4. Gardes, Y., A. D. May, J. Dahlgren, and A. Skabardonis. *Bay Area Simulation and Ramp Metering Study.* California PATH Research Report 1055-1425, UCB-ITS-PRR-2003-9, University of California, Berkeley, 2003.
5. Hellinga, B. R. *Requirements for the Calibration of Traffic Simulation Models.* Canadian Society of Civil Engineering, University of Waterloo, 1998.
6. 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.
7. Prassas, E. S. Comparison of Delay on Arterials by Simulation and Highway Capacity Manual Computations. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1802,* TRB, National Research Council, Washington, D.C., 2002, pp. 69–76.
8. Rakha, H., B. Hellinga, M. Van Aerde, and W. Perez. Systematic Verification, Validation, and Calibration of Traffic Simulation Models. Presented at 75th Annual Meeting of the Transportation Research Board, Washington, D.C., 1996.
9. Skabardonis, A. Simulation of Freeway Weaving Areas. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1802,* TRB, National Research Council, Washington, D.C., 2002, pp. 115–124.
10. 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.
11. *Highway Capacity Manual.* TRB, National Research Council, Washington, D.C., 2000.
12. *Freeway System Operational Assessment.* Technical Report I-33: Paramics Calibration & Validation Guidelines (Draft). Wisconsin Department of Transportation, District 2, Milwaukee, June 2002.
13. *Traffic Appraisal in Urban Areas: Highways Agency, Manual for Roads & Bridges,* Vol. 12. Department for Transport, London, May 1996. www.archive.official-documents.co.uk/document/ha/dmrb/index.htm.