Calibration Validation

A. Compare the multiple model runs to each other (confidence in the model)

It is critical that the micro-simulation results are compared to one another to ensure that no outliners appear in the results. It's important to determine the true average of the model. The model validation step does not mean that the model is representative of real world conditions but it is a critical step in determining the true average of the model.

To ensure that the results reported for the model are a true statistical representation of the average, the following formula for a 95th percent confidence interval shall be applied:

$$N = \left(2 * t_{0.025, N-1} \frac{s}{R} \right)^2$$

R = 95-percent confidence interval for the true mean

$t_{0.025,N-1}$ = Student's t-statistic for two-sided error of 2.5 percent (totals 5 percent) with N-1 degrees of freedom

s = standard deviation of about the mean for selected MOE

N = number of required simulation runs with different random seeds

If N < 10 then 10 intervals shall be run with different random seeds. The calculation should be run for one or two measures of effectiveness that are deemed most important to the outcome of the specific project. The value of the student's t statistic can be found in any statistics manual but for a data set of 10 runs, t=2.3. The calculation above is required to determine how many simulation runs are needed for the base model scenario. To be consistent throughout the process, all alternatives shall use the same number of simulation runs.

B. Show calibration targets and results for the existing model compared against the field data for the volumes, travel time, and queue lengths.

Field data on the above goals have already been collected in the data collection phase of the project and will be compared to the model results in order to calibrate the existing model. Some project may require a multi-hour analysis. The calibration criteria below should be applied to the peak.

Calibration sheets are required for review and must be presented with the deliverable of Phase 2 of the VISSIM schedule. Example calibration tables are provided.

When calibrating a micro-simulation model, the below targets shall be met. For micro-simulation models prepared for LADOTD, the following MOE's, at a minimum, shall be used as calibration goals for all models:

Criteria and Measures	*Calibration Acceptance Targets*
1) Hourly Flows, Model vs. Count Data	
a. Individual Link Flows	
i. Within 100 vehicles per hour (veh/hr) for flow < 700 veh/hr	> 85% of cases
ii. Within 15% for 700 veh/hr < flow < 2700 veh/hr	> 85% of cases
iii. Within 400 veh/hr for flows > 2700 veh/hr	> 85% of cases
iv. Sum of all link flows	Within 5% of sum of all link flows
v. GEH Statistic < 5 for Individual Link Flows*	> 85% of cases
vi. GEH Statistic for Sum of All Link Flows*	GEH < 4 for sum of all link counts
2) Travel Times, Model Versus Observed Journey Times, Network Within 15% (or 1 minute, if higher)	> 85% of cases
3) Queue Comparison, Model Versus Observed	Visual

*The GEH statistic is computed as follows:

$$GEH = \sqrt{\frac{(E-V)^2}{(E+V)/2}} \quad (4)$$

where:

E = model estimated volume
V = field count

Source: "Freeway System Operational Assessment," *Paramics Calibration and Validation Guidelines* (Draft), Technical Report I-33, Wisconsin DOT, District 2, June 2002.

Volume Calibration

The volume calibrations should not exceed 10% of the **count traffic volume** and GEH <5.

Caution: A frequent error noted is the use of the balanced traffic volume network for calibration of a VISSIM model. This is an incorrect calibration method. Calibration should not be made using the demand volume (i.e. the balanced volume network); rather they should meet the throughput measured in the field (i.e. raw data count).

Table 1: Example Volume Calibration Table

Segment	Example Volume Calibration Table						
	1.Demand Volume (vehicles)	2.Count Volume (vehicles)	3.Simulated Volume (vehicles)	4.Difference (% of 2&3)	5.Difference <10%?	6.GEH	7.GEH <5?
Corridor							
A-B	2,765	2,620	2,628	0%	Yes	0.2	Yes
B-C	4,050	3,500	4,086	-17%	No	9.5	No

Calibration threshold not met.

Travel Time/Speed Calibration

Calibration of the network using travel times or speed must report short segment data in addition to overall corridor travel time/speed. DOTD requires a ± 10 percent travel time variation for small segments (no more than 1-mile-long) and ± 5 percent travel time variation over the entire corridor analyzed. Exceptions permitted on a case-by-case basis with justification.

For facility spanning more than 1 mile, it is recommended to break the facility into segments based on obvious breakpoints (ex. between signalized intersections, or at ramps). These new smaller segments would then be calibrated at ± 10 percent variation with an overall corridor calibration of ± 5 percent.

To calibrate to travel times or speeds, average car method shall be used. Do **not** mix the calibration of travel times from average car runs with speeds collected from an outside source. Two options are available if multiple data sets are available:

1. Average the speed data with the travel time runs into one data set (i.e. convert speeds into travel time runs or vice versa and calibrate the VISSIM outputs to the average of the two).
2. Use only one data set, either travel time runs from the average car runs, or the speed data from an outside source, and keep the other data source for validation.

Table 2: Travel Time Calibration

Table 2: Travel Time Calibration							
Travel Segments	**Distance (miles)**	**Speeds (MPH)**		**Travel Times (sec)**		**Difference**	
		Field (mph)	**Simulated (mph)**	**Field (sec)**	**Simulated (sec)**	**Difference (sec)**	**Difference (%)**
Corridor	2	51	56.3	141.2	128.5	12.7	10%
A-B	1	60	59.7	60	60.3	0.3	1%
B-C	1	30	52.8	120	68.2	-51.8	-43%

Calibration threshold not met.

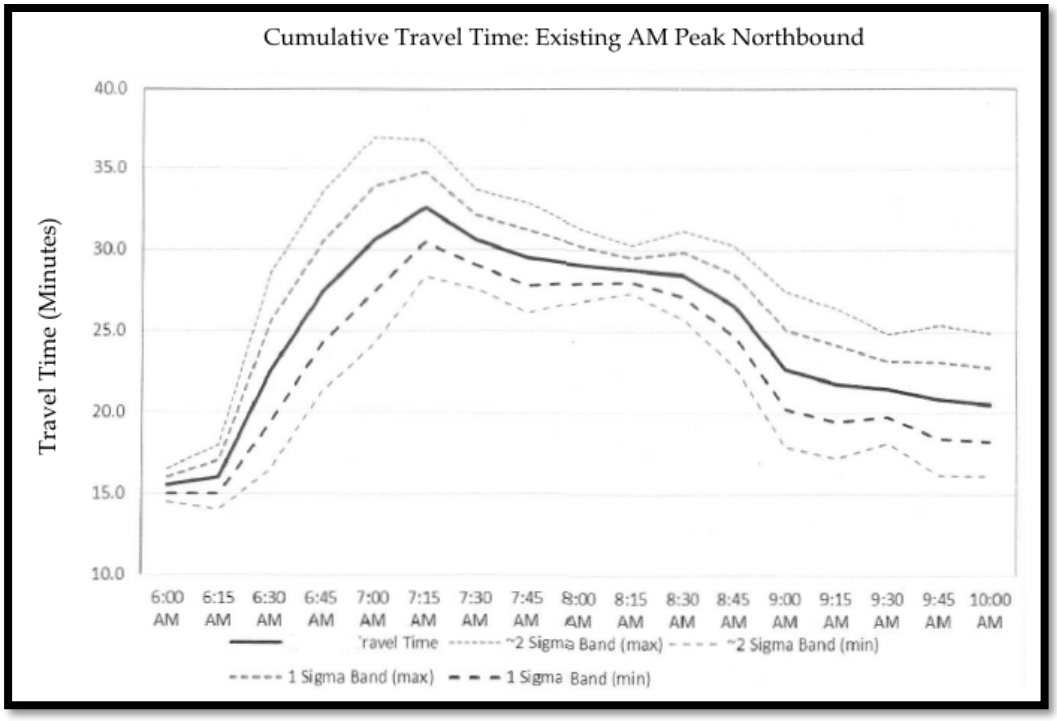

Figure 1: Example of Cumulative Travel Times

Queue Length Comparison

Queueing and bottleneck areas shall be compared on an aerial (see example below) once the other calibration targets are met. If the model does not reflect real world observations, an explanation shall be provided.

Figure 2: Example Queue Length Comparison