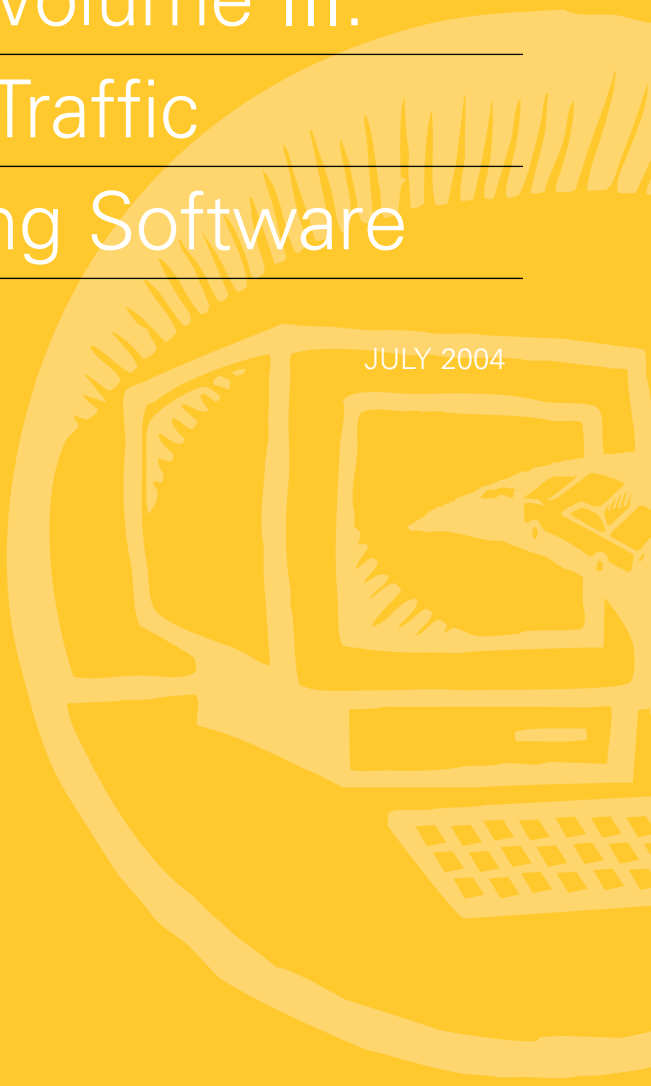
Traffic Analysis Toolbox Volume III: Guidelines for Applying Traffic Microsimulation Modeling Software

PUBLICATION NO. FHWA-HRT-04-040

JULY 2004



U.S. Department of Transportation
Federal Highway Administration

Research, Development, and Technology
Turner-Fairbank Highway Research Center
6300 Georgetown Pike
McLean, VA 22101-2296

Foreword

Traffic simulation software has become increasingly more popular as a traffic analysis tool used in transportation analyses. One reason for this increase in the use of simulation is the need to model and analyze the operation of complex transportation systems under congested conditions. Where some analytical techniques break down under these types of conditions, simulation shines. However, despite the widespread use of traffic simulation software, there are a variety of conflicting thoughts and practices on how simulation should be used.

The purpose of the *Guidelines for Applying Traffic Microsimulation Modeling Software* is to provide a recommended process for using traffic simulation software in transportation analyses. The guidelines provide the reader with a seven-step process that begins with project scope and ends with the final project report. The process is generic, in that it is independent of the specific software tool used in the analysis. In fact, the first step in the process involves picking the appropriate tool for the job at hand. It is hoped that these guidelines will assist the transportation community in creating a more consistent process in the use of traffic simulation software.

This document serves as Volume III in the Traffic Analysis Toolbox. Other volumes currently in the toolbox include: Volume I: *Traffic Analysis Tools Primer* and Volume II: *Decision Support Methodology for Selecting Traffic Analysis Tools.*

The intended audience for this report includes the simulation analyst, the reviewer of simulation analyses, and the procurer of simulation services.

Jeffery A. Lindley, P.E.
Director
Office of Transportation Management

Technical Report Documentation Page

1. Report No. FHWA-HRT-04-040	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Traffic Analysis Toolbox Volume III: Guidelines for Applying Traffic Microsimulation Software		5. Report Date June 2004 6. Performing Organization Code
7. Author(s) Richard Dowling, Alexander Skabardonis, Vassili Alexiadis		8. Performing Organization Report No.
9. Performing Organization Name and Address Dowling Associates, Inc. 180 Grand Avenue, Suite 250 Oakland, CA 94612		10. Work Unit No. 11. Contract or Grant No. DTFH61-01-C-00181
12. Sponsoring Agency Name and Address Office of Operations Federal Highway Administration 400 7th Street, S.W. Washington, D.C. 20590		13. Type of Report and Period Covered Final Report, May 2002 – August 2003 14. Sponsoring Agency Code

15. Supplementary Notes
FHWA COTR: John Halkias, Office of Transportation Management

16. Abstract

This report describes a process and acts as guidelines for the recommended use of traffic microsimulation software in transportation analyses. The seven-step process presented in these guidelines highlights the aspects of microsimulation analysis from project start to project completion. The seven steps in the process include: 1) scope project, 2) data collection, 3) base model development, 4) error checking, 5) compare model MOEs to field data (and adjust model parameters), 6) alternatives analysis, and 7) final report. Each step is described in detail and an example problem applying the process is carried through the entire document.

The appendices to report contain detailed in information covering areas such as: a) traffic microsimulation fundamentals, b) confidence intervals, c) estimation of simulation initialization period, d) simple search algorithms for calibration, e) hypothesis testing of alternatives, and f) demand constraints.

This is the third volume in a series of volumes in the Traffic Analysis Toolbox. The other volumes currently in the Traffic Analysis Toolbox are:
Volume I: Traffic Analysis Tools Primer (FHWA-HRT-04-038)
Volume II: Decision Support Methodology for Selecting Traffic Analysis Tools (FHWA-HRT-04-039)

17. Key Words	18. Distribution Statement
Traffic simulation, traffic analysis tools, highway capacity, simulation guidelines	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No of Pages	22. Price
Unclassified	Unclassified	146	

Form DOT F 1700.7 (8-72) **Reproduction of completed pages authorized**

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
°C	Celsius	1.8C+32	Fahrenheit	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380.
(Revised March 2003)

Table of Contents

Table of Contents (continued)

Table of Contents (continued)

List of Figures

List of Tables

Introduction

Microsimulation is the modeling of individual vehicle movements on a second or subsecond basis for the purpose of assessing the traffic performance of highway and street systems, transit, and pedestrians. The last few years have seen a rapid evolution in the sophistication of microsimulation models and a major expansion of their use in transportation engineering and planning practices. These guidelines provide practitioners with guidance on the appropriate application of microsimulation models to traffic analysis problems, with an overarching focus on existing and future alternatives analysis.

The use of these guidelines will aid in the consistent and reproducible application of microsimulation models and will further support the credibility of the tools of today and tomorrow. As a result, practitioners and decisionmakers will be equipped to make informed decisions that will account for current and evolving technology. Depending on the project-specific purpose, need, and scope, elements of the process described in these guidelines may be enhanced or adapted to support the analyst and the project team. It is strongly recommended that the respective stakeholders and partners consult prior to and throughout the application of any microsimulation model. This further supports the credibility of the results, recommendations, and conclusions, and minimizes the potential for unnecessary or unanticipated tasks.

Organization of Guidelines

These guidelines are organized into the following chapters and appendixes:

Introduction (this chapter) highlights the key guiding principles of microsimulation and provides an overview of these guidelines.

Chapter 1.0 addresses the management, scope, and organization of microsimulation analyses.

Chapter 2.0 discusses the steps necessary to collect and prepare input data for use in microsimulation models.

Chapter 3.0 discusses the coding of input data into the microsimulation models.

Chapter 4.0 presents error-checking methods.

Chapter 5.0 provides guidance on the calibration of microsimulation models to local traffic conditions.

Chapter 6.0 explains how to use microsimulation models for alternatives analysis.

Chapter 7.0 provides guidance on the documentation of microsimulation model analysis.

Appendix A provides an introduction to the fundamentals of microsimulation model theory.

Appendix B provides guidance on the estimation of the minimum number of microsimulation model run repetitions necessary for achieving a target confidence level and confidence interval.

Appendix C provides guidance on the estimation of the duration of the initialization (warmup) period, after which the simulation has stabilized and it is then appropriate to begin gathering performance statistics.

Appendix D provides examples of some simple manual search algorithms for optimizing parameters during calibration.

Appendix E summarizes the standard statistical tests that can be used to determine whether two alternatives result in significantly different performance. The purpose of these tests is to demonstrate that the improved performance in a particular alternative is not a result of random variation in the simulation model results.

Appendix F describes a manual method for constraining external demands to available capacity. This method is useful for adapting travel demand model forecasts for use in microsimulation models.

Appendix G provides useful references on microsimulation.

Guiding Principles of Microsimulation

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

- Use of the appropriate tool is essential. Do not use microsimulation analysis when it is not appropriate. Understand the limitations of the tool and ensure that it accurately represents the traffic operations theory. Confirm that it can be applied to support the purpose, needs, and scope of work, and can address the question that is being asked.

- *Traffic Analysis Toolbox, Volume II: Decision Support Methodology for Selecting Traffic Analysis Tools* (Federal Highway Administration (FHWA), presents a methodology for selecting the appropriate type of traffic analysis tool for the task.

- Do not use microsimulation if sufficient time and resources are not available. Misapplication can degrade credibility and can be a focus of controversy or disagreement.

- Good data are critical for good microsimulation model results.

- It is critical that the analyst calibrate any microsimulation model to local conditions.

- Output of a microsimulation model is different from that of the *Highway Capacity Manual* (HCM) (Transportation Research Board (TRB)). Definitions of key terms, such as "delay" and "queues," are different at the microscopic level of microsimulation models than at the macroscopic level typical of the HCM.

- Prior to embarking on the development of a microsimulation model, establish its scope among the partners, taking into consideration expectations, tasks, and an understanding of how the tool will support the engineering decision. Identify known limitations.

- To minimize disagreements between partners, embed interim periodic reviews at prudent milestones in the model development and calibration processes.

Additional information on traffic microsimulation fundamentals is provided in appendix A.

Terminology Used in These Guidelines

Calibration: Process where the analyst selects the model parameters that cause the model to best reproduce field-measured local traffic operations conditions.

Microsimulation: Modeling of individual vehicle movements on a second or subsecond basis for the purpose of assessing the traffic performance of highway and street systems.

Model: Specific combination of modeling software and analyst-developed input/ parameters for a specific application. A single model may be applied to the same study area for several time periods and several existing and future improvement alternatives.

Project: To reduce the chances of confusing the analysis of a project with the project itself, this report limits the use of the term "project" to the physical road improvement being studied. The evaluation of the impact of a project will be called an "analysis."

Software: Set of computer instructions for assisting the analyst in the development and application of a specific microsimulation model. Several models can be developed using a single software program. These models will share the same basic computational algorithms embedded in the software; however, they will employ different input and parameter values.

Validation: Process where the analyst checks the overall model-predicted traffic performance for a street/road system against field measurements of traffic performance, such as traffic volumes, travel times, average speeds, and average delays. Model validation is performed based on field data not used in the calibration process. This report presumes that the software developer has already completed this validation of the software and its underlying algorithms in a number of research and practical applications.

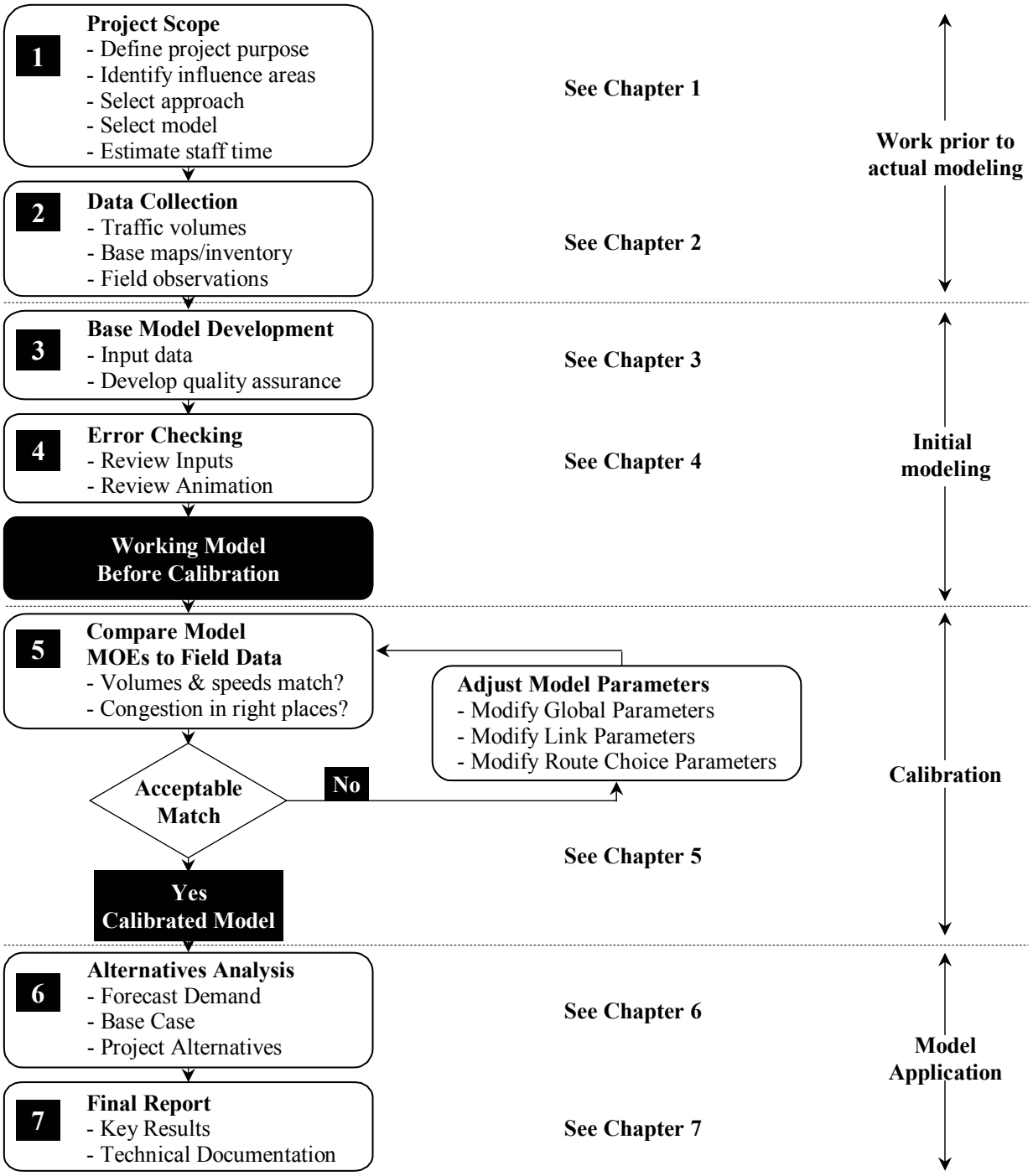
Verification: Process where the software developer and other researchers check the accuracy of the software implementation of traffic operations theory. This report provides no information on software verification procedures.

Microsimulation Model Development and Application Process

The overall process for developing and applying a microsimulation model to a specific traffic analysis problem consists of seven major tasks:

1. Identification of Study Purpose, Scope, and Approach.
2. Data Collection and Preparation.
3. Base Model Development.
4. Error Checking.
5. Calibration.
6. Alternatives Analysis.
7. Final Report and Technical Documentation.

Each task is summarized below and described in more detail in subsequent chapters. A flow chart, complementing the overall process, is presented in figure 1. It is intended to be a quick reference that will be traceable throughout the document. This report's chapters correspond to the numbering scheme in figure 1.



Developed by the FHWA Traffic Analysis Tools Team and later adapted from *Advanced Corsim Training Manual*, Short, Elliott, Hendrickson, Inc., Minnesota Department of Transportation, September 2003.

Figure 1. Microsimulation model development and application process.

To demonstrate the process, an example problem is also provided. The example problem involves the analysis of a corridor consisting of a freeway section and a parallel arterial. For simplicity, this example assumes a proactive traffic management strategy that includes ramp metering. Realizing that each project is unique, the analyst and project manager may see a need to revisit previous tasks in the process to fully address the issues that arise.

Organization and management of a microsimulation analysis require the development of a "scope" for the analysis. This scope includes identification of project objectives, available resources, assessment of verified and available tools, quality assurance plan, and identification of the appropriate tasks to complete the analysis.

Task 1: Microsimulation Analysis Organization/Scope

The key issues for the management of a microsimulation study are:

- Securing sufficient expertise to develop and/or evaluate the model.
- Providing sufficient time and resources to develop and calibrate the microsimulation model.
- Developing adequate documentation for the model development process and calibration results.

Task 2: Data Collection and Preparation

This task involves the collection and preparation of all of the data necessary for the microsimulation analysis. Microsimulation models require extensive input data, including:

- Geometry (lengths, lanes, curvature).
- Controls (signal timing, signs).
- Existing demands (turning volumes, origin-destination (O-D) table).
- Calibration data (capacities, travel times, queues).
- Transit, bicycle, and pedestrian data.

In support of tasks 4 and 5, the current and accurate data required for error checking and calibration should also be collected. While capacities can be measured at any time, it is crucial that the other calibration data (travel times, delays, and queues) be gathered simultaneously with the traffic counts.

Task 3: Base Model Development

The goal of base model development is a model that is verifiable, reproducible, and accurate. It is a complex and time-consuming task with steps that are specific to the software used to perform the microsimulation analysis. The details of model development are best covered in software-specific user's guides, and, for this reason, the development process may vary. This report provides a general outline of the model development task.

The method for developing a microsimulation model can best be thought of as the building up of several layers of the model until the model has been completed. The first layer (the link/node diagram) sets the foundation for the model. Additional data on traffic controls and link operations are then added on top of this foundation. Travel demand and traveler behavior data are then added to the basic network. Finally, the simulation run control data are input to complete the model development task. The model development process does not have to follow this order exactly; however, each of these layers is required in some form in any simulation model. The model development task should also include the development and implementation of a quality assurance/quality control (QA/QC) plan to reduce the introduction of input coding errors into the model.

Task 4: Error Checking

The error-checking task is necessary to identify and correct model coding errors so that they do not interfere with the model calibration task. Coding errors can distort the model calibration process and cause the analyst to adopt incorrect values for the calibration parameters. Error checking involves various tests of the coded network and the demand data to identify input coding errors.

Task 5: Microsimulation Model Calibration

Each microsimulation software program has a set of user-adjustable parameters that enable the practitioner to calibrate the software to better match specific local conditions. These parameter adjustments are necessary because no microsimulation model can include all of the possible factors (both onstreet and offstreet) that might affect capacity and traffic operations. The calibration process accounts for the impact of these "unmodeled" site-specific factors through the adjustment of the calibration parameters included in the software for this specific purpose.

Therefore, model calibration involves the selection of a few parameters for calibration and the repeated operation of the model to identify the best values for those parameters. This can be a time-consuming process. It should be well documented so that later reviewers of the model can understand the rationale for the various parameter changes made during calibration. For example, the car-following sensitivity factor for a specific freeway segment (link 20 to 21) has been modified to a value of 95 to match the observed average speed in this freeway segment.

The key issues in calibration are:

- Identification of necessary model calibration targets.
- Allocation of sufficient time and resources to achieve calibration targets.
- Selection of the appropriate calibration parameter values to best match locally measured street, highway, freeway, and intersection capacities.
- Selection of the calibration parameter values that best reproduce current route choice patterns.
- Calibration of the overall model against overall system performance measures, such as travel time, delay, and queues.

Task 6: Alternatives Analysis With Microsimulation Models

This is the first model application task. The calibrated microsimulation model is run several times to test various project alternatives. The first step in this task is to develop a baseline demand scenario. Then the various improvement alternatives are coded into the simulation model. The analyst then determines which performance statistics will be gathered and runs the model for each alternative to generate the necessary output. If the analyst wishes to produce HCM level-of-service (LOS) results, then sufficient time should be allowed for post-processing the model output to convert microsimulation results into HCM-compatible LOS results.

The key issues in an alternatives analysis are:

- Forecasting realistic future demands.
- Selecting the appropriate performance measures for evaluation of the alternatives.
- Accurate accounting of the full congestion-reduction benefits of each alternative.
- Properly converting the microsimulation results to HCM LOS (reporting LOS is optional).

Task 7: Final Report/Technical Documentation

This task involves summarizing the analytical results in a final report and documenting the analytical approach in a technical document. This task may also include presentation of study results to technical supervisors, elected officials, and the general public.

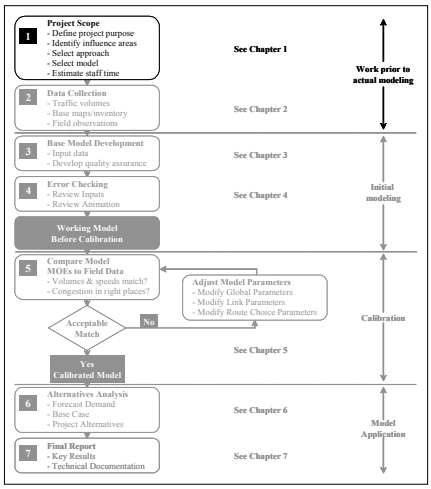
The final report presents the analytical results in a form that is readily understandable by the decisionmakers for the project. The effort involved in summarizing the results for the

final report should not be underestimated, since microsimulation models produce a wealth of numerical output that must be tabulated and summarized.

Technical documentation is important for ensuring that the decisionmakers understand the assumptions behind the results and for enabling other analysts to reproduce the results. The documentation should be sufficient so that given the same input files, another analyst can understand the calibration process and repeat the alternatives analysis.

1.0 Microsimulation Study Organization/Scope

Microsimulation can provide a wealth of information; however, it can also be a very time-consuming and resource-intensive effort. It is critical that the manager effectively coordinate the microsimulation effort to ensure a cost-effective outcome to the study. The primary component of an effective management plan is the study scope, which defines the objectives, breadth, approach, tools, resources, and time schedule for the study. This chapter presents the key components of an overall management plan for achieving a cost-effective microsimulation analysis.



1.1 Study Objectives

Before embarking on any major analytical effort, it is wise to assess exactly what it is the analyst, manager, and decisionmakers hope to accomplish. One should identify the study objectives, the breadth of the study, and the appropriate tools and analytical approach to accomplish those objectives.

Study objectives should answer the following questions:

Precise study objectives will ensure a cost-effective analysis.

It can be just as useful to identify what WILL NOT be analyzed, as it is to identify what WILL be analyzed.

- Why is the analysis needed?
- What questions should the analysis answer? (What alternatives will be analyzed?)
- Who are the intended recipients/decisionmakers for the results?

Try to avoid broad, all-encompassing study objectives. They are difficult to achieve with the limited resources normally available and they do not help focus the analysis on the top-priority needs. A great deal of study resources can be saved if the manager and the analyst can identify upfront what WILL NOT be achieved by the analysis. The objectives for the analysis should be realistic, recognizing the resources and time that may be available for their achievement.

1.2 Study Breadth

Once the study objectives have been identified, the next step is to identify the scope or breadth of the analysis—the geographic and temporal boundaries. Several questions related to the required breadth of the analysis should be answered at this time:

- What are the characteristics of the project being analyzed? How large and complex is it?

- What are the alternatives to be analyzed? How many of them are there? How large and complex are they?

- What measures of effectiveness (MOEs) will be required to evaluate the alternatives and how can they be measured in the field?

- What resources are available to the analyst?

- What are the probable geographic and temporal scopes of the impact of the project and its alternatives (now and in the future)? How far and for how many hours does the congestion extend? The geographic and temporal boundaries selected for the analysis should encompass all of the expected congestion to provide a valid basis for comparing alternatives.[1]

The geographic and temporal scopes of a microsimulation model should be sufficient to completely encompass all of the traffic congestion present in the primary influence area of the project during the target analysis period.

- What degree of precision do the decisionmakers require? Is a 10-percent error tolerable? Are hourly averages satisfactory? Will the impact of the alternatives be very similar or very different from those of the proposed project? How disaggregate an analysis is required? Is the analysis likely to produce a set of alternatives where the decisionmakers must choose between varying levels of congestion (as opposed to a situation where one or more alternatives eliminate congestion, while others do not)?

Development of a logical terminus of an improvement project versus a model has been debated since the early days of microsimulation; in the end, it is a matter of balancing study objectives and study resources. Therefore, the modeler needs to understand the operation of the improvement project to develop logical termini.

The study termini will be dependent on the "zone of influence," and the project manager will probably make that determination in consultation with the project stakeholders. Once

---

[1] The analyst should try to design the model to geographically and temporally encompass all significant congestion to ensure that the model is evaluating demands rather than capacity; however, the extent of the congestion in many urban areas and resource limitations may preclude 100 percent achievement of this goal. If this goal cannot be achieved 100 percent, then the analyst should attempt to encompass as much of the congestion as is feasible within the resource constraints and be prepared to post-process the model's results to compensate for the portion of congestion not included in the model.

that has been completed, the modeler then needs to look at the operation of the proposed facility. When determining the zone of influence, the modeler needs to understand the operational characteristics of the facility in the proposed project. This could be one intersection beyond the project terminus at one end of the project or a major generator 3.2 kilometers (km) (2 miles (mi)) away from the other end of the project. Therefore, there is no geographical guidance that can be given. However, some general guidelines can be summarized as follows:

Interstate Projects: The model network should extend up to 2.4 km (1.5 mi) from both termini of the improvement being evaluated and up to 1.6 km (1 mi) on either side of the interstate route. This will allow sufficient time and distance for the model to better develop the traffic-stream characteristics reflected in the real world.

Arterial Projects: The model network should extend at least one intersection beyond those within the boundaries of the improvement and should consider the potential impact on arterial coordination as warranted. This will capture influences such as the upstream metering of traffic and the downstream queuing of traffic.

The model study area should include areas that might be impacted by the proposed improvement strategies. For example, if an analysis is to be conducted of incident management strategies, the model study area should include the area impacted by the diverted traffic. All potential problem areas should be included in the model network. For example, if queues are identified in the network boundary areas, the analyst might need to extend the network further upstream.

Microsimulation is like a microscope. It is a very effective tool when you have it pointed at the correct subject. Tightly focused scopes of work ensure cost-effective microsimulation analyses.

A study scope that has a tight geographic focus, little tolerance for error, and little difference in the performance of the alternatives will tend to point to microsimulation as the most cost-effective approach.[2] A scope with a moderately greater geographic focus and with a timeframe 20 years in the future will tend to require a blended travel demand model and microsimulation approach. A scope that covers large geographic areas and long timeframes in the future will tend to rule out microsimulation and will instead require a combination of travel demand models and HCM analytical techniques.

[2] Continuing improvements in data collection, computer technology, and software will eventually enable microsimulation models to be applied to larger problems.

1.3 Analytical Approach Selection

Traffic Analysis Toolbox, Volume II: Decision Support Methodology for Selecting Traffic Analysis Tools[3] (a separate document) provides detailed guidance on the selection of an appropriate analytical approach. This section provides a brief summary of the key points.

Microsimulation takes more effort than macroscopic simulation, and macroscopic simulation takes more effort than HCM-type analyses. The analyst should employ only the level of effort required by the problem being studied.

Microsimulation models are data-intensive. They should only be used when sufficient resources can be made available and less data-intensive approaches cannot yield satisfactory results.

The following are several situations where microsimulation is the best technical approach for performing a traffic analysis:

- Conditions violate the basic assumptions of the other available analytical tools.
- Conditions are not covered by the other available analytical tools.
- There is testing of vehicle performance, guidance, and driver behavior modification options.

For example, most of the HCM procedures assume that the operation of one intersection or road segment is not adversely affected by conditions on the adjacent roadway.[4] Long queues from one location interfering with another location would violate this assumption. Microsimulation would be the superior analytical tool for this situation.

Because they are sensitive to different vehicle performance characteristics and differing driver behavior characteristics, microsimulation models are useful for testing intelligent transportation system (ITS) strategies designed to modify these characteristics. Traveler information systems, automated vehicle guidance, triple- or quadruple-trailer options, new weight limits, new propulsion technologies, etc., are all excellent candidates for testing with microsimulation models. The HCM procedures, for example, are not designed to be sensitive to new technology options, while microsimulation allows prediction of what the effect of new technology might be on capacity before the new technology is actually in place.

---

[3] Available at http://ops.fhwa.dot.gov/Travel/Traffic_Analysis_Tools/traffic_analysis_tools.htm.

[4] The one exception to this statement is the recently developed freeway systems analysis methodology presented in *Highway Capacity Manual 2000* (HCM 2000), which does explicitly treat oversaturated flow situations.

1.4 Analytical Tool Selection (Software)

Software selection is like picking out a suit of clothes. There is no single suit that fits all people and all uses. You must know what your software needs are and the capabilities and training of the people that you expect to use the software. Then you can select the software that best meets your needs and best fits the capabilities of the people who will use it.

The selection of the appropriate analytical tool is a key part of the study scope and is tied into the selection of the analytical approach. Some of the key criteria for software selection are technical capabilities, input/output/interfaces, user training/support, and ongoing software enhancements.[5]

Generally, it is a good idea to separate the selection of the appropriate analytical tool from the actual implementation of the tool. This can be accomplished through a selection process that is independent from any project-level analytical activities. *Traffic Analysis Toolbox, Volume II: Decision Support Methodology for Selecting Traffic Analysis Tools*, identifies several criteria that should be considered in the selection of an appropriate traffic analysis tool and helps identify the circumstances when a particular type of tool should be used. A methodology also is presented to guide the users in the selection of the appropriate tool category. This report includes worksheets that transportation professionals can use to select the appropriate tool category and assistance in identifying the most appropriate tool within the selected category. An automated tool that implements this methodology can be found at the FHWA Traffic Analysis Tools Web site at:

http://ops.fhwa.dot.gov/Travel/Traffic_Analysis_Tools/traffic_analysis_tools.htm

1.4.1 Technical Capabilities

The technical capabilities of the software are related to its ability to accurately forecast the traffic performance of the alternatives being considered in the analysis. The manager must decide if the software is capable of handling the size of problems being evaluated in the study. Are the technical analytical procedures that are incorporated into the software sensitive to the variables of concern in the study? The following is a general list of technical capabilities to be considered in the selection of software:

---

[5] Additional discussions on software selection criteria can be found in *Traffic Analysis Toolbox, Volume II: Decision Support Methodology for Selecting Traffic Analysis Tools*, and in Shaw, J.W., and D.H. Nam, "Microsimulation, Freeway System Operational Assessment, and Project Selection in Southeastern Wisconsin: Expanding the Vision" (paper presented at the TRB Annual Meeting, Washington, D.C., 2002).

- Maximum problem size (the software may be limited by the maximum number of signals that can be in a single network or the maximum number of vehicles that may be present on the network at any one time).

- Vehicle movement logic (lane changing, car following, etc.) that reflects the state of the art.

- Sensitivity to specific features of the alternatives being analyzed (such as trucks on grades, effects of horizontal curvature on speeds, or advanced traffic management techniques).

- Model parameters available for model calibration.

- Variety and extent of prior successful applications of the software program (should be considered by the manager).

1.4.2 Input/Output/Interfaces

Input, output, and the ability of the software to interface with other software that will be used in the study (such as traffic forecasting models) are other key considerations. The manager should review the ability of the software to produce reports on the MOEs needed for the study. The ability to customize output reports can also be very useful to the analyst. It is essential that the manager or analyst understand the definitions of the MOEs as defined by the software. This is because a given MOE may be calculated or defined differently by the software in comparison to how it is defined or calculated by the HCM.

1.4.3 User Training/Support

User training and support requirements are another key consideration. What kind of training and support is available? Are there other users in the area that can provide informal advice?

1.4.4 Ongoing Software Enhancements

Finally, the commitment of the software developer to ongoing enhancements ensures that the agency's investment in staff training and model development for a particular software tool will continue to pay off over the long term. Unsupported software can become unusable if improvements are made to operating systems and hardware.

1.5 Resource Requirements

The resource requirements for the development, calibration, and application of microsimulation models will vary according to the complexity of the project, its

geographic scope, temporal scope, number of alternatives, and the availability and quality of the data.[6]

In terms of training, the person responsible for the initial round of coding can be a beginner or on an intermediate level in terms of knowledge of the software. They should have supervision from an individual with more experience with the software. Error checking and calibration are best done by a person with advanced knowledge of microsimulation software and the underlying algorithms. Model documentation and public presentations can be done by a person with an intermediate level of knowledge of microsimulation software.

A prototype time schedule for the various model development, calibration, and application tasks is presented in figure 2, which shows the sequential nature of the tasks and their relative durations. Data collection, coding, error checking, and calibration are the critical tasks for completing a calibrated model. The alternatives analysis cannot be started until the calibrated model has been reviewed and accepted.

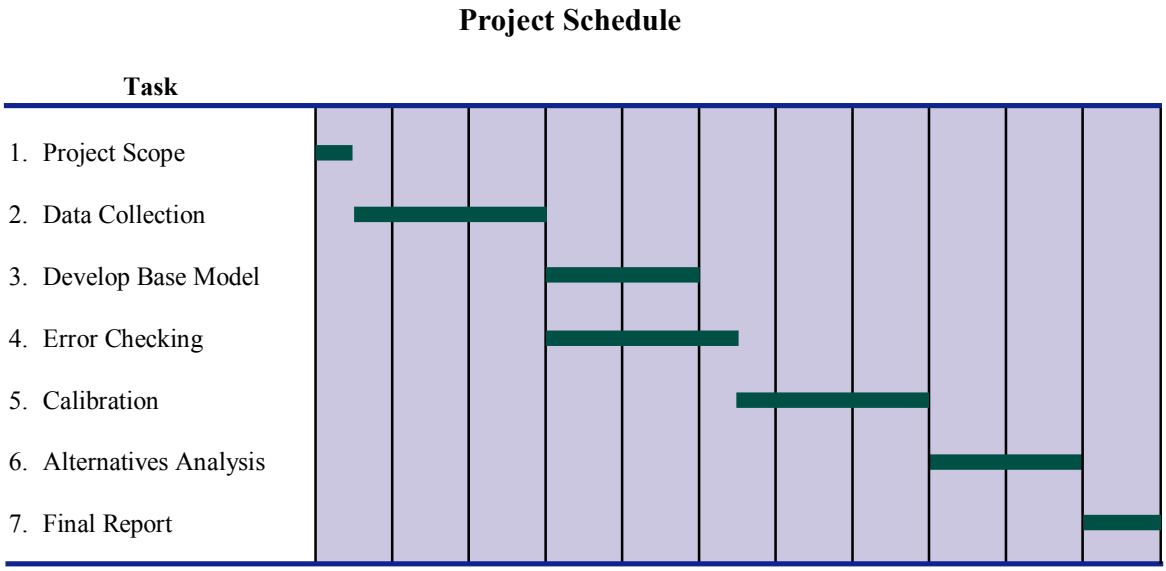


Figure 2. Prototypical microsimulation analysis task sequence.

[6] Some managers might devote about 50 percent of the budget to the tasks that lead up to and include coding of the simulation model, including data collection. Another 25 percent of the budget might go toward calibration. The remaining 25 percent might then go toward alternatives analysis and documentation. Others might divide the resources into one-third each for data collection and model coding, calibration, and alternatives analysis and documentation.

1.6 Management of a Microsimulation Study

Much of the management of a microsimulation study is the same as managing any other highway design project: establish clear objectives, define a solid scope of work and schedule, monitor milestones, and review deliverables. The key milestones and deliverables for a microsimulation study are shown in table 1:

Table 1. Milestones and deliverables for a microsimulation study.

Milestone	Deliverable	Contents
1. Study scope	1. Study scope and schedule 2. Proposed data collection plan 3. Proposed calibration plan 4. Coding quality assurance plan	Study objectives, geographic and temporal scope, alternatives, data collection plan, coding error-checking procedures, calibration plan and targets
2. Data collection	5. Data collection results report	Data collection procedures, quality assurance, summary of results
3. Model development	6. 50% coded model	Software input files
4. Error checking	7. 100% coded model	Software input files
5. Calibration	8. Calibration test results report	Calibration procedures, adjusted parameters and rationale, achievement of calibration targets
6. Alternatives analysis	9. Alternatives analysis report	Description of alternatives, analytical procedures, results
7. Final report	10. Final report	Summary tables and graphics highlighting key results
	11. Technical documentation	Compilation of prior reports documenting model development and calibration, software input files

Two problems are often encountered when managing microsimulation models developed by others:

- Insufficient managerial expertise for verifying the technical application of the model.
- Insufficient data/documentation for calibrating the model.

The study manager may choose to bring more expertise to the review of the model by forming a technical advisory panel. Furthermore, use of a panel may support a project of regional importance and detail, or address stakeholder interests regarding the acceptance of new technology. The panel may be drawn from experts at other public agencies, consultants, or from a nearby university. The experts should have had prior experience developing simulation models with the specific software being used for the particular model.

The manager (and the technical advisory panel) must have access to the input files and the software for the microsimulation model. There are several hundred parameters involved in the development and calibration of a simulation model. Consequently, it is impossible to assess the technical validity of a model based solely on its printed output and visual animation of the results. The manager must have access to the model input files so that he or she can assess the veracity of the model by reviewing the parameter values that go into the model and looking at its output.

Finally, good documentation of the model calibration process and the rationale for parameter adjustments is required so that the technical validity of the calibrated model can be assessed. A standardized format for the calibration report can expedite the review process.

1.7 Example Problem: Study Scope and Purpose

The example problem is a study of the impact of freeway ramp metering on freeway and surface-street operations. Ramp metering will be operational during the afternoon peak period on the eastbound on-ramps at two interchanges.[7]

Study Objectives: To quantify the traffic operation benefits and the impact of the proposed afternoon peak-period ramp metering project on both the freeway and nearby surface streets. The information will be provided to technical people at both the State department of transportation (DOT) and the city public works department.

Study Breadth (Geographic): The ramp metering project is expected to impact freeway and surface-street operations several miles upstream and downstream of the two metered on-ramps. However, the most significant impact is expected in the immediate vicinity of the meters (the two interchanges, the closest parallel arterial streets, and the cross-connector streets between the interchanges and the parallel arterial). If the negative impact

[7] This example problem is part of a larger project involving metering of several miles of freeway. The study area in the example problem was selected to illustrate the concepts and procedures in the microsimulation guide.

of the ramp meters is acceptable in the immediate vicinity of the project, then the impact farther away should be of a lower magnitude and, therefore, also acceptable.[8]

There is no parallel arterial street north of the freeway, so only the parallel arterial on the south side needs to be studied. Figure 3 illustrates the selected study area. There are six signalized intersections along the study section of the Green Bay Avenue parallel arterial. The signals are operating on a common 90-second (s) cycle length. The signals at the ramp junctions on Wisconsin and Milwaukee Avenues also operate on a 90-s cycle length.

The freeway and the adjacent arterial are currently not congested during the afternoon peak period and, consequently, the study area only needs to be enlarged to encompass projected future congestion. The freeway and the arterial are modeled 1.6 km (1 mi) east and west of the two interchanges to include possible future congestion under the "no meter" and "meter" alternatives.

Study Breadth (Temporal): Since there is no existing congestion on the freeway and surface streets, and the ramp metering will occur only in the afternoon peak hour, the afternoon peak hour is selected as the temporal boundaries for the analysis.[9]

Analytical Approach: There are concerns that: (1) freeway traffic may be diverted to city streets causing unacceptable congestion, and (2) traffic at metered on-ramps may backup and adversely impact surface-street operations. A regional travel demand model would probably not be adequate to estimate traffic diversion caused by a regional ramp metering system because it could not model site- and time-specific traffic queues accurately and would not be able to predict the impact of the meters on surface-street operations. HCM methods could estimate the capacity impact of ramp metering; however, because these methods are not well adapted to estimating the impact of downstream congestion on upstream traffic operations, they are not sufficient by themselves for the analysis of the impact of ramp meters on surface-street operations. Microsimulation would probably assist the analyst in predicting traffic diversions between the freeway and surface streets, and would be the appropriate analytical tool for this problem.

Analytical Tool: The analyst selects a microsimulation tool that either incorporates a procedure for the rerouting of freeway traffic in response to ramp metering, or the analyst

---

[8] Of course, if the analyst does not believe this to be true, then the study area should be expanded accordingly.

[9] If existing or future congestion on either the freeway or arterial was expected to last more than 1 h, then the analysis period would be extended to encompass all of the existing and future congestion in the study area.

supplements the microsimulation tool with a separate travel demand model analysis to predict the rerouting of traffic.[10]

Resource Requirements and Schedule: The resource requirements and schedule are estimated at this time to ensure that the project can be completed with the selected approach to the satisfaction of the decisionmakers. The details of the resources and scheduling are not discussed here because they are specific to the resources available to the analyst and the time available for completion.

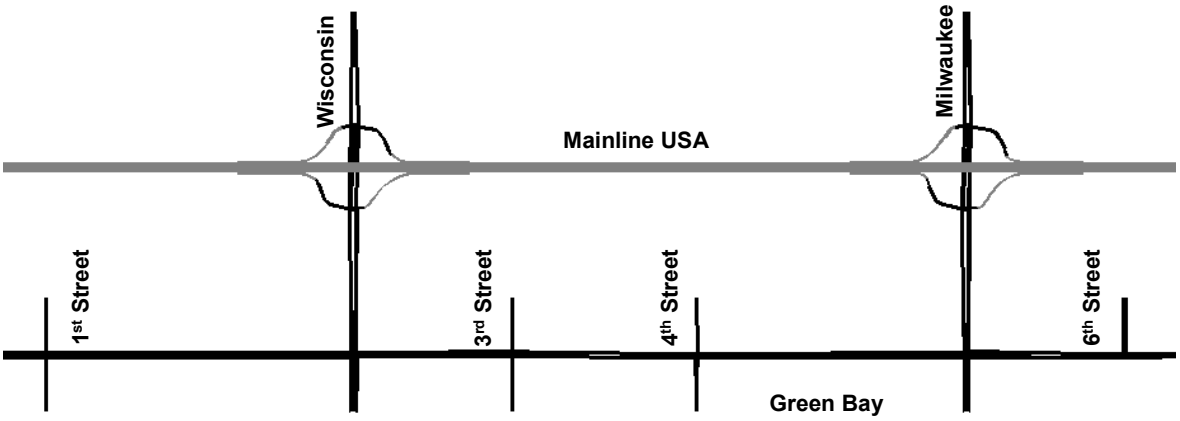
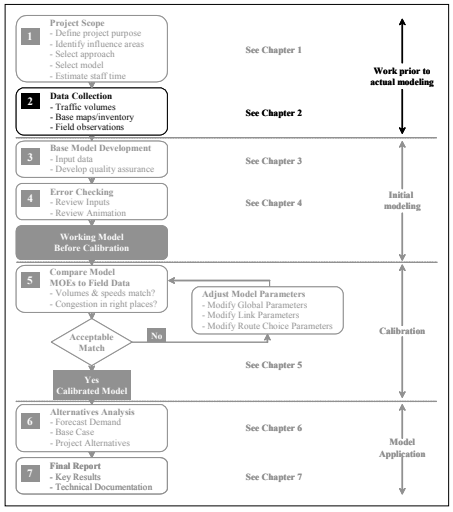


Figure 3. Example problem study network.

[10]For this example problem, a microsimulation tool without rerouting capabilities was selected. It was supplemented with estimates of traffic diversions from a regional travel demand model.

2.0 Data Collection/Preparation



This chapter provides guidance on the identification, collection, and preparation of the data sets needed to develop a microsimulation model for a specific project analysis, and the data needed to evaluate the calibration and fidelity of the model to real-world conditions present in the project analysis study area. There are agency-specific techniques and guidance documents that focus on data collection, which should be used to support project-specific needs.

A selection of general guides on the collection of traffic data includes:

Introduction to Traffic Engineering: A Manual for Data Collection and Analysis, T.R. Currin, Brooks/Cole, 2001, 140 pp., ISBN No. 0-534378-67-6.

Manual of Transportation Engineering Studies, H. Douglas Robertson, Joseph E. Hummer, and Donna C. Nelson, Institute of Transportation Engineers, Washington, DC, 1994, ISBN No. 0-13-097569-9.

Highway Capacity Manual 2000, TRB, 2000, 1200 pp., ISBN No. 0-309067-46-4.

These sources should be consulted regarding appropriate data collection methods (they are not all-inclusive on the subject of data collection). The discussion in this chapter focuses on data requirements, potential data sources, and the proper preparation of data for use in microsimulation analysis.

If the amount of available data does not adequately support the project objectives and scope identified in task 1, then the project team should return to task 1 and redefine the objectives and scope so that they will be sufficiently supported by the available data.

2.1 Required Data

The precise input data required by a microsimulation model will vary by software and the specific modeling application as defined by the study objectives and scope. Most microsimulation analytical studies will require the following types of input data:[11]

[11] The forecasting of future demands (turning volumes, O-D table) is discussed later under the Alternatives Analysis task.

- Road geometry (lengths, lanes, curvature).
- Traffic controls (signal timing, signs).
- Demands (entry volumes, turning volumes, O-D table).
- Calibration data (traffic counts and performance data such as speed, queues).

In addition to the above basic input data, microsimulation models also require data on vehicle and driver characteristics (vehicle length, maximum acceleration rate, driver aggressiveness, etc.). Because these data can be difficult to measure in the field, it is often supplied with the software in the form of various default values.

Each microsimulation model will also require various control parameters that specify how the model conducts the simulation. The user's guide for the specific simulation software should be consulted for a complete list of input requirements. The discussion below describes only the most basic data requirements shared by the majority of microsimulation model software.

2.1.1 Geometric Data

The basic geometric data required by most models consist of the number of lanes, length, and free-flow speed.[12] For intersections, the necessary geometric data may also include the designated turn lanes and their vehicle storage lengths. These data can usually be obtained from construction drawings, field surveys, geographical information system (GIS) files, or aerial photographs.

2.1.2 Control Data

Control data consist of the locations of traffic control devices and signal-timing settings.[13] These data can best be obtained from the files of the agencies operating the traffic controls or from field inspection.

2.1.3 Demand Data

The basic travel demand data required for most models consist of entry volumes (traffic entering the study area) and turning movements at intersections within the study area.

---

[12]Some microsimulation models may allow (or require) the analyst to input additional geometric data related to the grades, horizontal curvature, load limits, height limits, shoulders, onstreet parking, pavement condition, etc.

[13]Some models may allow the inclusion of advanced traffic control features. Some models require the equivalent fixed-time input for traffic-actuated signals. Others can work with both fixed-time and actuated-controller settings.

Some models require one or more vehicular O-D tables, which enable the modeling of route diversions. Procedures exist in many demand modeling software and some microsimulation software for estimating O-D tables from traffic counts.

Count Locations and Duration

Traffic counts should be conducted at key locations within the microsimulation model study area for the duration of the proposed simulation analytical period. The counts should ideally be aggregated to no longer than 15-minute (min) time periods; however, alternative aggregations can be used if dictated by circumstances.[14]

If congestion is present at a count location (or upstream of it), care should be taken to ensure that the count measures demand and not capacity. The count period should ideally start before the onset of congestion and end after the dissipation of all congestion to ensure that all queued demand is eventually included in the count.

The counts should be conducted simultaneously if resources permit so that all count information is consistent with a single simulation period. Often, resources do not permit this for the larger simulation areas, so the analyst must establish one or more control stations where a continuous count is maintained over the length of the data collection period. The analyst then uses the control station counts to adjust the counts collected over several days into a single consistent set of counts representative of a single typical day within the study area.

Estimating Origin-Destination (O-D) Trip Tables

For some simulation software, the counts must be converted into an estimate of existing O-D trip patterns. Other software programs can work with either turning-movement counts or an O-D table. An O-D table is required if it is desirable to model route choice shifts within the microsimulation model.

Local metropolitan planning organization (MPO) travel demand models can provide O-D data; however, these data sets are generally limited to the nearest decennial census year and the zone system is usually too macroscopic for microsimulation. The analyst must usually estimate the existing O-D table from the MPO O-D data in combination with other data sources, such as traffic counts. This process will probably require consideration of O-D pattern changes resulting from the time of day, especially for simulations that cover an extended period of time throughout the day.

A license plate matching survey is the most accurate method for measuring existing O-D data. The analyst establishes checkpoints within and on the periphery of the study area

---

[14] Project constraints, traffic counter limitations, or other considerations (such as a long simulation period) may require that counts be aggregated to longer or shorter periods.

and notes the license plate numbers of all vehicles passing by each checkpoint. A matching program is then used to determine how many vehicles traveled between each pair of checkpoints. However, license plate surveys can be quite expensive. For this reason, the estimation of the O-D table from traffic counts is often selected.[15]

Vehicle Characteristics

The vehicle characteristics typically include vehicle mix, vehicle dimensions, and vehicle performance characteristics (maximum acceleration, etc.).[16]

Vehicle Mix: The vehicle mix is defined by the analyst, often in terms of the percentage of total vehicles generated in the O-D process. Typical vehicle types in the vehicle mix might be passenger cars, single-unit trucks, semi-trailer trucks, and buses.

Default percentages are usually included in most software programs; however, the vehicle mix is highly localized and national default values will rarely be valid for specific locations. For example, the percentage of trucks in the vehicle mix can vary from a low of 2 percent on urban streets during rush hour to a high of 40 percent of daily weekday traffic on an intercity interstate freeway.

It is recommended that the analyst obtain one or more vehicle classification studies for the study area for the time period being analyzed. Vehicle classification studies can often be obtained from nearby truck weigh station locations.

Vehicle Dimensions and Performance: The analyst should attempt to obtain the vehicle fleet data (vehicle mix, dimensions, and performance) from the local State DOT or air quality management agency. National data can be obtained from the Motor and Equipment Manufacturers Association (MEMA), various car manufacturers, FHWA, and the U.S. Environmental Protection Agency (EPA). In the absence of data from these sources, the analyst may use the defaults shown in table 2.

---

[15] There are many references in the literature on the estimation of O-D volumes from traffic counts. Appendix A, chapter 29, HCM 2000 provides a couple of simple O-D estimation algorithms. Most travel demand software and some microsimulation software have O-D estimation modules available.

[16] The software-supplied default vehicle mix, dimensions, and performance characteristics should be reviewed to ensure that they are representative of local vehicle fleet data, especially for simulation software developed outside the United States.

Table 2. Vehicle characteristic defaults that can be used in the absence of better data.

Vehicle Type	Length (ft)	Maximum Speed (mi/h)	Maximum Accel. (ft/s^2)	Maximum Decel. (ft/s^2)	Jerk (ft/s^3)
Passenger car	14	75	10	15	7
Single-unit truck	35	75	5	15	7
Semi-trailer truck	53	67	3	15	7
Double-bottom trailer truck	64	61	2	15	7
Bus	40	65	5	15	7

1 ft = 0.305 m, 1 mi/h = 1.61 km/h, 1 ft/s^2 = 0.305 m/s^2, 1 ft/s^3 = 0.305 m/s^3

Sources: CORSIM and SimTraffic™ technical documentation

Notes:
Maximum Speed: Maximum sustained speed on level grade in miles per hour (mi/h)
Maximum Accel.: Maximum acceleration rate in feet per second squared (ft/s^2) when starting from zero speed
Maximum Decel.: Maximum braking rate in ft/s^2 (vehicles can actually stop faster than this; however, this is a mean comfort-based maximum)
Jerk: Maximum rate of change in acceleration rates in feet per second cubed (ft/s^3)

2.2 Calibration Data

System performance data (travel times, delays, queues, speeds) must be gathered simultaneously with the traffic counts.

Calibration data consist of measures of capacity, traffic counts, and measures of system performance such as travel times, speeds, delays, and queues. Capacities can be gathered independently of the traffic counts (except during adverse weather or lighting conditions); however, travel times, speeds, delays, and queue lengths must be gathered simultaneously with the traffic counts to be useful in calibrating the model.[17] If there are one or more continuous counting stations in the study area, it may be possible to adjust the count data to match the conditions present when the calibration data were collected; however, this introduces the potential for

[17] It is not reasonable to expect the simulation model to reproduce observed speeds, delays, and queues if the model is using traffic counts of demand for a different day or time period than when the system performance data were gathered.

additional error in the calibration data and weakens the strength of the conclusions that can be drawn from the model calibration task.

Finally, the analyst should verify that the documented signal-timing plans coincide with those operating in the field. This will confirm any modifications resulting from a signal retiming program.

2.2.1 Field Inspection

It is extremely valuable to observe existing operations in the field during the time period being simulated. Simple visual inspection can identify behavior not apparent in counts and floating car runs. Video images may be useful; however, they may not focus on the upstream conditions causing the observed behavior, which is why a field visit during peak conditions is always important. A field inspection is also valuable for aiding the modeler in identifying potential errors in data collection.

2.2.2 Travel Time 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 analytical period and the mean travel time is computed. The number of vehicle runs required to establish a mean travel time within a 95-percent confidence level 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.

The minimum number of floating car runs needed to determine the mean travel time within a desired 95-percent confidence interval depends on the width of the interval that is acceptable to the analyst. If the analyst wishes to calibrate the model to a very tight tolerance, then a very small interval will be desirable and a large number of floating car runs will be required. The analyst might aim for a confidence interval of ± 10 percent of the mean travel time. Thus, if the mean travel time were 10 min, the target 95-percent confidence interval would be 2 min. The number of required floating car runs is obtained from equation 1:

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

where:

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 (for four runs, t = 3.2; for six runs, t = 2.6; for 10 runs, t = 2.3)

(Note: There is one less degree of freedom than car runs when looking up the appropriate value of t in the statistical tables.)

s = standard deviation of the floating car runs

N = number of required floating car runs

For example, if the floating car runs showed a standard deviation of 1.0 min, a minimum of seven floating car runs would be required to achieve a target 95-percent confidence interval of 2.0 min (± 1.0 min) for the mean travel time.

The analyst is advised that the standard deviation is unknown prior to runs being conducted. In addition, the standard deviation is typically higher under more congested conditions.

2.2.3 Point Speed Data

Traffic management centers (TMCs) are a good source of simultaneous speed and flow data for urban freeways. Loop detectors, though, may be subject to failure, so the data must be reviewed carefully to avoid extraneous data points. Loop detectors are typically spaced 0.5 to 0.8 km (0.3 to 0.5 mi) apart and their detection range is limited to 3.7 meters (m) (12 feet (ft)). Under congested conditions, much can happen 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; however, 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; however, this calibration requires some method 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. Overall, 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; however, they may not be an accurate measure of speed.

2.2.4 Capacity and Saturation Flow Data

Capacity and saturation flow data are particularly valuable calibration data since they determine when the system goes from uncongested to congested conditions:

- Capacity can be measured in the field on any street segment immediately downstream of a queue of vehicles. The queue should ideally last for a full hour; however,

reasonable estimates of capacity can be obtained if the queue lasts only 0.5 hour (h). The analyst would simply count the vehicles passing a point on the downstream segment for 1 h (or for a lesser time period if the queue does not persist for a full hour) to obtain the segment capacity.

- Saturation flow rate is defined as "the equivalent hourly rate at which previously queued vehicles can traverse an intersection approach under prevailing conditions, assuming that the green signal is available at all times and no lost times are experienced, in vehicles per hour or vehicles per hour per lane."[18] The saturation flow rate should be measured (using procedures specified in the HCM) at all signalized intersections that are operating at or more than 90 percent of their existing capacity. At these locations, the estimation of saturation flow and, therefore, capacity will critically affect the predicted operation of the signal. Thus, it is cost-effective to accurately measure the saturation flow and, therefore, capacity at these intersections.

2.2.5 Delay and Queue Data

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 mainline; however, they 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 since 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; however, the cost may be prohibitive.

Comprehensive measures of intersection delay can be obtained from surveys of stopped delay on the approaches to an intersection (see the HCM for the procedure). 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 the control delay using the procedure in appendix A of chapter 16 in the HCM.

2.3 Data Preparation/Quality Assurance

Data preparation consists of review, error checking, and the reduction of the data collected in the field. Data reduction is already well covered in other manuals on data

[18]HCM 2000.

collection. This section consequently focuses on review and error checking of the data. The following checks of the data should be made during the data preparation step:

- Geometric and control data should be reviewed for apparent violations of design standards and/or traffic engineering practices. Sudden breaks in geometric continuity (such as a short block of a two-lane street sandwiched in between long stretches of a four-lane street) may also be worth checking with people who are knowledgeable about local conditions. Breaks in continuity and violations of design standards may be indicative of data collection errors.

- Internal consistency of counts should be reviewed. Upstream counts should be compared to downstream counts. Unexplained large jumps or drops in the counts should be reconciled.[19]

- Floating car run results should be reviewed for realistic segment speeds.

- Counts of capacity and saturation flow should be compared to the HCM estimates for these values. Large differences between field measurements and the HCM warrant double-checking the field measurements and the HCM computations.[20]

2.4 Reconciliation of Traffic Counts

Inevitably, there will be traffic counts at two or more nearby adjacent locations that do not match. This may be a result of counting errors, counting on different days (counts typically vary by 10 percent or more on a daily basis), major traffic sources (or sinks) between the two locations, or queuing between the two locations. In the case of a freeway, a discrepancy between the total traffic entering the freeway and the total exiting it may be caused by storage or discharge of some of the vehicles in growing or shrinking queues on the freeway.

The analyst must review the counts and determine (based on local knowledge and field observations) the probable causes of the discrepancies. Counting errors and counts made

---

[19] There is no guidance on precisely what constitutes a "large" difference in counts. The analyst might consider investigating any differences of greater than 10 percent between upstream and downstream counts for locations where no known traffic sources or sinks (such as driveways) exist between the count locations. Larger differences are acceptable where driveways and parking lots could potentially explain the count differences.

[20] There is no guidance on precisely how large differences can be between the HCM and field measurements before the field measurements may be suspect. The analyst might consider 25-percent differences to be within the normal range of accuracy of the HCM and take a second look at the calculations and field measurements if the differences are greater than that.

on different days are treated differently than counting differences caused by midblock sources/sinks or midblock queuing.

Discrepancies in the counts resulting from counting errors or counts made on different days must be reconciled before proceeding to the model development task. Inconsistent counts make error checking and model calibration much more difficult. Differing counts for the same location should be normalized or averaged assuming that they are reasonable. This is especially true for entry volumes into the model network. Intersection turning volumes should be expressed as percentages based on an average of the counts observed for that location. This will greatly assist with calibrating the model later.

Differences in counts caused by midblock sources (such as a parking lot) need not be reconciled; however, they must be dealt with by coding midblock sources and sinks in the simulation model during the model development task.

Differences in entering and exiting counts that are caused by queuing in between the two count locations suggest that the analyst should extend the count period to ensure that all demand is included in both counts.

Accurate vehicle classification counts and accurate travel speeds can also affect the traffic volumes. Trucks and other large vehicles and inaccurate speeds can skew the volume counts.

2.5 Example Problem: Data Collection and Preparation

The same ramp metering example problem discussed in chapter 1.0 is continued here. Now the task is to identify, gather, and prepare the data for the study area and the afternoon peak-hour analytical period.

Road Geometry: Aerial photographs, construction drawings, and field inspections are used to obtain the lengths and the number of lanes for each section of the freeway, ramps, and surface streets. Turn lanes and pocket lengths are determined for each intersection. Transition lengths for lane drops and additions are determined. Lane widths are measured if they are not standard widths. Horizontal curvature and curb return radii are determined if the selected software tool is sensitive to these features of the road and freeway design. Free-flow speeds are estimated based on the posted speed limits for the freeway and the surface streets.

Traffic Controls: Existing signal settings were obtained from the agencies' records and verified in the field. The controllers at the interchange ramp terminals are all fixed-time, having a cycle length of 90 s. The signals along Green Bay Avenue are traffic-actuated.

Demands: Field measurements of traffic volumes on the freeway mainline and ramps, and turning-movement counts at each intersection were conducted for a 2-h period during the

afternoon peak period (4:00 to 6:00 p.m.). The peak hour was determined to be between 5:00 and 6:00 p.m., and the highest 15-min volumes occurred between 5:30 and 5:45 p.m.

Vehicle Characteristics: The default vehicle mix and characteristics provided with the microsimulation software were used in this analysis.[21]

Calibration Data: The model will be calibrated against estimates of capacity and traffic counts, and the system performance will be assessed against travel time data.

Saturation flows for protected movements at traffic signals were estimated using the HCM 2000 methodology and verified through field observations at a few approaches with long queues.[22]

Capacity values for basic freeway segments were estimated using the HCM 2000 procedures.[23]

The traffic count data have already been discussed.

The system performance calibration data were obtained at the same time as the traffic counts, which consisted of travel times obtained from floating car runs, delays at traffic signals, and speeds on the freeway.

Data Preparation: The input data were reviewed by the analyst for consistency and accuracy prior to data entry. The turning, ramp, and freeway mainline counts were reconciled by the analyst to produce a set of consistent counts for the entire study area. After completion of the reconciliation, all volumes discharged from upstream locations are equal to the volumes arriving at the downstream location. Based on local knowledge of the area, the analyst determined that midblock sources were not required for the surface streets in the study area.

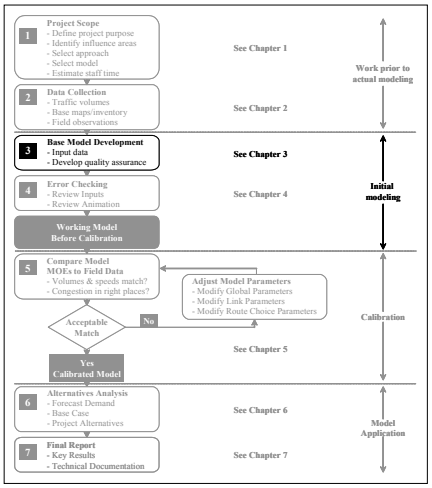
---

[21] The selected software had been developed with defaults appropriate to the United States, and this was considered to be sufficiently accurate for this planning study of ramp metering.

[22] The HCM was used to estimate surface-street saturation flow rates because none of the intersections were currently operating nor were they expected to operate within 90 percent of their capacity. Thus, an estimate is sufficiently accurate. Field measurements were not required.

[23] The HCM was used to estimate freeway capacity because there is no freeway congestion under existing conditions that would enable field measurement of capacities. Another method would be to observe capacities on a similar type of facility in the area that operated under congested conditions. This method may be preferred if a comparable facility can be found.

3.0 Base Model Development



This chapter provides general guidance on the procedures for developing a microsimulation model. There are many software programs for performing this task and each has its own unique coding methodology. The guidance provided here is intended to give general advice on model development; however, the analyst should consult the specific microsimulation software documentation for information on available data input tools and techniques.

Building a model is analogous to building a house. You begin with a blueprint and then you build each element in sequence—the foundation, the frame, the roof, the utilities and drywall, and finally the interior details. The development of a successful simulation model is similar in that you must begin with a blueprint (the link-node diagram) and then you proceed to build the model in sequence—coding links and nodes, filling in the link geometries, adding traffic control data at appropriate nodes, coding travel demand data, adding traveler behavior data, and finally selecting the model run control parameters.

3.1 Link-Node Diagram: Model Blueprint

The link-node diagram is the blueprint for constructing the microsimulation model. The diagram identifies which streets and highways will be included in the model and how they will be represented. An example link-node diagram is shown in figure 4. This step is critical and the modeler should always prepare a link-node diagram for a project of major complexity.

The link-node diagram can be created directly in the microsimulation software or offline using various types of computer-aided design (CAD) software. If the diagram is created in the microsimulation software, then it is helpful to import a map or aerial photograph into the software over which the link-node diagram can be overlaid. If the diagram is created offline using CAD software, then it is helpful to import a map or photograph into the CAD software.

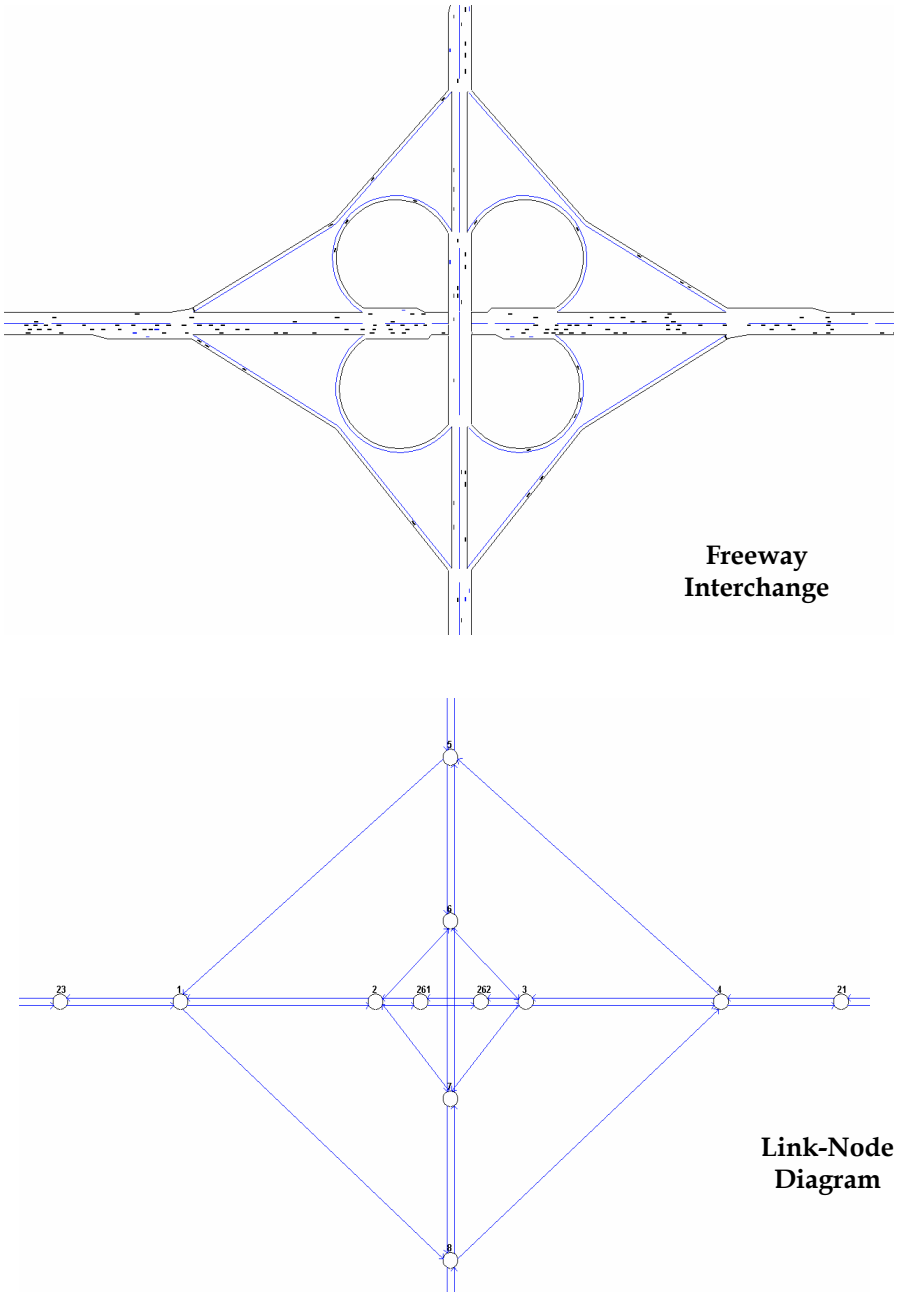


Figure 4. Example of link-node diagram.

Nodes are the intersection of two or more links. Nodes are usually placed in the model using x-y coordinates and they can be at a place that represents an intersection or a location where there is a change in the link geometry. Some simulation software may also warrant consideration of a z coordinate. The node locations can be obtained from aerial photographs, maps, or physical measurements.

Links are one-directional segments of surface streets or freeways.[24] Links represent the length of the segment and usually contain data on the geometric characteristics of the road or highway between the nodes. Ideally, a link represents a roadway segment with uniform geometry and traffic operation conditions.[25]

The analyst should consider establishing a node-numbering scheme to facilitate error checking and the aggregation of performance statistics for groups of links related to a specific facility or facility type (see table 3 for an example of a node-numbering scheme designed to enable the rapid determination of the facility type in text output). Much of the output produced by microsimulation software is text, with the results identified by the node numbers at the beginning and end points of each link. A node-numbering convention can greatly facilitate the search for results in the massive text files commonly produced by simulation software.[26]

If the link-node diagram was created offline (using some other software besides the microsimulation software), then the information in the diagram needs to be entered (or imported) into the microsimulation software. The x-y coordinates and identification numbers of the nodes (plus any feature points needed to represent link curvature) are entered. The nodes are then connected together to develop the link-node framework for the model.

---

[24]Some software programs do not always use a link-node scheme, while others allow the analyst to code both directions of travel with a single link. The two-way links coded by the user are then represented internally (inside the software) as two one-way links.

[25]The analyst may find it desirable to split links for performance reporting purposes. For example, it may be desirable to report density separately for the 460 m (1500-ft) influence area on the freeway near a ramp merge or diverge. However, the analyst should be cautious about the potential impact of split links on the ability of the software to accurately simulate vehicle behavior.

[26]Some software programs provide analytical modules that assist the analyst in displaying and aggregating the results for specific groups of links. This feature reduces the necessity of adopting a node-numbering convention; however, a numbering convention can still result in a significant labor savings when reviewing text output or when importing text results into other software programs for analytical purposes.

Table 3. Example node-numbering convention.

Segments	Range		Description
	From	To	
0's	1	99	Miscellaneous
100's	100	199	Northbound Freeway Mainline
200's	200	299	Northbound Freeway Ramps
300's	300	399	Southbound Freeway Mainline
400's	400	499	Southbound Freeway Ramps
500's	500	599	Eastbound Freeway Mainline
600's	600	699	Eastbound Freeway Ramps
700's	700	799	Westbound Freeway Mainline
800's	800	899	Westbound Freeway Ramps
900's	900	999	Arterials

3.2 Link Geometry Data

Once the modeler has completed coding the link-node diagram, then the modeler must input the physical and operational characteristics of the links into the model. This phase is when the modeler must code the following:

- Number of lanes.
- Lane width.
- Link length.
- Grade.
- Curvature.
- Pavement conditions (dry, wet, etc.).
- Sight distance.
- Bus stop locations.
- Crosswalks and other pedestrian facilities.

- Bicycle lanes/paths.
- Others.

The specific data to be coded for the links will vary according to the microsimulation software.

3.3 Traffic Control Data at Intersections and Junctions

Most microsimulation uses a time-step simulation to describe traffic operations (which is usually 1 s or less). Vehicles are moved according to car-following logic in response to traffic control devices and in response to other demands. Traffic control devices for microsimulation models will vary. Listed below are some examples:

- No control.
- Yield signs.
- Stop signs.
- Signals (pretimed, actuated, real-time traffic adaptive).
- Ramp metering.

3.4 Traffic Operations and Management Data for Links

Traffic operations and management data for links consist of the following:

- Warning data (incidents, lane drops, exits, etc.).
- Regulatory data (speed limits, variable speed limits, high-occupancy vehicles (HOVs), high-occupancy toll (HOT), detours, lane channelizations, lane use, etc.).
- Information (guidance) data (dynamic message signs and roadside beacons).
- Surveillance detectors (type and location).

3.5 Traffic Demand Data

Traffic demands are defined as the number of vehicles and the percentage of vehicles of each type that wish to traverse the study area during the simulation time period. Furthermore, it may be necessary to reflect the variation in demand throughout the simulation time period. In most software programs, the traffic entering into the network is usually defined by some parameter, and traffic leaving the network is usually computed based on parameters internal to the network (turning movements, etc.). The modeler must code the traffic volume by first starting from the external nodes (this is where the traffic is put into the model). Once all of the entering traffic volumes at the external nodes are

coded, the modeler will then go into the model and define the turning movements and any other parameter related to route choice. The key traffic demand data are:

- Entry volumes by vehicle type and turning fractions at all intersections or junctions (random walk simulators).
- O-D/path-specific and vehicle data (path-specific simulators).
- Bus operations (routes and headways/schedules).
- Bicycle and pedestrian demand data.

3.6 Driver Behavior Data

Default values for driver behavior data are usually provided with the microsimulation software. Users may override the default values of any driver behavior parameters if valid observed data are available (e.g., desired free-flow speed, discharge headway, startup lost time at intersections, etc.). Driver behavior data include:

- Driver's aggressiveness (for minimum headway in car-following, gap acceptance for lane changing, response to yellow interval).
- Availability of (real-time) information for the driver.
- Driver's response to information (for pretrip planning and/or en route switching).

3.7 Events/Scenarios Data

Events data include:

- Blockages and incidents.
- Work zones.
- Parking activity in curb lane.

This data is optional and will vary according to the specific application being developed by the analyst.

3.8 Simulation Run Control Data

All simulation software contain run control parameters to enable the modeler to customize the software operation for their specific modeling needs. These parameters will vary between software programs. They generally include:

- Length of simulation time.
- Selected MOEs or output (e.g., reports, animation files, or both).
- Resolution of simulation results (e.g., temporal and special resolution).
- Other system parameters to run the model.

3.9 Coding Techniques for Complex Situations

Microsimulation software allows the modeler to develop a model to represent the real-world situation. However, not all possible real-world situations were necessarily contemplated when the software was originally written. This is when the modeler, with a good understanding of the operation of the software, can "extend" the software to simulate conditions not originally incorporated into the microsimulation software. However, this should only be attempted after the tools, resources, and skills available are fully appreciated. This is because new tools continue to become available that account for selected real-world situations.

Provided below are some examples of how basic microsimulation software can be applied to less standard situations, noting that calibration (discussed in chapter 5.0) must also consider the following situations:

- A curb lane is heavily used for parking, loading, and unloading activities. As a result, this lane may be blocked virtually all of the time. If the simulation software cannot correctly replicate the real situation, the modeler may consider removing this lane from the link.

- There is a drawbridge in the real-world network; however, the software does not have explicit "drawbridge" modeling capability. If boat traffic is very heavy and is equivalent to vehicular traffic, the modeler may consider coding the drawbridge as an intersection with fixed-time signal control. If boat traffic is irregular, the analyst might consider coding the drawbridge as a semi-actuated traffic signal-controlled intersection. If boat traffic is rare, then the drawbridge might be coded as a link with events that result in a 100-percent capacity loss with a certain probability of occurrence during the analytical period.

- Traffic regularly backs up on a freeway off-ramp, backing up onto the freeway. However, instead of stopping on the freeway and blocking a lane, the queue forms on the shoulder of the freeway, keeping the right-hand lane open. If this is the case in the study area, the modeler may artificially extend the off-ramp length to realistically model the traffic in the field.

3.10 Quality Assurance/Quality Control (QA/QC) Plan

Once the modeler is satisfied that the model is coded, the modeler must then vet the model for completeness. This could be a QA/QC process. Many of the models input are developed in a format that can be exported to a spreadsheet format. Therefore, the modeler can use this type of tool to aid in vetting the model for completeness and accuracy. The next chapter on error checking describes quality control tests in more detail.

3.11 Example Problem: Base Model Development

Continuing with the same example problem from the previous chapters, the task is now to code the base model.

The first step is to code the link-node diagram. How this is best done can be determined from the software user's guide. Figure 5 shows the link-node diagram for the example problem.

The next steps are to input:

1. Link geometry data.
2. Intersection traffic control data.
3. Traffic operations and management data.
4. Traffic count data.
5. Run control parameters.[27]

[27] The software-provided default values for driver behavior (aggressiveness, awareness, etc.) were used in this example problem. They were considered to be sufficiently accurate for this analysis.

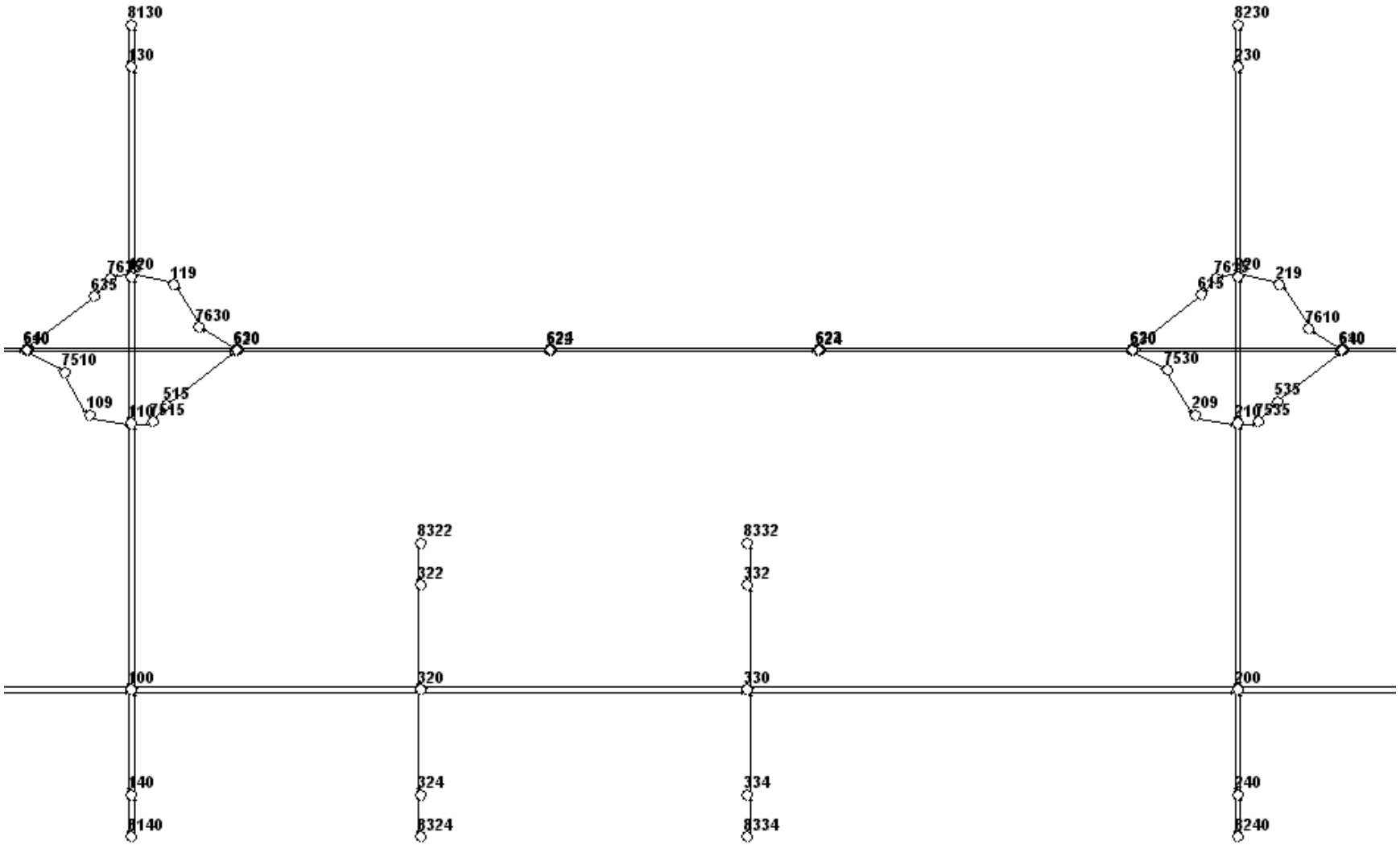
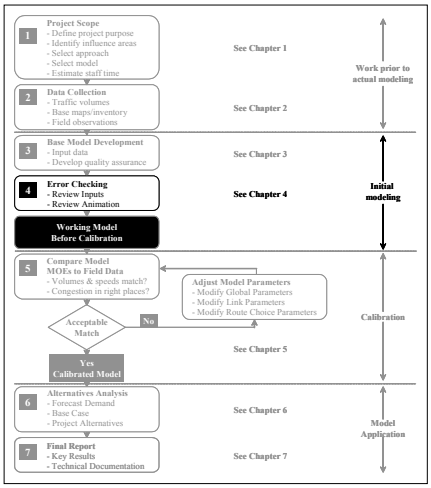


Figure 5. Link-node diagram for example problem.

4.0 Error Checking

The error correction step is essential in developing a working model so that the calibration process does not result in parameters that are distorted to compensate for overlooked coding errors. The succeeding steps of the calibration process rely on the elimination of all major errors in demand and network coding before calibration.



Error checking involves various reviews of the coded network, coded demands, and default parameters. Error checking proceeds in three basic stages: (1) software error checking, (2) input coding error checking, and (3) animation review to spot less obvious input errors.

4.1 Review Software Errors

The analyst should review the software and user group Web sites to ensure that he or she is aware of the latest known "bugs" and user workarounds for the software. The analyst should ensure that he or she is using the latest version and "patch" of the software.

4.2 Review Input

A checklist for verifying the accuracy of the coded input data is provided below:

1. Link and node network:
    a. Check basic network connectivity (are all connections present?).
    b. Check link geometry (lengths, number of lanes, free-flow speed, facility type, etc.).
    c. Check intersection controls (control type, control data).
    d. Check for prohibited turns, lane closures, and lane restrictions at the intersections and on the links.
2. Demand:
    a. Check vehicle mix proportions at each entry node/gate/zone.
    b. Check identified sources and sinks (zones) for traffic.
    c. Verify zone volumes against traffic counts.
    d. Check vehicle occupancy distribution (if modeling HOVs).

e. Check turn percentages (if appropriate).

f. Check O-Ds of trips on the network.

3. Traveler behavior and vehicle characteristics:

   a. Check and revise, as necessary, the default vehicle types and dimensions.

   b. Check and revise the default vehicle performance specifications.

The following techniques may be useful to increase the efficiency and effectiveness of the error-checking process:[1]

- Overlay the coded network over aerial photographs of the study area to quickly verify the accuracy of the coded network geometry.

- If working with software that supports three-dimensional modeling, turn on the node numbers and look for superimposed numbers. They are an indication of unintentionally superimposed links and nodes. Two or more nodes placed in the same location will look like a single node in a three-dimensional model when viewed in two dimensions. The links may connect to one of the nodes, but not to the other.

- For a large network, a report summarizing the link attributes should be created so that their values can be easily reviewed.

- Use color codes to identify links by the specific attribute being checked (e.g., links might be color-coded by free-flow speed range). Out-of-range attributes can be identified quickly if given a particular color. Breaks in continuity can also be spotted quickly (e.g., a series of 56-km/h (35-mi/h) links with one link coded as 40 km/h (25 mi/h)).

4.3 Review Animation

Animation output enables the analyst to see the vehicle behavior that is being modeled and assess the reasonableness of the microsimulation model itself. Running the simulation model and reviewing the animation, even with artificial demands, can be useful to identify input coding errors. The analyst inputs a very low level of demand and then follows individual vehicles through the network. Aberrant vehicle behavior (such as unexpected braking or stops) is a quick indicator of possible coding errors.

A two-stage process can be followed in reviewing the animation output:

1. Run the animation at an extremely low demand level (so low that there is no congestion). The analyst should then trace single vehicles through the network and see where they unexpectedly slow down. These will usually be locations of minor

[1] Some of these techniques may not be available or necessary for some software programs.

network coding errors that disturb the movement of vehicles over the link or through the node. This test should be repeated for several different O-D zone pairs.

2. Once the extremely low demand level tests have been completed, then run the simulation at 50 percent of the existing demand level. At this level, demand is usually not yet high enough to cause congestion. If congestion appears, it may be the result of some more subtle coding errors that affect the distribution of vehicles across lanes or their headways. Check entry and exit link flows to verify that all demand is being correctly loaded and moved through the network.

The animation should be observed in close detail at key congestion points to determine if the animated vehicle behavior is realistic. If the observed vehicle behavior appears to be unrealistic, the analyst should explore the following potential causes of the unrealistic animation in the order shown below:

Error in Analyst Expectations: The analyst should first verify in the field the correct vehicle behavior for the location and time period being simulated before deciding that the animation is showing unrealistic vehicle behavior. Many times, analyst expectations of realistic vehicle behavior are not matched by actual behavior in the field.[2] Field inspection may also reveal the causes of vehicle behavior that are not apparent when coding the network from plans and aerial photographs. These causes need to be coded into the model if the model is expected to produce realistic behavior.[3]

Analyst Data Coding Errors: The analyst should check for data coding errors that may be causing the simulation model to represent travel behavior incorrectly. Subtle data coding errors are the most frequent cause of unrealistic vehicle behavior in commercial microsimulation models that have already been subjected to extensive validation. Subtle coding errors include apparently correctly coded input that is incorrect because of how it is used in the model to determine vehicle behavior.[4]

---

[2] Analysts should not expect classic macroscopic traffic-flow concepts to apply at the microscopic individual-vehicle level. Macroscopic flow concepts (e.g., no variance in mean speed at low flow rates) do not apply to the behavior of an individual vehicle over the length of the highway. An individual vehicle's speed may vary over the length of the highway and between vehicles, even at low flow rates. Macroscopic flow theory refers to the average speed of all vehicles being relatively constant at low flow rates, not individual vehicles.

[3] A TMC with high-density camera spacing will be very helpful in reviewing the working model. Many TMCs are now providing workstations for traffic analysis/simulation staff.

[4] For example, it could be that the warning sign for an upcoming off-ramp is posted in the real world 0.40 km (0.25 mi) before the off-ramp; however, because the model uses warning signs to identify where people start positioning themselves for the exit ramps, the analyst may have to code the warning sign at a different location (the location where field observations indicate that the majority of the drivers start positioning themselves for the off-ramp).

A comparison of model animation to field design and operations cannot be overemphasized. Some of the things to look for include:

- Overlooked data values that need refinement.
- Aberrant vehicle operations (e.g., drivers using shoulders as turning or travel lanes, etc.).
- Previously unidentified points of major ingress or egress (these might be modeled as an intersecting street).
- Operations that the model cannot explicitly replicate (certain operations in certain tools/models), such as a two-way center turn lane (this might be modeled as an alternating short turn bay).
- Unusual parking configurations, such as median parking (this might be modeled operationally by reducing the free-flow speed to account for this friction).
- Average travel speeds that exceed posted or legal speeds (use the observed average speed in the calibration process).
- Turn bays that cannot be fully utilized because of being blocked by through traffic.
- In general, localized problems that can result in a systemwide impact.

4.4 Residual Errors

If the analyst has field-verified his or her expectations of traffic performance and has exhausted all possible input errors, and the simulation still does not perform to the analyst's satisfaction, there are still a few possibilities. The desired performance may be beyond the capabilities of the software, or there may be a software error.

Software limitations can be identified through careful review of the software documentation. If software limitations are a problem, then the analyst will have to work around the limitations by "tricking" or "forcing" the model to produce the desired performance.[5] If the limits are too great, the analyst might seek an alternate software program without the limitations. Advanced analysts can also write their own software interface with the microsimulation software (called an "application program interface" (API)) to overcome the limitations and produce the desired performance.

Software errors can be tested by coding simple test problems (such as a single link or intersection) where the result (such as capacity or mean speed) can be computed manually and compared to the model. Software errors can only be resolved by working with the software developer.

---

[5] For example, a drawbridge that opens regularly might be coded as a traffic signal.

4.5 Key Decision Point

The completion of error checking is a key decision point. The next task—model calibration—can be very time-consuming. Before embarking upon this task, the analyst should confirm that error checking has been completed, specifically:

- All input data are correct.
- Values of all initial parameters and default parameters are reasonable.
- Animated results look fine based on judgment or field inspection.

Once the error checking has been completed, the analyst has a working model (though it is still not calibrated).

4.6 Example Problem: Error Checking

Continuing with the same example problem from the previous chapters, the task is now to error check the coded base model.

Software: The latest version of the software was used. Review of the model documentation and other material in the software and user groups' Web sites indicated that that there were no known problems or bugs related to the network under study and the scenarios to be simulated.

Review of Input Data and Parameters: The coded input data were verified using the input files, the input data portion of the output files, static displays, and animation.

First, the basic network connectivity was checked, including its consistency with coded geometry and turning restrictions. All identified errors were corrected. For example, one link with exclusive left- and right-turn lanes (no through traffic) was coded as feeding a downstream through link.

Static network displays were used extensively to verify the number of lanes, lane use, lane alignment (i.e., lane number that feeds the downstream through link), and the location of lane drops. At this step, the consistency of link attributes was checked. For example, is the free-flow speed of 87 km/h (55 mi/h) coded for all freeway links?

Next, the traffic demand data were checked. Input volumes at the network entrances were specified in four time slices. The input values were checked against the collected data.

Special attention was given to the traffic patterns at the interchange ramp terminals to avoid unrealistic movement. The software provisions (and options) were exercised to force the model not to assign movements to travel paths that were not feasible.[6]

The vehicle characteristics and performance data were reviewed.

The model displays and animation were used to verify the input data and operation of the traffic signals coded at each intersection. For fixed-time signals, the phasing and signal settings were checked (see figure 6). For actuated signals, initial simulations were performed with reduced, but balanced, volumes to ensure that all phases were activated by the traffic demand. This was done because often the inappropriate settings of the phase flags cause signals to malfunction within the simulation and produce unreasonable results. This step also involves checking the location of the detectors and their association with the signal phases.

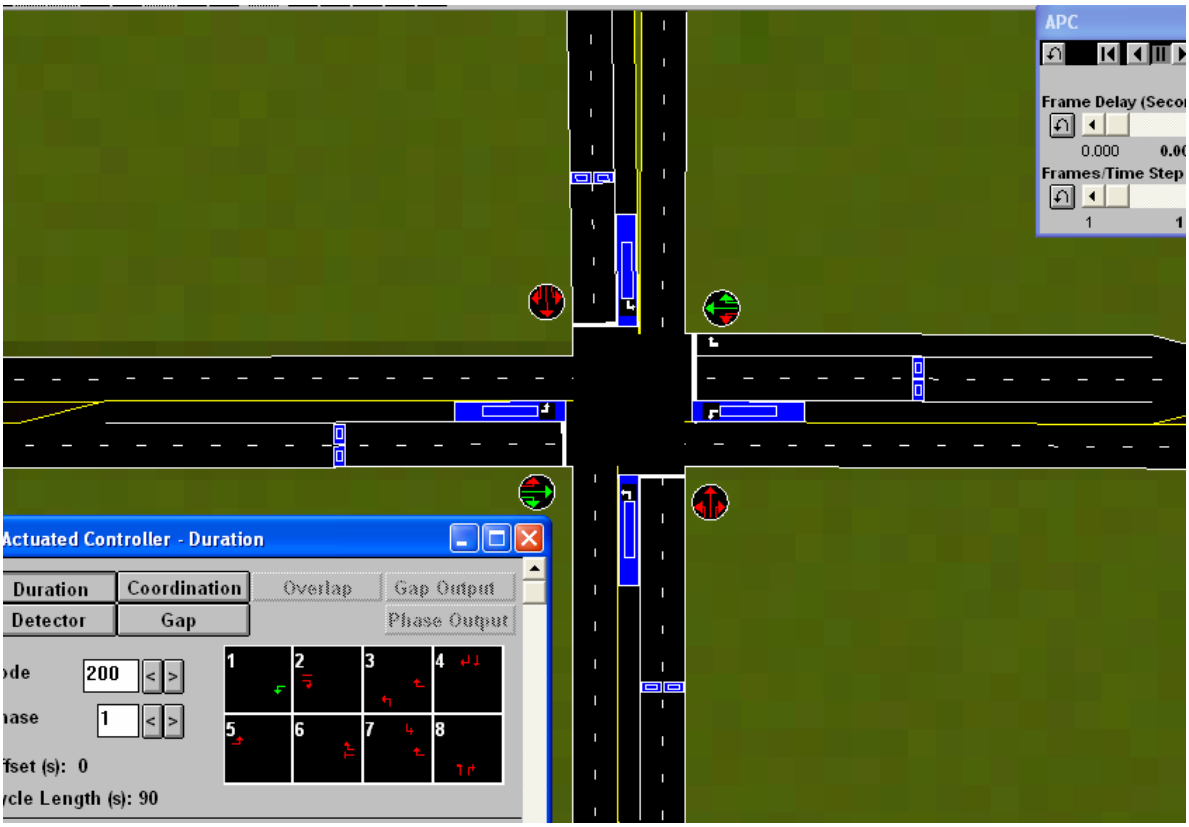


Figure 6. Checking intersection geometry, phasing, and detectors.

Review Animation: Following the checking of the input data, the model was run using very low demand volumes to verify that all of the vehicles travel the network without

[6] An example of this would be coding to address freeway drivers who do or do not make U-turns at an interchange (i.e., get off and then get back on the freeway in the opposite direction).

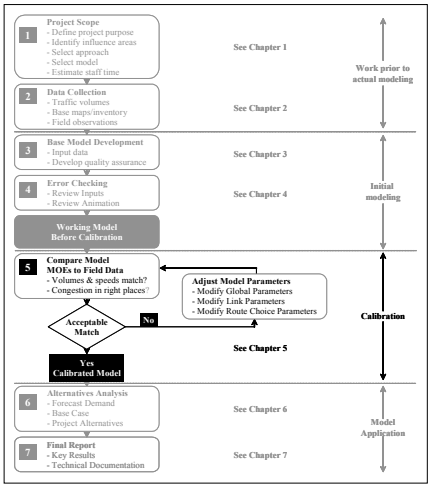
slowdowns. This step uncovered minor errors in the links alignments that needed to be adjusted.

Next, the traffic demands were specified to about 50 percent of the actual volumes and the simulation model was rerun. Animation was used to verify that all demands were properly loaded in the network links and the traffic signals were properly operating. The link and system performance measures (travel time, delay) were also checked for reasonableness (i.e., they should reflect free-flow conditions).

Careful checking of the animation revealed subtle coding problems. For example, the coded distance of a warning sign (for exiting vehicles) or the distance from the start of the link to the lane drop affects the proper simulation of driver behavior. These problems were corrected.

Key Decision Point: The model, as revised throughout the error-checking process, was run with all the input data (actual demands) and the default model parameters. The output and animation were also reviewed and discussed with other agency staff who were familiar with the study area. The conclusion is that the model is working properly.

5.0 Calibration of Microsimulation Models



Upon completion of the error-checking task, the analyst has a working model. However, without calibration, the analyst has no assurance that the model will correctly predict traffic performance for the project.

Calibration is the adjustment of model parameters to improve the model's ability to reproduce local driver behavior and traffic performance characteristics. Calibration is performed on various components of the overall model.

The importance of calibration cannot be overemphasized. Recent tests of six different software programs found that calibration differences of 13 percent in the predicted freeway speeds for existing conditions increased to differences of 69 percent in the forecasted freeway speeds for future conditions.[7]

5.1 Objectives of Calibration

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. Since no single model can include the whole universe of variables, every model must be adapted to local conditions.

The objective of calibration is to improve the ability of the model to accurately reproduce local traffic conditions.

Every microsimulation software program comes with a set of user-adjustable parameters for the purpose of calibrating the model to local conditions. Therefore, the objective of calibration is to find the set of parameter values for the model that best reproduces local traffic conditions.

For the convenience of the analyst, the software developers provide suggested default values for the model parameters. However, only under very rare circumstances will the model be able to produce accurate results for a specific area using only the default parameter values. The analyst should *always* perform some calibration tests to ensure that the coded model accurately reproduces local traffic conditions and behavior.

---

[7] Bloomberg, L., M. Swenson, and B. Haldors, *Comparison of Simulation Models and the Highway Capacity Manual*, Preprint, Annual Meeting, TRB, Washington, DC, 2003.

The fundamental assumption of calibration is that the travel behavior models in the simulation model are essentially sound.[8] There is no need to verify that they produce the correct delay, travel time, and density when they are given the correct input parameters for a link. Therefore, the only remaining task for the analyst is to adjust the parameters so that the models correctly predict local traffic conditions.

5.2 Calibration Approach

Calibration involves the review and adjustment of potentially hundreds of 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 that a new one occurs somewhere else. Therefore, it is essential to break the calibration process into a series of logical, sequential steps—a strategy for calibration.

To make calibration practical, the parameters must be divided into categories and each category must be dealt with separately. The analyst should divide the available calibration parameters into the following two basic categories:

- Parameters that the analyst is certain about and does not wish to adjust.
- Parameters that the analyst is less certain about and willing to adjust.

The analyst should attempt to keep the set of adjustable parameters as small as possible to minimize the effort required to calibrate them. Whenever practical, the analyst should use observed field data to reflect local conditions. This observed data will serve as the nonadjustable values for certain calibration parameters, thus leaving the set of adjustable parameters to a minimum. However, the tradeoff 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, then the route choice adjustment parameters are calibrated.

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.

[8] The analyst determines the soundness of the simulation software when selecting it for use in the study.

The following three-step strategy is recommended for calibration:

Calibration Strategy:

1. *Calibrate capacity parameters.*
2. *Calibrate route choice parameters.*
3. *Calibrate overall model performance.*

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 performed first, followed by link-specific fine-tuning. The HCM can be used as an alternative source of capacity target values if field measurements are not feasible.

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 performed first, followed by link-specific fine-tuning.

3. **System Performance Calibration:** Finally, the overall model estimates of system performance (travel times and queues) are compared to the field measurements for travel times and queues. Fine-tuning adjustments are made to enable the model to better match the field measurements.

5.3 Step 1: Calibration for Capacity

The capacity calibration step adjusts the global and link-specific capacity-related parameters in the simulation model to best replicate local field measurements for capacity. This is an important step because capacity has a significant effect on predicted system performance (delay and queues).[9] The objective of this calibration step is to find a set of model parameters that causes the model to come as close as possible to matching the field measurements for traffic capacity.

Capacity calibration has two phases:

1. *Global calibration*
2. *Link fine-tuning*

The capacity calibration step consists of two phases: (1) a global calibration phase and (2) a fine-tuning phase. Global calibration is first performed to identify the appropriate network-wide value of the capacity parameter(s) that best reproduces

[9] 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 in the calibration process) only on the critical counts at the bottlenecks and to get the model to reproduce these counts correctly. Once this has been 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 have already been checked against the counts at the entry gates as part of the first step (error checking).

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 bottleneck.[10]

The capacity calibration procedure is as follows:

1. Collect field measurements of capacity.
2. Obtain model estimates of capacity.
3. Select calibration parameters.
4. Set calibration objective function.
5. Perform search for optimal parameter value.
6. Fine-tune the calibration.

The following subsections explain this procedure in more detail.

5.3.1 Collect 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 nonsignalized 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 where the queue discharges. This observed flow rate is measured only while an upstream queue is present. It is totaled 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 procedures outlined in appendix H to the signalized intersection chapter of HCM. The capacity of the signalized intersection approach is then the saturation flow multiplied by the portion of green time in the signal cycle:

$$c = s\frac{g}{C} \tag{2}$$

[10]It is certainly possible to use link-specific parameters exclusively during the capacity calibration step; however, this eliminates the benefits of the global parameter adjustment. The global adjustment ensures the accuracy of the model-predicted capacities on all links (even those not currently congested). Adjustment of link-specific parameters ensures the model accuracy only for the specific link.

where:

c = capacity (vehicles per hour (veh/h))

s = saturation flow (veh/h per green phase)

g = effective green time

C = cycle length

Higher saturation flows and lower startup lost times result in higher capacities.

Several measurements of maximum flow rates should be made in the field and averaged. Procedures are provided in the appendix for estimating how many measurements are required to estimate capacity within a desired confidence interval (see Appendix B: Confidence Intervals). If capacity cannot be measured in the field, then the HCM methodology can be used to estimate capacity. The HCM methods should not be considered a default technique since the estimates are not as accurate as direct field measurements.

5.3.2 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. Thus, 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 initially shows no congestion at the bottlenecks, the demand must be temporarily increased before calibration can proceed.

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 after the capacity calibration has been completed, but before the route choice calibration step.

All "false" bottlenecks in the model must be cleared before calibration can proceed.

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 nonsignalized 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 that the queue is present. The result is then divided by the number of lanes and is 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 the green phase. 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.[11] 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 were averaged, the model runs should be repeated several times and the maximum flow rate at each location should be 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 the procedures provided in appendix B (Confidence Intervals).

5.3.3 Select Calibration Parameter(s)

Only the model parameters that directly affect capacity are calibrated at this time. Each microsimulation software program 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.

This chapter does not intend to describe all of the parameters for all of the simulation models that are available. An illustrative list of capacity-related parameters for freeways and signalized arterials is given below:

Freeway Facilities:

- Mean following headway.
- Driver reaction time.
- Critical gap for lane changing.
- Minimum separation under stop-and-go conditions.

Signalized Intersections:

- Startup lost time.
- Queue discharge headway.
- Gap acceptance for unprotected left turns.

For example, in CORSIM, the two parameters that most globally affect capacity are "headway factor by vehicle type" (record type 58) and "car following

Mean headway is a good global parameter for calibrating capacity. It may have different names in different software.

[11] The headways for the first three vehicles are discarded.

sensitivity factor" (record type 68). There are numerous other parameters in CORSIM that affect capacity; however, they mostly apply only to specific conditions (e.g., "acceptable gap in near-side cross traffic for vehicles at a stop sign" (record type 142)).

For the fine-tuning phase in CORSIM, the link-specific capacity calibration parameter is "mean queue discharge headway," which can be found in record type 11 for city streets, and in record type 20 for freeways ("link-specific car-following sensitivity multiplier").

5.3.4 Set Calibration Objective Function

It is recommended that the analyst seek to minimize the mean square 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. Each set of repetitions has a single set of model parameter values p with different random number seeds[12] for each repetition within the set.[13]

Select a set of model parameters p to minimize:

$$MSE = \frac{1}{R}\sum_{r}(M_{ltpr} - F_l)^2 \quad (3)$$

Subject to:

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

where:

MSE = mean square 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

[12]Microsimulation models assign driver-vehicle characteristics from statistical distributions using random numbers. The sequence of random numbers generated depends on the initial value of the random number (random number seed). Changing the random number seed produces a different sequence of random numbers, which, in turn, produces different values of driver-vehicle characteristics.

[13]Some researchers have calibrated models using the percent MSE 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 squared 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.

R = number of repetitive model runs with fixed parameter values p_m and different random number seeds[14]

p_m = value of model parameter number m

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 the laws of physics, vehicle performance limits, and driver behavior extremes)

5.3.5 Perform 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, 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.

There are many software programs available for identifying the optimal combination of calibration parameters for minimizing the squared error between the field observations and the simulation model. The Optimization Technology Center Web site of the Argonne National Laboratory and Northwestern University lists several software programs for nonlinear least-squares parameter estimation. See the following:

www-fp.mcs.anl.gov/otc/Guide/SoftwareGuide/Categories/nonlinleastsq.html

Appendix D describes a few simple search algorithms that can be applied manually or with the aid of spreadsheets. Figure 7 illustrates how a single parameter search (m = 1) might look using one of these simple search algorithms for the number of model runs (R) set to 4.

[14]Since the objective is to minimize error, dividing by a constant R (the number of repetitions) would have no effect on the results. 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.

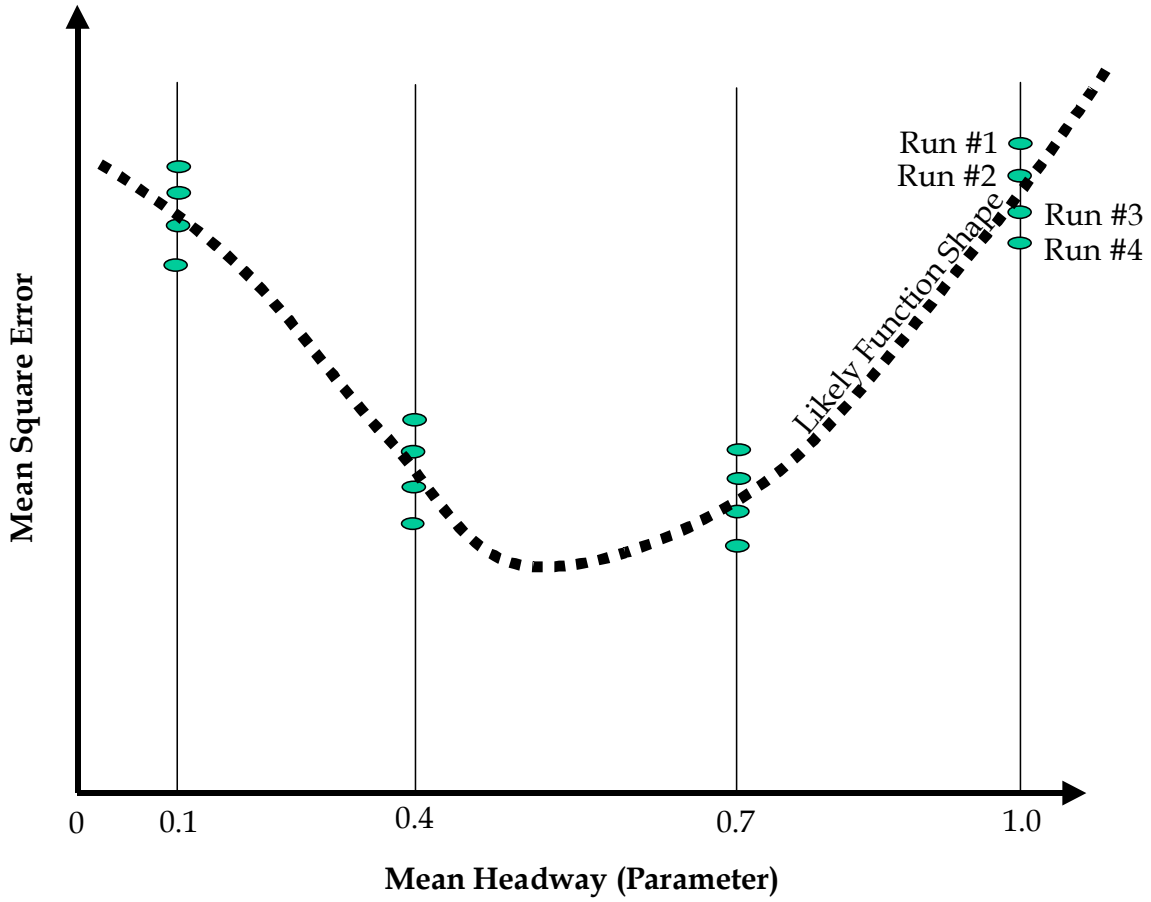


Figure 7. Minimization of MSE.

5.3.6 Fine-Tune the Calibration

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

Link-specific adjustment factors are often still necessary after global calibration; however, they should be used sparingly.

Link-specific capacity adjustments account for roadside factors that affect capacity, but are not typically coded in the microsimulation network input data set (such as presence of onstreet parking, numerous driveways, or narrow shoulders). Most simulation software programs 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 since they are not behavior-based. They are fixed adjustments that will be carried through all future runs of the simulation model.

5.4 Step 2: Route Choice Calibration

Once the analyst is satisfied that the model reproduces as closely as possible the field-measured capacities, the next step is to then calibrate the route choice parameters in the model to better match the observed flows. The temporary demand adjustments used in the previous capacity calibration step are reversed. The model-predicted volumes are then compared to the field counts and the analyst adjusts the route choice algorithm parameters until the best volume fit is achieved.[15]

If the model network consists of only a single facility, then no route choice calibration is possible or needed. This step is skipped.[16] This step is also skipped for microsimulation software that does not have route choice capabilities.

Route choice calibration proceeds in two phases: (1) global calibration and (2) link-specific fine-tuning.

5.4.1 Global Calibration

Global calibration of route choice consists of the application of a route choice algorithm and associated parameters.[17] The specific parameters vary by algorithm and software program, but usually involve weightings placed on the actual cost and travel time for each route. Additional parameters may be related 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.

5.4.2 Fine-Tuning

Once the global calibration has been completed, link-specific adjustments to cost or speed are made during the fine-tuning phase. The fine-tuning has been completed when the calibration targets for volumes have been met (see section 5.6 on calibration targets).

---

[15] The specific route choice algorithm parameters will vary by software program. They generally relate to the driver's awareness of, perception of, and sensitivity to travel time, delay, and the cost of alternate routes.

[16] For a single-facility network, if there are still some remaining volume errors after the capacity calibration step, then the input demands should be checked for errors.

[17] Some software programs allow selection of the algorithm and its associated parameters.

5.5 Step 3: System Performance Calibration

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

5.6 Calibration Targets

The objective of model calibration is to obtain the best match possible between model performance estimates and the 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 4 provides an example of calibration targets that were developed by Wisconsin DOT for their Milwaukee freeway system simulation model. They are based on guidelines developed in the United Kingdom.[18]

[18] "Traffic Appraisal in Urban Areas, Highways Agency," *Design Manual for Roads and Bridges: Volume 12, Section 2*, Department for Transport (formerly Department of Environment, Transport, and the Regions), London, England, May 1996 (www.official-documents.co.uk/document/deps/ha/dmrb/index.htm)

Table 4. Wisconsin DOT freeway model calibration criteria.

Criteria and Measures	Calibration Acceptance Targets
Hourly Flows, Model Versus Observed	
Individual Link Flows	
Within 15%, for 700 veh/h < Flow < 2700 veh/h	> 85% of cases
Within 100 veh/h, for Flow < 700 veh/h	> 85% of cases
Within 400 veh/h, for Flow > 2700 veh/h	> 85% of cases
Sum of All Link Flows	Within 5% of sum of all link counts
GEH Statistic < 5 for Individual Link Flows*	> 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 1 min, 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

*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.

Another example of suggested calibration targets is "Theil's Inequality Coefficient," which is broken down into three parts, each of which provides information on the differences between the model measures and the target measures.[19]

5.7 Example Problem: Model Calibration

The same example problem from the previous chapters is continued. The task is now to calibrate the model.

A review of the results produced by the working model at the end of the model development task indicates that there are discrepancies between the observed and simulated traffic performance. The purpose of the calibration process is to adjust the model parameters to better match field conditions.

Calibration for Capacity – Step 1: Global Calibration

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 2000 procedures. The estimated value for the saturation flow for protected movements at signalized intersections was 1800 vehicles per hour of green per lane (vphgpl) and the capacity of the freeway links was 2100 vphgpl.

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 for the throughput values to represent capacity.

Because of the lack of 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/veh (which is equivalent to a saturation flow of 2000 vphgpl

---

[19]Further discussions of Theil's Inequality Coefficient can be found in the *Advanced CORSIM Training Manual*, Minnesota DOT, 2002, and in Hourdakis, J., P. Michalopoulos, and J. Kottommannil, "Practical Procedure for Calibrating Microscopic Traffic Simulation Models," TRB, TRR 1852, Washington, D.C., 2003.

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 that 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 saturation flow/capacity and the field estimates.

Table 5 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-generated saturation flow within 5 percent of its true value. The mean values of the saturation flows were used in the calculation of MSE. Figure 8 shows the values of MSE for each parameter value. The results indicate that a value of 1.9 s produces the minimum MSE for this data set.

Table 5. Impact of queue discharge headway on predicted saturation flow.

Parameter Value	1	2	3	4	5	6	7	8	9	10	Sample Mean	Sample St Dev
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

Note: Each column labeled 1 through 10 represents a model run repetition using a different random number seed.

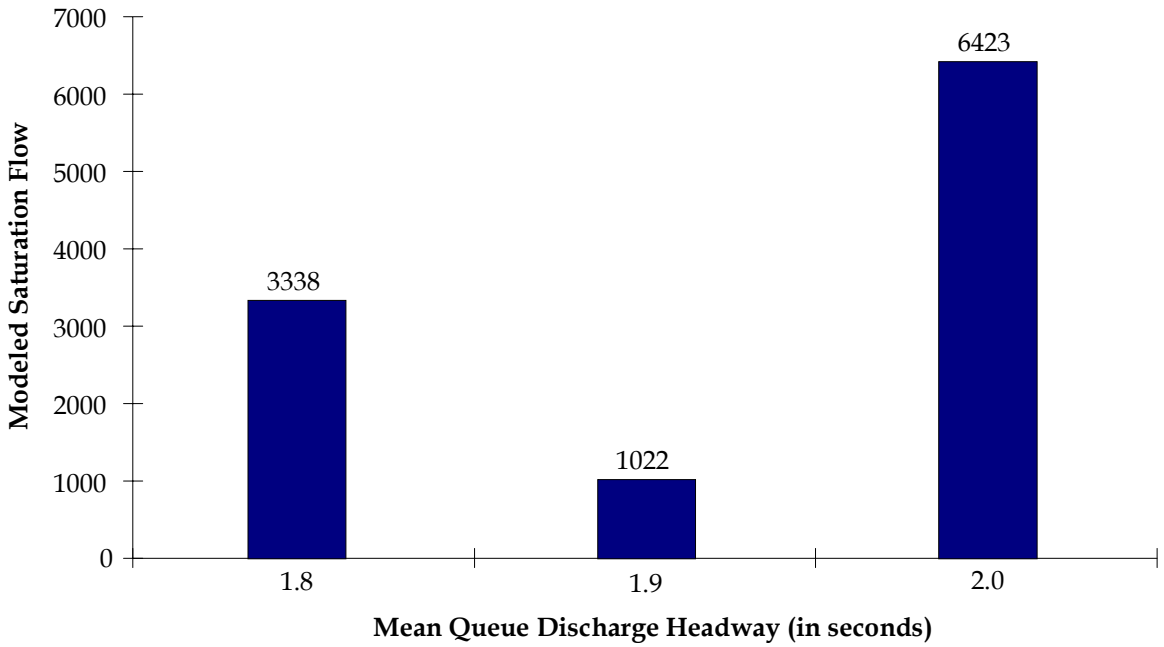


Figure 8. Impact of mean queue discharge headway on MSE of modeled saturation flow.

A similar 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 in CORSIM) must be reduced by 5 percent (i.e., 95 percent of its default value for the model throughput to match the observed value of 2100 vphgpl). This value minimizes the MSE.

Figures 9 and 10 illustrate the effect of this parameter on freeway performance. Before calibration (figure 9), the average speed on a series of freeway links is much lower than the speeds after calibration of the headways in the HCM 2000 (figure 10).

Calibration for Capacity—Step 2: Fine-Tuning

For this example, field capacity measurements were only available for two links on the network—one for the freeway and one for the surface street. These measures were used in the previous step to perform the global calibration. Given that no other field capacity data were available, no further fine-tuning (or link-level calibration) is needed.

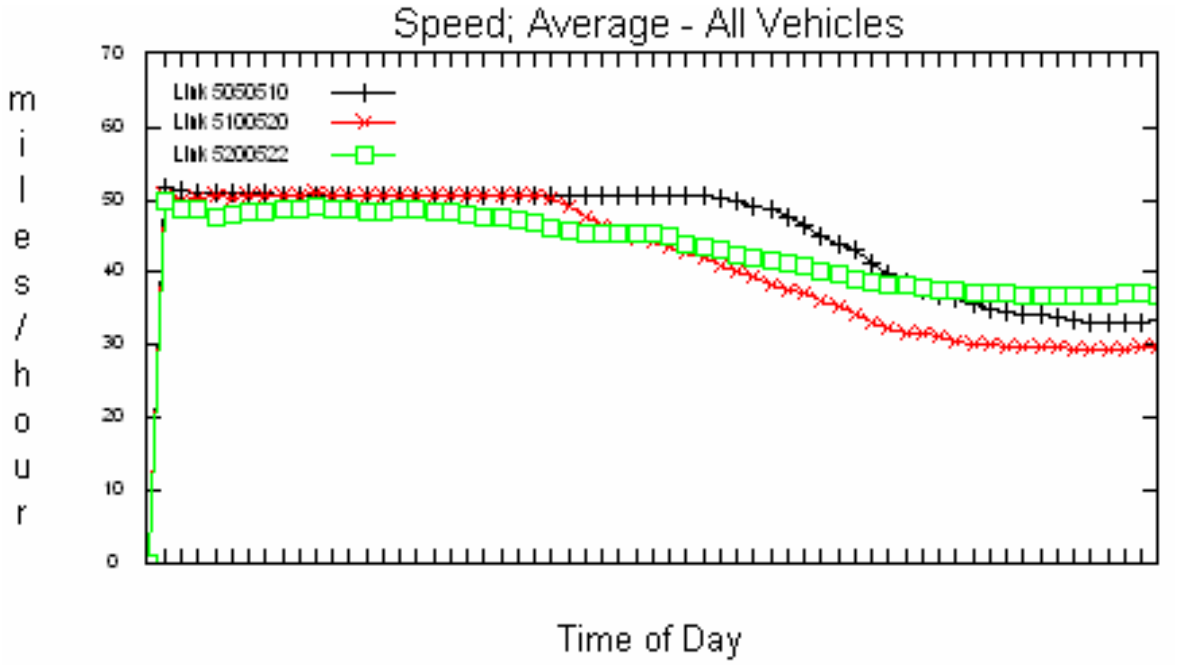


Figure 9. Average speed by minute of simulation (three links): Before calibration.

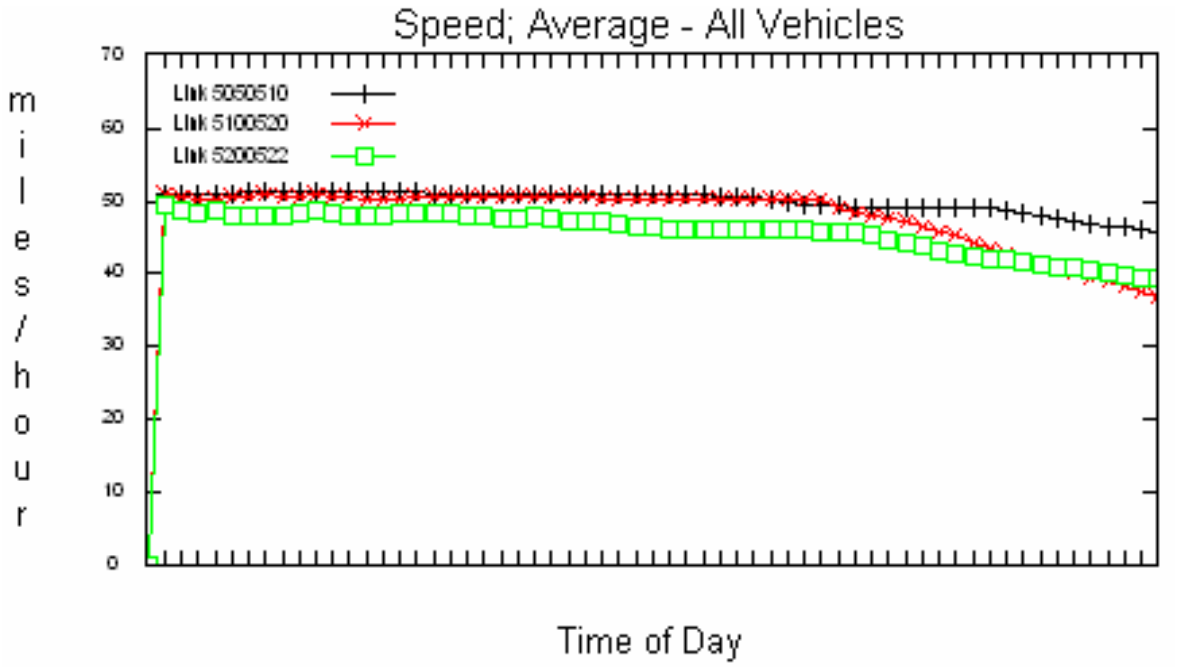


Figure 10. Average speed by minute (three links): After calibration.

Calibration for Route Choice

The prior calibration step has ensured that the model is predicting the correct capacities. With the correct capacities, the model should no longer have bottlenecks at the incorrect locations with too low or too high discharge rates. The upstream and downstream traffic volumes on each of the links should now better match the observed counts. Any remaining differences (assuming that error checking caught errors in the source node volumes) should be the result of route choice errors in the model. Therefore, the purpose of route choice calibration is to better match the counts at the non-bottleneck locations on the network.

The particular software selected for this example problem does not predict changes in route choice (no O-D table was provided for the model). Therefore, there is no route choice to be calibrated.[20] Assuming that the analyst is confident in the quality of the turn counts entered into the model (and converted inside the model into equivalent turn probabilities for routing traffic), any remaining errors in the link volumes can be corrected by manual adjustments to midblock source and sink node volumes.

In this example problem, the freeway, ramp, and intersection counts were all adjusted at the data preparation stage to produce an internally consistent set of volumes for every link and node. There are no unaccounted for increases or decreases in volumes between nodes. Thus, no midblock source/sink nodes were required. The model-predicted link volumes based on the source node volumes and the intersection turn percentages are consistent with the link counts (within the range of variation expected because of the random components of the simulation model).

System Performance Calibration

Once calibration has been completed to the analyst's satisfaction, the next step is to calibrate the overall performance of the model.

The model predictions were compared with the field data on speed, travel time, and delay on the freeways and arterials.

Based on this comparison, the free-flow speed distribution for the freeway was adjusted so that the model-generated free-flow speeds were within a more narrow range than the default distribution. This reflects commuter traffic behavior.

No other adjustments were made to the model parameters.

---

[20] Actually, the analyst could (in theory) calibrate the percentage turns that were input at each intersection and ramp junction to better match the observed link flows downstream from the source nodes; however, this is currently impractical (there are too many variables to be adjusted without the assistance of a good computerized optimization algorithm).

Ten repetitions of the calibrated data set were performed and the link output was processed to produce performance measures to be compared with the field data.

The comparison is shown in table 6. The simulated values represent the mean value of the MOEs based on the 10 model runs for the third time slice of the simulated peak hour.

Comparisons with field measurement of delay are not shown here 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 programs than in the approaches commonly used for field measurements.

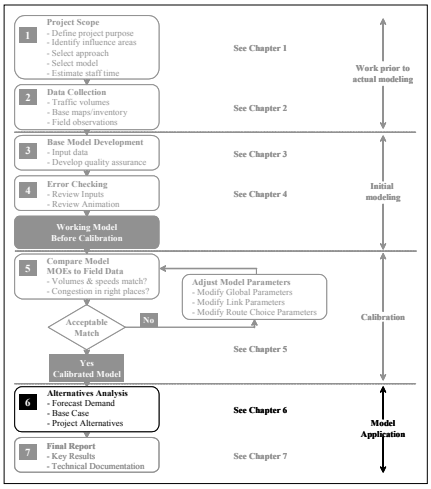
The results indicate that the model satisfies the criteria for calibration (shown in table 4).

Table 6. System performance results for example problem.

Freeway Segment	Travel Time (minutes)				Speed (mi/h)				Density (vehicle/mi/lane)			
	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 onto 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 onto 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%**

Arterial Segment	Travel Time (minutes)			
	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%**

6.0 Alternatives Analysis



Project alternatives analysis is the sixth task in the microsimulation analysis process. It is the reason for developing and calibrating the microsimulation model. The lengthy model development process has been completed and now it is time to put the model to work.

The analysis of project alternatives involves the forecasting of the future demand for the base case and the testing of various project alternatives against this baseline future demand. The analyst must run the model several times, review the output, extract relevant statistics, correct for biases in the reported results, and perform various analyses of the results. These analyses may include hypothesis testing, computation of confidence intervals, and sensitivity analyses to further support the conclusions of the analysis.

The alternatives analysis task consists of several steps:

1. Development of Baseline Demand Forecasts.
2. Generation of Project Alternatives for Analysis.
3. Selection of Measures of Effectiveness.
4. Model Application (Runs).
5. Tabulation of Results.
6. Evaluation of Alternatives.

6.1 Baseline Demand Forecast

This step consists of establishing the future level of demand to be used as a basis for evaluating project alternatives.

6.1.1 Demand Forecasting

Forecasts of future travel demand are best obtained from a travel demand model. These models require a great deal of effort and time to develop and calibrate. If one does not already exist, then the analyst may seek to develop demand forecasts based on historic growth rates. A trend-line forecast might be made, assuming that the recent percentage of growth in traffic will continue in the future. These trend-line forecasts are most reliable for relatively short periods of time (5 years or less). They do not take into account the potential of future capacity constraints to restrict the growth of future demand. Additional information and background regarding the development of traffic data for use in highway

planning and design may be found in National Cooperative Highway Research Program (NCHRP) Report 255, *Highway Traffic Data for Urbanized Area Project Planning and Design*.

6.1.2 Constraining Demand to Capacity

Regardless of which method is used to estimate future demand (regional model or trend line), care must be taken to ensure that the forecasts are a reasonable estimate of the actual amount of traffic that can arrive within the analytical period at the study area. Regional model forecasts are usually not well constrained to system capacity and trend-line forecasts are totally unconstrained. Appendix F provides a method for constraining future demands to the physical ability of the transportation system to deliver the traffic to the microsimulation model study area.

6.1.3 Allowance for Uncertainty in Demand Forecasts

All forecasts are subject to uncertainty. It is risky to design a road facility to a precise future condition given the uncertainties in the forecasts. There are uncertainties in both the probable growth in demand and the available capacity that might be present in the future. Slight changes in the timing or design of planned or proposed capacity improvements outside of the study area can significantly change the amount of traffic delivered to the study area during the analytical period. Changes in future vehicle mix and peaking can easily affect capacity by 10 percent. Similarly, changes in economic development and public agency approvals of new development can significantly change the amount of future demand. Thus, it is good practice to explicitly plan for a certain amount of uncertainty in the analysis. This level of uncertainty is the purpose of sensitivity testing (explained in a separate section below).

6.2 Generation of Alternatives

In this step, the analyst generates improvement alternatives based on direction from the decisionmakers and through project meetings. They will probably reflect operational strategies and/or geometric improvements to address the problems identified based on the baseline demand forecasts. The specifics of alternatives generation are beyond the scope of this report. Briefly, they consist of:

- Performing a microsimulation analysis[1] of the baseline demand forecast to identify the deficiencies.[2]
- Identifying alternative improvements that solve one or more of the identified problems.[3]

---

[1] A sketch-planning analysis using HCM techniques can be performed, if desired, to identify conceptual project alternatives prior to the development of a calibrated simulation model. The calibrated simulation model can then be used to refine the conceptual plans.

[2] Deficiencies are identified by comparing the measured performance against the project objectives and agency performance standards identified in task 1 (Project Scope).

6.3 Selection of Measures of Effectiveness (MOEs)

When reviewing model output, the analyst should look at a few key systemwide performance measures plus "hot spots" (indicators of localized breakdowns in the system).

MOEs are the system performance statistics that best characterize the degree to which a particular alternative meets the project objectives (which were determined in the Project Scope task). Thus, the appropriate MOEs are determined by the project objectives and agency performance standards rather than what is produced by the model. This section, however, discusses what output is typically produced by microsimulation software so that the analyst can appreciate what output might be available for constructing the desired MOEs.

Microsimulation, by its very nature, can bury the analyst in detailed microscopic output. The key is to focus on a few key indicators of system performance and localized breakdowns in the system (locations where queues are interfering with systems operation).

6.3.1 Candidate MOEs for Overall System Performance

As explained above, the selection of MOEs should be driven by the project objectives and the agency performance standards; however, many MOEs of overall system performance can be computed directly or indirectly from the following three basic system performance measures:

- Vehicle-miles traveled (VMT) or vehicle-kilometers traveled.
- Vehicle-hours traveled (VHT).
- Mean system speed.

These three basic performance measures can also be supplemented with other model output, depending on the objectives of the analysis. For example, total system delay is a useful overall system performance measure for comparing the congestion-relieving

---

[3] In support of the purpose and scope of the project, this identification of alternatives must take into consideration the range of demand conditions, weather conditions, incidents, and operational management strategies. Additional information and background may be found in the following documents:

Bunch, J., S. Hatcher, J. Larkin, G. Nelson, A. Proper, D. Roberts, V. Shah, and K. Wunderlich, *Incorporating ITS Into Corridor Planning: Seattle Case Study*, ITS Electronic Document Library (EDL)#11303, August 1999.

Wunderlich, K., *Incorporating ITS Into Planning Analyses: Summary of Key Findings From a Seattle 2020 Case Study*, Report No. FHWA-OP-02-031, May 2002.

effectiveness of various alternatives. The number of stops is a useful indicator for signal coordination studies.

6.3.2 Candidate MOEs for Localized Problems

In addition to evaluating overall system performance, the analyst should also be evaluating if and where there are localized system breakdowns ("hot spots"). A hot spot may be indicated by a persistent short queue that lasts too long, a signal phase failure (green time that fails to clear all waiting vehicles), or a blocked link (queue that backs up onto an upstream intersection).

A blocked link is the most significant indicator of localized breakdowns. A queue of vehicles that fills a link and blocks an upstream intersection can have a significant impact on system performance. A link queue overflow report can be developed to identify the links and times during the simulation period when the computed queue of vehicles equaled (and, therefore, probably actually exceeded) the storage capacity of the link.[4]

Signal phase failures, where the provided green time is insufficient to clear the queue, indicate potential operational problems if the queues continue to build over several cycles. The analyst should develop a signal phase failure report to indicate when and where signal green times are not sufficient to clear out all of the waiting queues of vehicles during each cycle.

At the finest level of detail, the analyst may wish to develop a report of the presence of persistent queues of a minimum length. This "hot spot" report would identify persistent queues on a lane-by-lane basis of a minimum number of vehicles that persist for a minimum amount of time. This report points the analyst to locations of persistent long queues (even those that do not overflow beyond the end of a link) during the simulation period.

6.3.3 Choice of Average or Worst Case MOEs

Microsimulation models employ random numbers to represent the uncertainty in driver behavior in any given population of drivers. They will produce slightly different results each time they are run, with a different random number seed giving a different mix of driver behaviors. The analyst needs to determine if the alternatives should be evaluated based on their average predicted performance or their worst case predicted performance. The average or mean performance is easy to compute and interpret statistically. The analyst runs the model several times for a given alternative, using different random number seeds each time. The results of each run are summed and averaged. The standard deviation of the results can be computed and used to determine the confidence interval for the results.

---

[4] This report will be less useful if the analyst has split long sections of roadway into arbitrary short links for other reporting purposes. The result may be many "false alarms" of blocked links that do not actually obstruct an upstream intersection.

The worst case result for each alternative is slightly more difficult to compute. It might be tempting to select the worst case result from the simulation model runs; however, the difficulty is that the analyst has no assurance that if the model were to be run a few more times, the model might not get an even worse result. Thus, the analyst never knows if he or she has truly obtained the worst case result.

The solution is to compute the 95th percentile probable worst outcome based on the mean outcome and an assumed normal distribution for the results.[5,6] The equation below can be used to make this estimate:

$$95\%\ Worst\ Result = m + 1.64 \bullet s \quad (5)$$

where:

m = mean observed result in the model runs

s = standard deviation of the result in the model runs

6.4 Model Application

The calibrated microsimulation model is applied in this step to compute the MOEs for each alternative. Although model operation is adequately described in the software user's guide, there are a few key considerations to be taken into account when using a microsimulation model for the analysis of alternatives.

6.4.1 Requirement for Multiple Repetitions

Microsimulation models rely on random numbers to generate vehicles, select their destination and route, and determine their behavior as they move through the network. No single simulation run can be expected to reflect any specific field condition. The results from individual runs can vary by 25 percent and higher standard deviations may be expected for facilities operating at or near capacity.[7] It is necessary to run the model several times *with different random number seeds*[8] to get the necessary output to determine

---

[5] Many statistical phenomena approximate a normal distribution at large sample sizes. Even though most microsimulation analysts usually work with relatively few model repetitions, the assumption of normal distribution is usually good enough for most analyses.

[6] Note that when computing the 95th percentile queue on the macroscopic level, it is typically assumed that the arrival of the vehicles are Poisson distributed. Microsimulation models predict the arrival patterns of vehicles, so the Poisson distribution assumption is not necessary when estimating 95th percentile queues using microsimulation data.

[7] See example in appendix B.

[8] Those with a more sophisticated statistical aptitude may elect to use variance reduction techniques that employ a single common random number seed to reduce the number of required repetitions. These techniques are described in Joshi, S.S., and A.K. Rathi, "Statistical Analysis and

(Footnote continued on next page...)

mean, minimum, and maximum values. The analyst must then post-process the runs to obtain the necessary output statistics (see appendix B for guidance on the computation of confidence intervals and the determination of the minimum number of repetitions of model runs).

6.4.2 Exclusion of Initialization Period

The initialization (warmup) period before the system reaches equilibrium for the simulation period should be excluded from the tabulated statistics (see appendix C for guidance on identifying the initialization period).

6.4.3 Avoiding Bias in the Results

The simulation geographic and temporal limits should be sufficient to include all congestion related to the base case and all of the alternatives. Otherwise, the model will not measure all of the congestion associated with an alternative, thus causing the analyst to underreport the benefits of an alternative. See the subsection below on the methods for correcting congestion bias in the results.

6.4.4 Impact of Alternatives on Demand

The analyst should consider the potential impact of alternative improvements on the base case forecast demand. This should take into consideration the effects of a geometric alternative, an operational strategy, and combinations of both. The analyst should then make a reasonable effort to incorporate any significant demand effects within the microsimulation analysis.[9]

6.4.5 Signal/Meter Control Optimization

Most simulation models do not currently optimize signal timing or ramp meter controls. Thus, if the analyst is testing various demand patterns or alternatives that significantly change the traffic flows on specific signalized streets or metered ramps, he or she may need to include a signal and meter control optimization substep within the analysis of each alternative. This optimization might be performed offline using a macroscopic signal timing or ramp metering optimization model. Or, the analyst may run the simulation model multiple times with different signal settings and manually seek the signal setting that gives the best performance.

Validation of Multi-Population Traffic Simulation Experiments," *Transportation Research Record 1510*, TRB, Washington, DC, 1995.

[9] The analyst might consult *Predicting Short-Term and Long-Term Air Quality Effects of Traffic-Flow Improvement Projects*, *Final Report*, NCHRP Project 25-21, TRB, Washington, DC, 2003.

6.5 Tabulation of Results

Microsimulation models typically produce two types of output: (1) animation displays and (2) numerical output in text files.[10] The animation display shows the movement of individual vehicles through the network over the simulation period. Text files report accumulated statistics on the performance of the network. It is crucial that the analyst reviews both numerical and animation output (not just one or the other) to gain a complete picture of the results.[11] This information can then be formatted for inclusion in the final report.

6.5.1 Reviewing Animation Output

Animation output is powerful in that it enables the analyst to quickly see and qualitatively assess the overall performance of the alternative. However, the assessment can only be qualitative. In addition, reviewing animation results can be time-consuming and tedious for numerous model repetitions, large networks, and long simulation periods. The analyst should select one or more model run repetitions for review and then focus his or her attention on the key aspects of each animation result.

Selection of Representative Repetition

The analyst has to decide whether he or she will review the typical case output, or the worst case output, or both. The typical case might give an indication of the average conditions for the simulation period. The worst case is useful for determining if the transportation system will experience a failure and for viewing the consequences of that failure.

The next question that the analyst must decide is how to identify which model repetition represents typical conditions and which repetition reflects worst case conditions. The total VHT may be a useful indicator of typical and worst case conditions. The analyst might also select other measures, such as the number of occurrences of blocked links (links with queue overflows) or delay.

If VHT is selected as the measure and the analyst wishes to review the typical case, then he or she would pick the model run repetition that had the total VHT that came closest to falling within the median of the repetitions (50 percent of the repetitions had a VHT less

---

[10] Some software programs also produce static graphs that can be very useful for gaining insight into the input or the results.

[11] Animation output shows the results from just one run of the simulation model. Drawing conclusions about traffic system performance from reviewing just one animation result is like trying to decide if the dice are fair from just one roll. One needs to roll the dice several times, tabulate the results, and compute the mean and standard deviation of the results to have the information needed to determine if the dice are fair.

than that, and 50 percent has more VHT). If the analyst wished to review the worst case, then he or she would select the repetition that had the highest VHT.

The pitfall of using a global summary statistic (such as VHT) to select a model run repetition for review is that overall average system conditions does not mean that each link and intersection in the system is experiencing average conditions. The median VHT repetition may actually have the worst performance for a specific link. If the analyst is focused on a specific link or intersection, then he or she should select some statistic related to vehicle performance on that specific link or intersection for selecting the model run repetition for review.

Review of Key Events in Animation

The key event to look for in reviewing animation is the formation of persistent queues. Cyclical queues at signals that clear each cycle are not usually as critical unless they block some other traffic movement. The analyst should not confuse the secondary impact of queues (one queue blocking upstream movement and creating a secondary queue) with the root cause of the queuing problem. Eliminating the cause of the first or primary queue may eliminate all secondary queuing. Thus, the analyst should focus on the few minutes just prior to formation of a persistent queue to identify the causes of the queuing.

6.5.2 Numerical Output

To avoid pitfalls, it is critical that the analyst understand how the results have been accumulated by the software.

Microsimulation software reports the numerical results of the model run in text output files called "reports." Unless the analyst is reviewing actual vehicle trajectory output, the output reports are almost always a summary of the vehicle activity simulated by the model. The results may be summarized over time and/or space. It is critical that the analyst understands how the software has accumulated and summarized the results to avoid pitfalls in interpreting the numerical output.

Microsimulation software may report instantaneous rates (such as speed) observed at specific instances of time, or may accumulate the data over a longer time interval and either report the sum, the maximum, or the average. Depending on the software program, vehicle activity that occurs between time steps (such as passing over a detector) may not be tallied, accumulated, or reported.

It is good practice to cross-check output to ensure that the analyst understands how it is computed by the software.

Microsimulation software may report the results for specific points on a link in the network or aggregated for the entire link. The point-specific output is similar to what would be reported by detectors in the field. Link-specific values of road performance are accumulated over the length of the link and, therefore, will vary from the point data.

The key to correctly interpreting the numerical output of a microsimulation model is to understand how the data were accumulated by the model and summarized in the report. The report headings may give the analyst a clue as to the method of accumulation used; however, these short headings cannot usually be relied on. The method of data accumulation and averaging can be determined through a detailed review of the model documentation of the reports that it produces, and, if the documentation is lacking, by querying the software developers themselves.

An initial healthy skepticism is valuable when reviewing reports until the analyst has more experience with the software. It helps to cross-check output to ensure that the analyst understands how the data is accumulated and reported by the software.

6.5.3 Correcting Biases in the Results

To make a reliable comparison of the alternatives, it is important that vehicle congestion for each alternative be accurately tabulated by the model. This means that congestion (vehicle queues) should not extend physically or temporally beyond the geographic or temporal boundaries of the simulation model. Congestion that overflows the time or geographic limits of the model will not normally be reported by the model, which can bias the comparison of alternatives.

The tabulated results should also exclude the initial and unrealistic initialization period when vehicles are first loaded on the network.

Ideally, the simulation results for each alternative would have the following characteristics:

- All of the congestion begins and ends within the simulation study area.
- No congestion begins or ends outside of the simulation period.
- No vehicles are unable to enter the network from any zone (source node) during any time step of the simulation.

It may not always be feasible to achieve all three of these conditions, so it may be necessary to adjust for congestion that is missing from the model tabulations of the results.

Correction of Output for Blocked Vehicles

If simulation alternatives are severely congested, then the simulation may be unable to load vehicles onto the network. Some may be blocked from entering the network on the periphery. Some may be blocked from being generated on internal links. These blocked vehicles will not typically be included in the travel time (VHT) or delay statistics for the model run.[12] The best solution is to extend the network back to include the maximum back

[12]The analyst should verify with the software documentation or developer how statistics on blocked vehicles are accumulated in the travel time and delay summaries.

of the queue. If this is not feasible, then the analyst should correct the reported VHT to account for the unreported delay for the blocked vehicles.

Microsimulation software will usually tally the excess queue that backs up outside the network as "blocked" vehicles (vehicles unable to enter the network) for each time step. The analyst totals the number of software-reported blocked vehicles for each time step of the simulation and multiplies this figure by the length of each time step (in hours) to obtain the vehicle-hours of delay. The delay resulting from blocked vehicles is added to the model-reported VHT for each model run.

Correction of Output for Congestion Extending Beyond the End of the Simulation Period

Vehicles queues that are present at the end of the simulation period may affect the accumulation of total delay and distort the comparison of alternatives (cyclical queues at signals can be neglected). The "build" project alternative may not look significantly better than the "no-build" option if the simulation period is not long enough to capture all of the benefits. The best solution is to extend the simulation period until all of the congestion that built up over the simulation period is served. If this is not feasible, the analyst can make a rough estimate of the uncaptured residual delay by computing how many vehicle-hours it would take to clear the queue using the equation given below:

$$VHT(Q) = \frac{Q^2}{2 \bullet C} \tag{6}$$

where:

VHT(Q) = extra VHT of delay attributable to a queue present at the end of the simulation period

Q = number of vehicles remaining in the queue at the end of the simulation period

C = discharge capacity of the bottleneck in veh/h

The equation computes the area of the triangle created by the queue and the discharge capacity after the end of the simulation period (see figure 11 below):

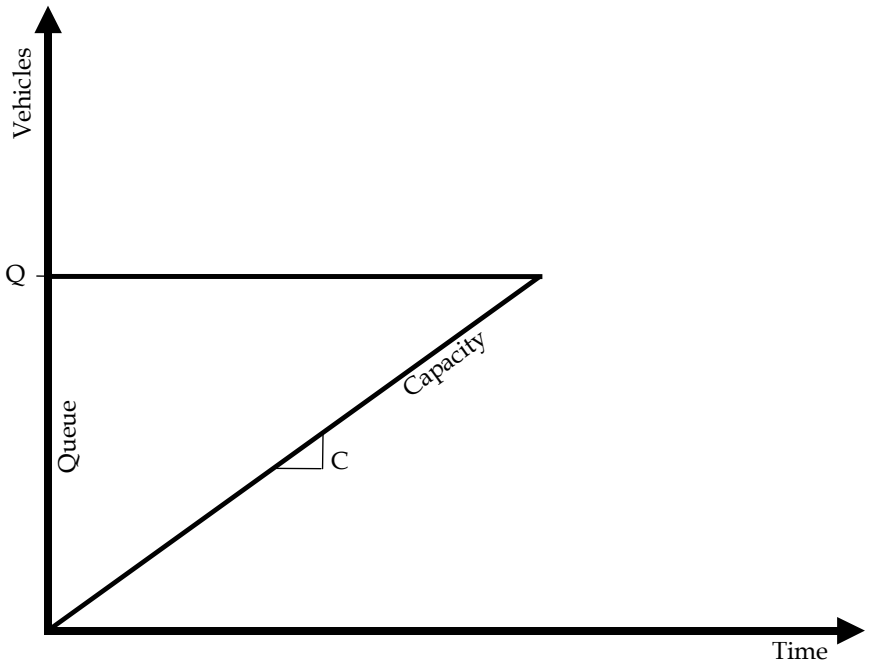


Figure 11. Computation of uncaptured residual delay at the end of the simulation period.

Note that this is not a complete estimate of the residual delay since it ignores the interaction of vehicles left over from the simulation period that interfere with traffic arriving during later time periods.

6.6 Evaluation of Alternatives

This step involves the evaluation of alternatives using the microsimulation model results. First, the interpretation of system performance results is discussed. Then, various analyses are discussed for assessing the robustness of the results. The ranking of alternatives and cost-effectiveness analyses are well documented in other reports and are not discussed here.

6.6.1 Interpretation of System Performance Results

This subsection explains how to interpret the differences between alternatives for the three basic system performance measures (VMT, VHT, and system speed).

VMT provides an indication of total travel demand (in terms of both the number of trips and the length of the trips) for the system.[13] Increases in VMT generally indicate increased demand (car, bus, and truck). VMT is computed as the product of the number of vehicles

[13] Person-miles traveled (PMT), if available, is the preferred measure of travel demand since it takes into account the number of people in each vehicle or mode.

traversing a link and the length of the link, summed over all links. Since VMT is computed as a combination of the number of vehicles on the system and their length of travel, it can be influenced both by changes in the number of vehicles and changes in the trip lengths during the simulation period. The following can cause changes in VMT between one alternative and the next:

- Random variations between one alternative and the next (use of the same random number seed in both runs can reduce or eliminate the variation).
- Changed demand.
- Changed congestion causing vehicles to take different paths (increased congestion may also reduce the number of vehicles that can complete their trip during the simulation period, also decreasing VMT).
- Inability of the model to load the coded demand onto the network within the simulation period (increased congestion may force the model to store some of the vehicle demand off-network because of bottlenecks at loading points; in this situation, increased congestion may actually lower VMT since stored vehicles cannot travel any distance during the simulation period).

VHT provides an estimate of the amount of time expended traveling on the system.[14] Decreases in VHT generally indicate improved system performance and reduced traveling costs for the public. VHT is computed as the product of the link volume and the link travel time, summed over all links. Since VHT is computed as a combination of the number of vehicles and the time spent traveling, it can be influenced both by changes in demand (the number of vehicles) and changes in congestion (travel time). Changes in VHT between one alternative and the next can be caused by the following:

- Random variations between one alternative and the next.
- Changed demand.
- Changed congestion.
- Demand stored off-network because of excessive congestion at load points (increased congestion that causes demand to be stored off-network may reduce VHT if the software does not accumulate delay for vehicles stored off-network).

Mean system speed is an indicator of overall system performance. Higher speeds generally indicate reduced travel costs for the public. The mean system speed is computed from the VMT and VHT as follows:

---

[14] Person-hours traveled (PHT), if available, provides a preferred measure of travel delay since it takes into account the number of people delayed in each vehicle. This is especially important for comparing the performance of HOV alternatives.

$$\text{Mean System Speed} = VMT/VHT \quad (7)$$

Changes in the mean system speed between one alternative and the next can be caused by the following:

- Random variations between one alternative and the next.
- Changed link speeds and delays caused by congestion.
- Changes in vehicle paths caused by congestion.
- Changes in vehicle demand.
- Changes in the number of vehicles stored off-network caused by excessive congestion at loading points.

Total system delay, if available, is useful because it reports the portion of total travel time that is most irritating to the traveling public. However, defining "total system delay" can be difficult. It depends on what the analyst or the software developer considers to be ideal (no delay) travel time. Some sources consider delay to include only the delay caused by increases in demand above some base uncongested (free-flow) condition. Others add in the base delay occurring at traffic control devices, even at low-flow conditions. Some include acceleration and deceleration delay. Others include only stopped delay. The analyst should consult the software documentation to ensure the appropriate use and interpretation of this measurement of system performance.

6.6.2 Hypothesis Testing

When the microsimulation model is run several times for each alternative, the analyst may find that the variance in the results for each alternative is close to the difference in the mean results for each alternative. How is the analyst to determine if the alternatives are significantly different? To what degree of confidence can the analyst claim that the observed differences in the simulation results are caused by the differences in the alternatives and not just the result of using different random number seeds? This is the purpose of statistical hypothesis testing. Hypothesis testing determines if the analyst has performed an adequate number of repetitions for each alternative to truly tell the alternatives apart at the analyst's desired level of confidence. Hypothesis testing is discussed in more detail in appendix E.

6.6.3 Confidence Intervals and Sensitivity Analysis

Confidence intervals are a means of recognizing the inherent variation in microsimulation model results and conveying them to the decisionmaker in a manner that clearly indicates the reliability of the results. For example, a confidence interval would state that the mean delay for alternative x lies between 35.6 s and 43.2 s, with a 95-percent level of confidence. If the 95-percent confidence interval for alternative y overlaps that of x, the decisionmaker

would consider the results to be less favorable for either alternative. They could be identical. Computation of the confidence interval is explained in appendix B.

A sensitivity analysis is a targeted assessment of the reliability of the microsimulation results, given the uncertainty in the input or assumptions. The analyst identifies certain input or assumptions about which there is some uncertainty and varies them to see what their impact might be on the microsimulation results.

Additional model runs are made with changes in demand levels and key parameters to determine the robustness of the conclusions from the alternatives analysis. The analyst may vary the following:

- Demand.
- Street improvements assumed to be in place outside of the study area.
- Parameters for which the analyst has little information.

A sensitivity analysis of different demand levels is particularly valuable when evaluating future conditions. Demand forecasts are generally less precise than the ability of the microsimulation model to predict their impact on traffic operations. A 10-percent change in demand can cause a facility to go from 95 percent of capacity to 105 percent of capacity, with a concomitant massive change in the predicted delay and queuing for the facility. The analyst should estimate the confidence interval for the demand forecasts and test the microsimulation at the high end of the confidence interval to determine if the alternative still operates satisfactorily at the potentially higher demand levels.

The analyst should plan for some selected percentage above and below the forecasted demand to allow for these uncertainties in future conditions. The analyst might consider at least a 10-percent margin of safety for the future demand forecasts. A larger range might be considered if the analyst has evidence to support the likelihood of greater variances in the forecasts.

To protect against the possibility of both underestimates and overestimates in the forecasts, the analyst might perform two sensitivity tests—one with 110 percent of the initial demand forecasts and the other with 90 percent of the initial demand forecasts—for establishing a confidence interval for probable future conditions.[15]

Street improvements assumed to be in place outside the simulation study area can also have a major impact on the simulation results by changing the amount of traffic that can

[15]Note that the percentage confidence interval (such as a 95 percent confidence interval) has not been stated here, so it cannot be claimed that there is a certain probability of the true value falling within this 10 percent range. This is merely a sensitivity test of the impact of the demands being 10 percent lower or 10 percent higher than that forecast, without knowing the likelihood of it actually happening.

enter or exit the facilities in the study area. Sensitivity testing would change the assumed future level of demand entering the study area and the assumed capacity of facilities leaving the study area to determine the impact of changes in the assumed street improvements.

The analyst may also run sensitivity tests to determine the effects of various assumptions about the parameter values used in the simulation. If the vehicle mix was estimated, variations in the percentage of trucks might be tested. The analyst might also test the effects of different percentages of familiar drivers in the network.

6.6.4 Comparison of Results to the HCM

It is often valuable when explaining microsimulation model results to the general public to report the results in terms of HCM levels of service. However, the analyst should be well aware of the differences between the HCM and the microsimulation analysis when making these comparisons.

Delay and Intersection Level of Service (LOS)

Delay is used in the HCM to estimate the LOS for signalized and unsignalized intersections. There are distinctions in the ways microsimulation software and the HCM define delay and accumulate it for the purpose of assessing LOS.

The HCM bases its LOS grades for intersections on estimates of mean control delay for the highest consecutive 15-min period within the hour. If microsimulation output is to be used to estimate LOS, then the results for each run must be accumulated over a similar 15-consecutive-minute time period and averaged over several runs with different random number seeds to achieve a comparable result.

This still may not yield a fully comparable result, because all microsimulation models assign delay to the segment in which it occurs. For example, the delay associated with a single approach to a traffic signal may be parceled out over several upstream links if the queues extend beyond one link upstream from the intersection. Thus, when analysts seek to accumulate the delay at a signal, they should investigate whether the delay/queues extend beyond the single approach links to the signal.

Finally, the HCM does not use total delay to measure signal LOS. It uses “control delay.” This is the component of total delay that results when a control signal causes a lane group to reduce speed or to stop. It is measured by comparison with the uncontrolled condition. The analyst needs to review the software documentation and seek additional documentation from the software vendor to understand how delay is computed by the software.

Density and Freeway/Highway LOS

If microsimulation model reports of vehicle density are to be reported in terms of their LOS implications, it is important to first translate the densities reported by the software into the densities used by the HCM to report LOS for uninterrupted flow facilities.[16]

HCM 2000 defines freeway and highway LOS based on the average density of passenger car equivalent vehicles in a section of highway for the peak 15-min period within an hour. For ramp merge and diverge areas, only the density in the rightmost two lanes is considered for LOS. For all other situations, the density across all lanes is considered. Trucks and other heavy vehicles must be converted to passenger car equivalents using the values contained in the HCM according to vehicle type, facility type, section type, and grade.

Queues

HCM 2000 defines a queue as: "A line of vehicles, bicycles, or persons waiting to be served by the system in which the flow rate from the front of the queue determines the average speed within the queue. Slowly moving vehicles or people joining the rear of the queue are usually considered part of the queue." These definitions are not implementable within a microsimulation environment since "waiting to be served" and "slowly" are not easily defined. Consequently, alternative definitions based on maximum speed, acceleration, and proximity to other vehicles have been developed for use in microsimulation.

Note also that for most microsimulation programs, the number of queued vehicles counted as being in a particular turn-pocket lane or through lane *cannot exceed the storage capacity of that lane*. Any overflow is reported for the upstream lane and link where it occurs, not the downstream cause of the queue. Unlike macroscopic approaches that assign the entire queue to the bottleneck that causes it, microsimulation models can only observe the presence of a queue; they currently do not assign a cause to it. So, to obtain the 95-percent queue length, it may be necessary to temporarily increase the length of the storage area so that all queues are appropriately tallied in the printed output.

6.7 Example Problem: Alternatives Analysis

The same example problem from the previous chapters is continued here. The task now is to apply the calibrated model to the analysis of the ramp metering project and its alternatives.

Step 1: Baseline Demand Forecast

A 5-year forecast was estimated using a straight-line growth approach assuming 2 percent growth per year uncompounded. The result was a forecasted 10-percent increase in traffic

[16]Note that density is NOT used as an LOS measurement for interrupted flow facilities, such as city streets with signals and intersections with stop signs.

demand for the corridor. The forecasted growth for individual links and ramps varied from this average value.

Since the existing conditions were uncongested and the growth forecast is a modest 10 percent growth, the forecasted demand was not constrained because of anticipated capacity constraints on the entry links to the corridor.

Since a 5-year forecast is being performed, it was considered to be fairly reliable for such a short time period. No extra allowance was added to the forecasts or subtracted from them to account for uncertainty in the demand forecasts.

Step 2: Generation of Alternatives

Two alternatives will be tested with the calibrated model—no-build and build. The build alternative consists of ramp metering on the two eastbound freeway on-ramps. The no-build alternative has no ramp metering. Figure 12 below illustrates the coding of one of the ramp meters.

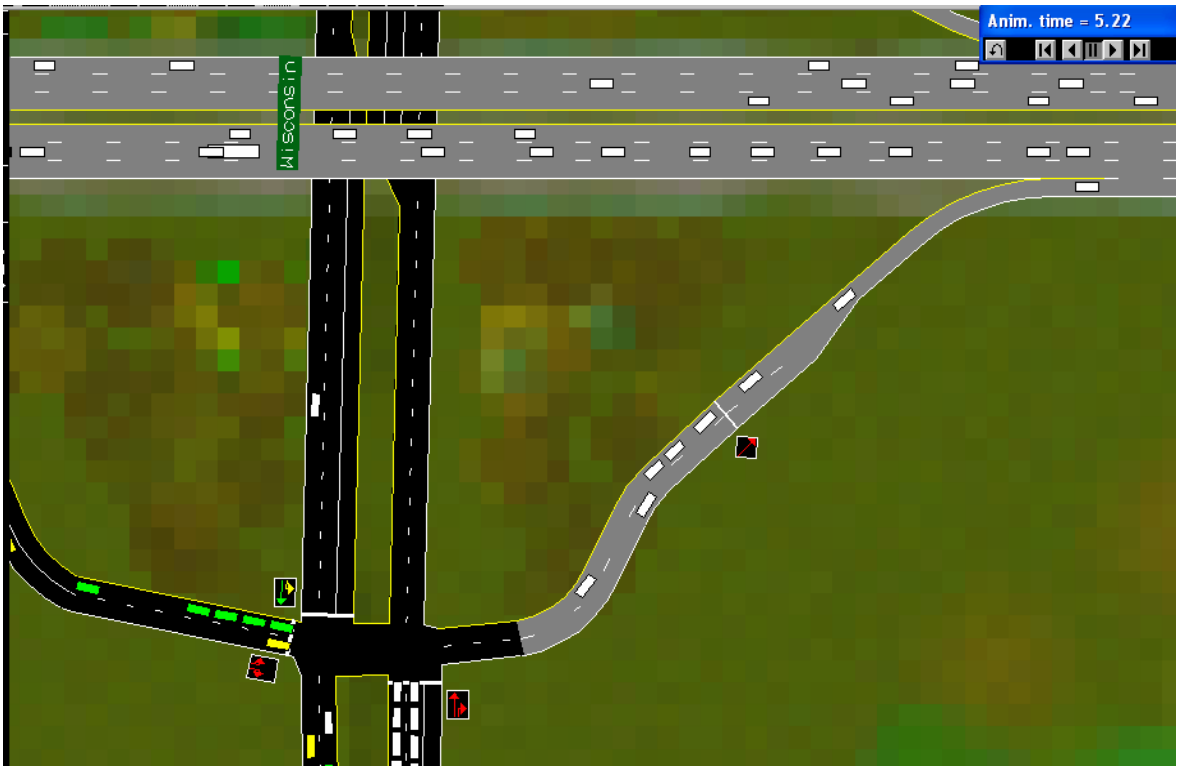


Figure 12. Ramp meter geometry.

Step 3: Selection of MOEs

The following system MOEs were selected for evaluation of the alternatives: VMT, VHT, and delay (vehicle-hours). The selected indicator of localized problems was a "blocked link," indicating that the queue filled up and overflowed the available storage in the link.

It was opted to report the mean results rather than the 95-percent worst case results.

Step 4: Model Application

The model was run 10 times for each alternative. The results were output into a spreadsheet and averaged for each alternative.

The impact of ramp meters on route choice was estimated outside of the model using a regional travel demand model to predict the amount of diversion. The regional model predicted that diversion was implemented in the simulation model by manually adjusting the turn percentages at the appropriate upstream intersections and ramp junctions.

The initialization period was automatically excluded from the tabulated results by the selected software program.

The results were reviewed to determine if building queues[17] were extending beyond the physical boundaries of the model or the temporal boundaries of the analytical period. None was found, so it was not necessary to correct the model results for untabulated congestion.

Because of the modest differences in congestion between the alternatives, induced demand was not considered to be a significant factor in this analysis. No adjustments were made to the baseline demand forecasts.

Signal/meter control optimization was performed outside of the model using macroscopic signal optimization software and ramp meter optimization software. The recommended optimal settings were input into the simulation model. Separate optimizations were performed for the no-build and build alternatives.

Step 5: Tabulation of Results

The model results for 10 repetitions of each alternative were output into a spreadsheet and averaged for each alternative (see table 7 below). A review of the animation output indicated that post-model corrections of untallied congestion were not necessary.[18]

---

[17] Recurrent queues at signals that occur during each cycle are not considered to be building queues indicative of unserved demand that might bias the results for an alternative.

[18] No increasing queues indicating underserved demand were found.

Table 7. Summary of analytical results.

Measure of Effectiveness	Existing	Future	
		No-Build	Build
VMT			
Freeway	35,530	39,980	40,036
Arterial	8,610	9,569	9,634
Total	44,140	49,549	49,670
VHT			
Freeway	681.6	822.3	834.2
Arterial	456.5	538.5	519.5
Total	1138.1	1360.8	1353.7
Delay (VHT)			
Freeway	33.1	90.4	101.0
Arterial	214.3	269.7	248.9
Total	247.4	360.1	349.9

Step 6: Evaluation of Alternatives

Under the no-build scenario, the total delay on the corridor increased by 46 percent over existing conditions. The VMT increased by 12 percent and the total travel time increased by 20 percent. Most of the delay increases were on the freeway mainline links.

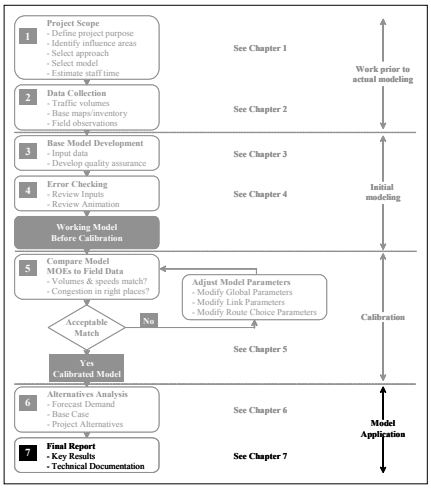
Under the improved scenario (ramp metering plus signal optimization), systemwide delay was reduced by about 3 percent (from the no-build scenario) with a slight increase in VMT.[19] Freeway mainline traffic conditions improved at the expense of the on-ramp traffic.

The improvements are operationally acceptable (no spillbacks from the ramp meters to the arterial network).

[19]The improvement is modest enough that it might be worth establishing confidence intervals for the results and performing some sensitivity analysis and hypothesis tests to confirm the robustness of the conclusion.

7.0 Final Report

This chapter discusses the documentation of the microsimulation analysis results in a final report and the technical report or appendix supporting the final report.



7.1 Final Report

The final report presents the assumptions, analytical steps, and results of the analysis in sufficient detail for decisionmakers to understand the basis for and implications of choosing among the project alternatives. The final report, however, will not usually contain sufficiently detailed information to enable other analysts to reproduce the results. That is the purpose of the technical report/appendix.

The effort involved in summarization of the results should not be underestimated, since microsimulation models produce a wealth of numerical output that must be tabulated and summarized. The final report should include the following:

1. Study Objectives and Scope.
2. Overview of Study Approach (tools used, methodology, rationale).
3. Data Collection (sources and methods).
4. Calibration Tests and Results (which parameters were modified and why).
5. Forecast Assumptions (assumed growth inside and outside of the study area, street improvements, etc.).
6. Description of Alternatives.
7. Results.

7.2 Technical Report/Appendix

The technical report/appendix documents the microsimulation analysis in sufficient detail to enable an analyst to reproduce the results (the version or release of the software used is included). It may be an appendix to the final report or a separate document.

The technical report/appendix is a vital step in preserving the rationale for the various decisions that were made in the process of developing, calibrating, and operating a microsimulation model. The documentation should be sufficient so that given the same

input files, another analyst can understand the calibration process and repeat the alternatives analysis.

The technical report/appendix should include the model input and output files (in electronic format) for the final model calibration run and alternatives analysis model runs. In addition to a diskette with the calibration and alternatives analysis files (input and output), the technical report/appendix should include a printed listing of the files on the diskette with a text description of the contents and purpose of each file.

Appendix A: Traffic Microsimulation Fundamentals

This section provides an overview of the technical aspects of microsimulation. It identifies and defines the terminology typically used, and what model types and procedures are generally employed in microsimulation. This section also provides a general introduction to traffic simulation models. Readers interested in additional details on the technical aspects of microsimulation may consult the *Revised Monograph on Traffic Flow Theory* (Chapter 10: Traffic Simulation) or other documents.

A.1 Overview

Simulation models are designed to emulate the behavior of traffic in a transportation network over time and space to predict system performance. Simulation models include the mathematical and logical abstractions of real-world systems implemented in computer software. Simulation model runs can be viewed as experiments performed in the laboratory rather than in the field.

A.2 System Update Mechanism

Traffic simulation models describe the changes in the system state through discrete intervals in time. There are generally two types of models, depending on whether the update time intervals are fixed or variable:

Discrete Time (Time-Scan) Models: The system is being updated at fixed-time intervals. For example, the model calculates the vehicle position, speed, and acceleration at 1-s intervals. The choice of the update time interval depends on how accurately the system needs to be simulated at the expense of computer-processing time. Most microsimulation models employ a 0.1-s resolution for updating the system. Traffic systems are typically modeled using time-scan models because they experience a continuous change in state.

Discrete Event (Event-Scan) Models: In these models, the time intervals vary in length and correspond to the intervals between events. For example, a pre-timed traffic signal indication (e.g., green) remains constant for 30 s until its state changes instantaneously to yellow. The operation of the signal is described by recording its changes in state when events occur, rather than monitoring the state of the signal each second. Event-scanning models typically achieve significant reductions in computer run time. However, they are suitable for simulating systems whose states change infrequently.

A.3 Traffic Stream Representation

Simulation models are typically classified according to the level of detail at which they represent the traffic stream. These include:

Microscopic Models: These models simulate the characteristics and interactions of individual vehicles. They essentially produce trajectories of vehicles as they move through the network. The processing logic includes algorithms and rules describing how vehicles move and interact, including acceleration, deceleration, lane changing, and passing maneuvers.

Mesoscopic Models: These models simulate individual vehicles, but describe their activities and interactions based on aggregate (macroscopic) relationships. Typical applications of mesoscopic models are evaluations of traveler information systems. For example, they can simulate the routing of individual vehicles equipped with in-vehicle, real-time travel information systems. The travel times are determined from the simulated average speeds on the network links. The average speeds are, in turn, calculated from a speed-flow relationship.

Macroscopic Models: These models simulate traffic flow, taking into consideration aggregate traffic stream characteristics (speed, flow, and density) and their relationships. Typically, macroscopic models employ equations on the conservation of flow and on how traffic disturbances (shockwaves) propagate in the system. They can be used to predict the spatial and temporal extent of congestion caused by traffic demand or incidents in a network; however, they cannot model the interactions of vehicles on alternative design configurations.

Microscopic models are potentially more accurate than macroscopic simulation models. However, they employ many more parameters that require calibration. Also, the parameters of the macroscopic models (e.g., capacity) are observable in the field. Most of the parameters of the microscopic models cannot be observed directly in the field (e.g., minimum distances between vehicles in car-following situations).

A.4 Randomness in Traffic Flow

Simulation models are also classified by how they represent the randomness in the traffic flow, including:

Deterministic Models: These models assume that there is no variability in the driver-vehicle characteristics. For example, it is assumed that all drivers have a critical gap of 5 s in which to merge into a traffic stream, or all passenger cars have a vehicle length of 4.9 m (16 ft).

Stochastic Models: These models assign driver-vehicle characteristics from statistical distributions using random numbers. For example, the desired speed of a vehicle is randomly generated from an assumed normal distribution of desired speeds, with a mean of 105 km/h (65 mi/h) and a standard deviation of 8 km/h (5 mi/h). Stochastic

simulation models have routines that generate random numbers. The sequence of random numbers generated depends on the particular method and the initial value of the random number (random number seed). Changing the random number seed produces a different sequence of random numbers, which, in turn, produces different values of driver-vehicle characteristics.

Stochastic models require additional parameters to be specified (e.g., the form and parameters of the statistical distributions that represent the particular vehicle characteristic). More importantly, the analysis of the simulation output should consider that the results from each model run vary with the input random number seed for otherwise identical input data. Deterministic models, in contrast, will always produce the same results with identical input data.

A.5 Microsimulation Process

Microsimulation models employ several submodels, analytical relationships, and logic to model traffic flow. Detailed descriptions of each submodel are beyond the scope of this section. Instead, this document focuses on some key aspects of the simulation process that will probably affect the choice of the particular tool to be used and the accuracy of the results.

Simulation models include algorithms and logic to:

- Generate vehicles into the system to be simulated.
- Move vehicles into the system.
- Model vehicle interactions.

A.5.1 Vehicle Generation

Generating Vehicles Into the Traffic Stream

At the beginning of the simulation run, the system is *empty*. Vehicles are generated at the entry nodes of the analytical network, based on the input volumes and an assumed headway distribution. Suppose that the specified volume is V = 600 veh/h for a 15-min analytical period and that the model uses a uniform distribution of vehicle headways. Then, a vehicle will be generated at time intervals:

$$H = \text{mean headway} = 3600/V = 3600/600 = 6 \text{ s} \quad (8)$$

If the model uses the shifted negative exponential distribution to simulate the arrival of vehicles at the network entry node instead of the uniform distribution, then vehicles will be generated as time intervals:

$$h = (H - h_{min}) [- \ln(1 - R)] + H - h_{min} \quad (9)$$

where:

h = headway (in seconds) separating each generated vehicle

h_{min} = specified minimum headway (e.g., 1.0 s)

R = random number (0 to 1.0)

Generating Driver-Vehicle Attributes

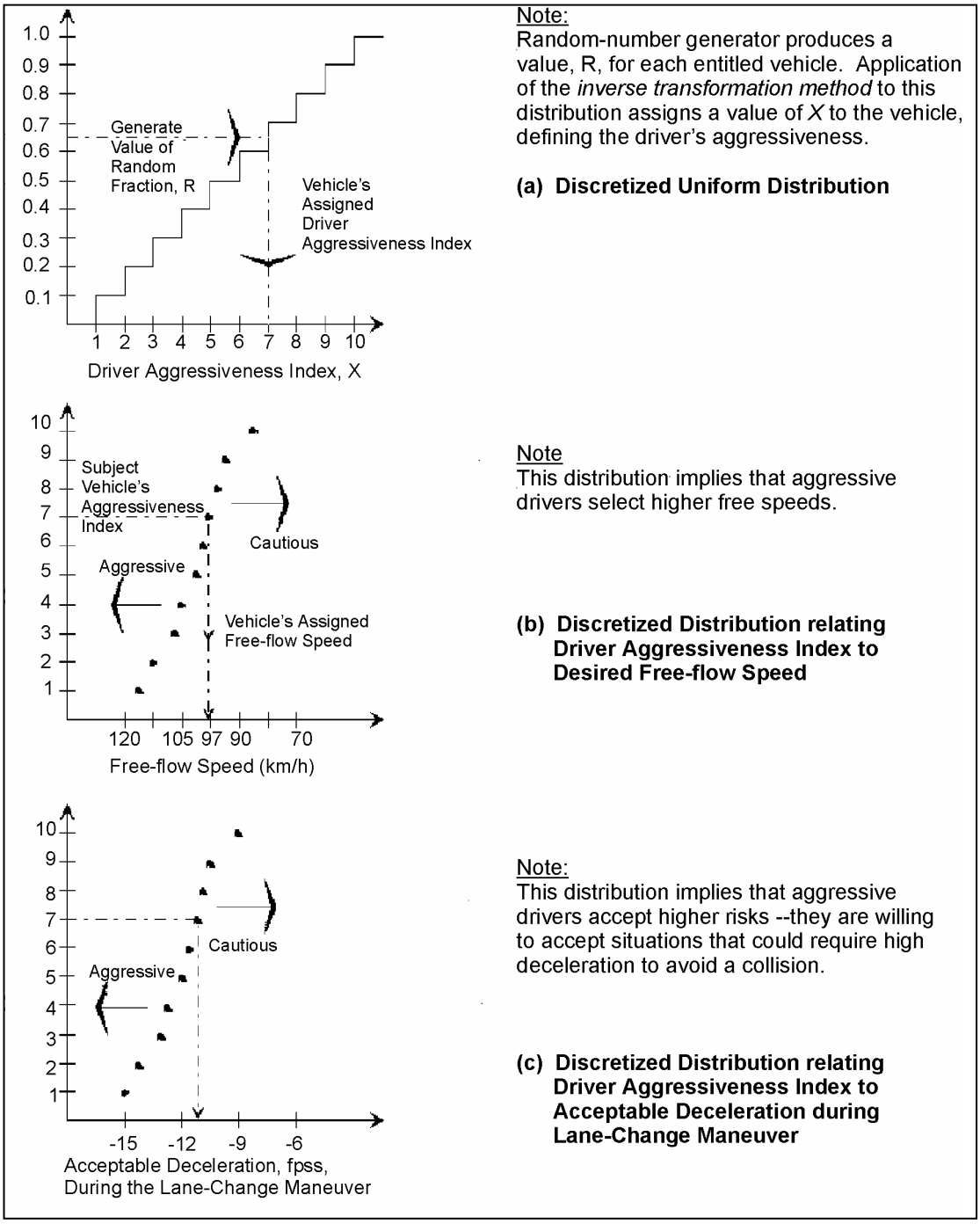
When a vehicle is generated at the entry of the network, the simulation model assigns driver-vehicle characteristics. The following characteristics or attributes are commonly generated for each driver-vehicle unit:

Vehicle: Type (car, bus, truck), length, width, maximum acceleration and deceleration, maximum speed, maximum turn radius, etc.

Driver: Driver aggressiveness, reaction time, desired speed, critical gaps (for lane changing, merging, crossing), destination (route), etc.

Note that the different models may employ additional attributes for each driver-vehicle unit to ensure that the model replicates real-world conditions. Each attribute may be represented in the model by constants (e.g., all passenger cars have a vehicle length of 4.9 m (16 ft)), functional relationships (e.g., maximum vehicle acceleration is a linear function of its current speed), or probability distributions (e.g., driver's desired speed is obtained from a normal distribution). Most microsimulation models employ statistical distributions to represent the driver-vehicle attributes. The statistical distributions employed to represent the variability of the driver-vehicle attributes and their parameters must be calibrated to reflect local conditions.

Figure 13 illustrates a generic process for generating driver-vehicle attributes for a stochastic microscopic simulation model. Drivers are randomly assigned an "aggressiveness index" ranging from 1 (very aggressive) to 10 (very cautious), drawn from a uniform distribution to represent the range of human behavior. Once the value of the aggressiveness has been assigned (X = 7 in this example), it is used to assign the value of the free-flow speed (97 km/h) and the acceptable deceleration (-3.4 m/s^2 (-11 ft/s^2)) from the corresponding distribution of speeds and accelerations.



Source: *Traffic Flow Theory Monograph* (Chapter 10: Simulation)

Figure 13. Generation of driver characteristics.

A.5.2 Vehicle Movement

This section briefly describes the process for simulating vehicle movement and how it is impacted by the physical environment. The physical environment—the transportation network under study—is typically represented as a network of links and nodes. Links are one-way roadways with fixed design characteristics and nodes that represent intersections or locations where the design characteristics of the links change. Simulation models have different limits on the network size (maximum number of links and nodes).

Vehicles, in the absence of impedance from other vehicles, travel at their desired speed on the network links. However, their speed may be affected by the link-specified geometry (horizontal and vertical alignment), pavement conditions, and other factors. For example, the simulation model computes the actual vehicle speed as the minimum value from the desired speed and the speed computed for the specified vertical and horizontal alignment. However, not all microsimulation tools model the sensitivity of vehicle speeds to link design characteristics.

Vehicles proceed through the network until they exit the system at their destination. Typically, there are two types of simulation models: turning-fraction based or O-D based. In O-D-based models, the O-D matrix is input, and when a vehicle is generated at an origin, it is assigned its destination. The vehicle then exits the network at the specified destination. In turning-fraction-based models, the vehicle destination is randomly assigned at the entry of the link, based on specified turning volumes (or fractions) at the downstream end of the link. For example, for a vehicle entering a signalized intersection approach, the vehicle destination (going through or turning) is randomly determined based on the input turning-volume fractions for the particular link. This also implies that turning-fraction models are not well suited for tracing the performance of individual vehicles throughout the network and evaluating the effectiveness of certain ITS options (e.g., a fraction of vehicles being equipped with real-time information systems).

Simulation models employ a number of approaches to guide vehicles within the network. They typically employ warning signs to advise the simulated vehicle to change lanes because it needs to exit at the downstream off-ramp, its lane is ending, or there is a blockage downstream. The location of the warning signs may significantly affect the accuracy of the simulation. For example, vehicles traveling on the freeway at a speed of 97 km/h (60 mi/h) (27 m/s (88 ft/s)) may miss their exit without this advance warning if they are traveling in the median lane of a multilane freeway and are required to make multiple lane changes to exit at the downstream end of a short link.

Vehicle Interactions

Microscopic models simulate the interactions of individual vehicles as they travel in the analytical network using car-following, lane-changing, and queue discharge algorithms.

Car Following

The interaction between a leader and follower pair of vehicles traveling in the same lane, is generally assumed to be a form of stimulus-response mechanism:

$$\text{Response follower}_{(t+T)} = (\text{sensitivity}) \bullet (\text{stimulus})_{t} \tag{10}$$

where:

T = reaction time (time lag) for the response of the following vehicle

Car-following models for highway traffic were proposed since the inception of the traffic-flow theory in the early 1950s. Extensive experimentation and theoretical work were performed at the General Motors (GM) laboratories and elsewhere. Most of the existing simulation models employ fail-safe car-following algorithms. Such algorithms simulate car-following behavior by maintaining a minimum safe distance between vehicles subject to stochastically generated vehicle characteristics (e.g., maximum acceleration and deceleration rates). The fail-safe car-following algorithms currently implemented in most simulation models consist of the following components:

1. An equation that calculates the acceleration (deceleration) of the following vehicle in response to the motion of its leader to maintain a *target headway (spacing)*, depending on the driver-vehicle characteristics.

   $$a_f = F(v_l, v_f, s, T, X_i) \tag{11}$$

   where:

   a_f = acceleration of the following vehicle after a reaction time T

   v_l and v_f = speeds of the leading and following vehicles, respectively

   s = distance between vehicles

   X_i = parameters specific to the particular car-following model

2. The computed acceleration rate above must not exceed the maximum acceleration rate for the specific vehicle type and must not result in a higher speed than the vehicle's desired speed.

3. Furthermore, the computed acceleration of the follower above must satisfy the "safe-following" rule (i.e., the follower should always maintain a *minimum separation* from the leader). If the value of the safe-following acceleration is smaller then the car-following acceleration computed above, the former is implemented.

Lane Changing

The modeling of lane changing is based on the gap-acceptance process. A vehicle may change lanes if the available gap in the target lane is greater than its critical gap. Typically, three types of lane changes are modeled:

1. A **mandatory lane change** occurs when a vehicle is required to exit the current lane. Merging from an on-ramp onto the freeway is an example of a mandatory lane change. A vehicle *must* execute a lane change under several circumstances:

    a. Current lane is ending (lane drop)

    b. Must move to the outer lane(s) to exit the freeway via an off-ramp

    c. Current lane will not accommodate its movement or vehicle type (e.g., HOV lane)

    d. Lane is blocked because of incidents or other events

2. A **discretionary lane change** occurs when a vehicle changes lanes to improve its position (i.e., travel at the desired speed). If a driver is impeded in the current lane because of a slower moving vehicle in front, he or she may consider changing lanes to improve his or her position (i.e., resume his or her desired speed). The lane-changing logic determines which of the adjacent lanes (if more than one is available) is the best candidate. To be a candidate, the lane's channelization must accommodate the subject vehicle, be free of blockage, and not end close by. Lane changing is performed subject to the availability of gaps in the candidate lane and the acceptability of the gaps by the subject vehicle.

    On a surface-street network, discretionary lane changes may occur because of various impedance factors perceived by the driver, including:

    a. Queue length at the stop line.

    b. Heavy truck traffic.

    c. Bus transit operations.

    d. Turning movements from shared lanes.

    e. Pedestrian interference with turning vehicles.

    f. Parking activities.

3. An **anticipatory lane change** occurs when a vehicle may change lanes in anticipation of slowdowns in its current lane further downstream as a result of merging or weaving. The decision to change lanes is based on the difference in speed *at the location* of the merge or weave, between the vehicles in the current lane and the adjacent lane not directly involved in the merge or weave. The lane-changing logic recognizes that a driver may accelerate or decelerate to create acceptable gaps to change lanes. Consider a vehicle A that desires to merge into a gap between vehicles B and C (B currently is the leader of C). Vehicle A will accept the gap if the time headway between it and vehicle B is greater than some critical value g(A,B) and the time headway

between it and vehicle C is greater than some critical value g(A,C). However, the critical time headway values are not constant, but are dependent on the following considerations:

- Critical time headways depend on vehicle speeds. Vehicle A will accept a smaller critical headway if it is going slower than vehicle B.

- Lane changing takes place over a finite period of time. During this time period, the driver who is changing lanes can adjust his or her position with respect to the new leader by decelerating.

- The new follower may cooperate with the driver who is changing lanes by decelerating to increase the size of the gap.

These considerations are typically combined in a *risk* measure. More aggressive drivers would accept a higher risk value (i.e., shorter gaps and higher acceleration/ deceleration rates) to change lanes. Moreover, the risk value may be further increased, depending on the type of lane change and the situation. For example, a merging vehicle reaching the end of the acceleration lane may accept much higher risk values (forced lane changes).

The model's lane-changing logic and parameters have important implications for traffic performance, especially for freeway merging and weaving areas. In addition, the time assumed for completion of the lane-change maneuver affects traffic performance because the vehicle occupies both lanes (original and target lanes) during this time interval. Furthermore, one lane change is allowed per simulation time interval.

Appendix B: Confidence Intervals

This section explains how to compute the confidence intervals for the microsimulation model results. It explains how to compute the minimum number of repeated field measurements or microsimulation model runs needed to estimate the mean with a certain level of confidence that the true mean actually falls within a target interval.

Three pieces of information are required: (1) sample standard deviation, (2) desired length of the confidence interval, and (3) desired level of confidence.

The required number of repetitions must be estimated by an iterative process. A preliminary set of repetitions is usually required to make the first estimate of the standard deviation for the results. The first estimate of the standard deviation is then used to estimate the number of repetitions required to make statistical conclusions about alternative highway improvements.

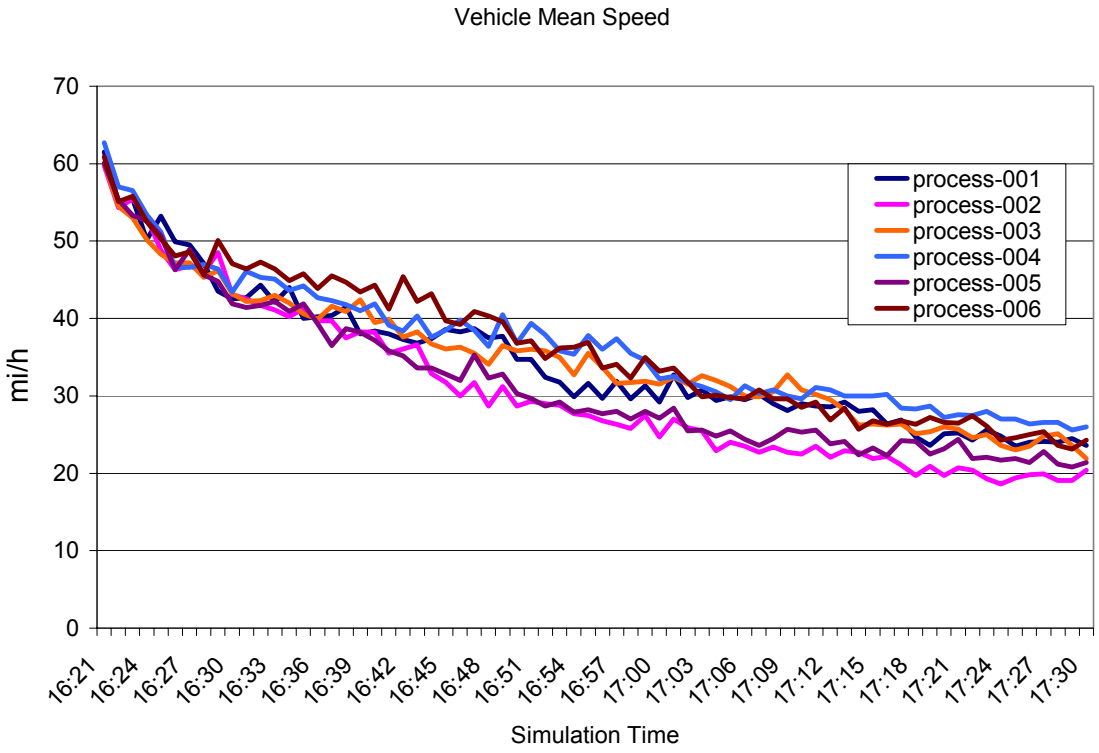
When all of the repetitions have been completed, the analyst then recomputes the standard deviation and the required number of repetitions based on all of the completed repetitions. If the required number of repetitions is less than or equal to the completed number of repetitions, then the analysis has been completed. If not, the analyst either relaxes the desired degree of confidence in the results or performs additional repetitions.

B.1 Why Are Multiple Model Runs Required?

Multiple repetitions of the same model are required because microsimulation results will vary depending on the random number seed used in each run. The random number seed is used to select a sequence of random numbers that are used to make numerous decisions throughout the simulation run (Should a vehicle be loaded on the network now or later? Should the driver be aggressive or timid? Which vehicle type should it be? Will the driver choose the shortest path or a slightly longer one instead?). The outcomes of all of these decisions will affect the simulation results. The results of each run will usually be close to the average of all of the runs; however, each run will be different from the other.

Microsimulation model results can vary by 25 percent between runs.

Figure 14 shows the range of results that can occur simply by varying the random number seed used in each run. The figure shows the mean system vehicle speed for a freeway-to-freeway interchange during each minute of the simulation run. A total of six simulation runs (called "processes" in the figure) were made. As can be seen, after the initial warmup period, results diverge. The range of mean speeds output by each process is about 8.1 to 9.7 km/h (5 to 6 mi/h) for most of the length of the simulation period. Running a longer simulation period does not reduce this variation. At the end of the simulation, the range of results is about 9.7 km/h (6 mi/h), or 25 percent of the mean system speed at the end of the simulation period.



Source: I-580/I-680 Freeway Interchange Model, Dowling Associates, 2002

1 mi/h = 1.61 km/h

Figure 14. Variation in the results between repetitions.

B.2 Estimation of Sample Standard Deviation

The standard deviation (an estimate of the variance) is required to estimate the number of repetitions. However, it takes a minimum number of repetitions to estimate the standard deviation in the first place. So either the analyst estimates the standard deviation directly (based on past experience) or executes a few model run repetitions (each run using a different random number seed) and uses the equation below to compute an initial estimate of the sample standard deviation.

$$s^2 = \frac{\sum\left(x - \bar{x}\right)^2}{N-1} \tag{12}$$

where:

s = standard deviation

x = variable (such as delay) for which the sample variance is desired

$\bar{x}$ = average value of the variable produced by the model runs

N = number of model runs

Unless the analyst already knows the standard deviation from experience, it is recommended that four repetitions be performed for the initial estimation of the standard deviation. This initial estimate is then revisited and revised later if and when additional repetitions are performed for the purposes of obtaining more precise estimates of mean values or for alternatives analysis.

B.3 Selection of Desired Confidence Level

The confidence level is the probability that the true mean lies within the target confidence interval. The analyst must decide to what degree he or she wishes to know the interval in which the true mean value lies. The usual approach is to pick a 95-percent level of confidence; however, analysts may choose higher or lower levels of confidence. Higher levels of confidence require more repetitions.

B.4 Selection of Desired Confidence Interval

The confidence interval is the range of values within which the true mean value may lie. The length of the interval is at the discretion of the analyst and may vary according to the purposes for which the results will be used. For example, if the analyst is testing alternatives that are very similar, then a very small confidence interval will be desirable to distinguish between the alternatives. If the analyst is testing alternatives with greater differences, then a larger confidence interval can be tolerated. Smaller confidence intervals require more repetitions to achieve a given level of confidence. Confidence intervals that are less than half the value of the standard deviation will require a large number of repetitions to achieve reasonable confidence levels.[20]

B.4.1 Computation of Minimum Repetitions

It is impossible to know in advance exactly how many model runs will be needed to determine a mean (or any other statistical value) to the analyst's satisfaction. However, after a few model runs, the analyst can make an estimate of how many more runs may be required to obtain a statistically valid result.

The required minimum number of model repetitions is computed using the following equation:

$$CI_{1-\alpha\%} = 2 * t_{(1-\alpha/2),N-1} \frac{s}{\sqrt{N}} \tag{13}$$

[20]With such a tight confidence interval, the analyst may be striving for a degree of precision not reflected under real-world conditions.

where:

$CI_{(1-alpha)\%}$ = (1-alpha)% confidence interval for the true mean, where alpha equals the probability of the true mean not lying within the confidence interval

$t_{(1-alpha/2),N-1}$ = Student's t-statistic for the probability of a two-sided error summing to alpha with N-1 degrees of freedom, where N equals the number of repetitions

s = standard deviation of the model results

Note that when solving this equation for N, it will be necessary to iterate until the estimated number of repetitions matches the number of repetitions assumed when looking up the t statistic. Table 8 shows the solutions to the above equation in terms of the minimum number of repetitions for various desired confidence intervals and desired degrees of confidence.

Table 8. Minimum number of repetitions needed to obtain the desired confidence interval.

Desired Range (CI/S)	Desired Confidence	Minimum Repetitions
0.5	99%	130
0.5	95%	83
0.5	90%	64
1.0	99%	36
1.0	95%	23
1.0	90%	18
1.5	99%	18
1.5	95%	12
1.5	90%	9
2.0	99%	12
2.0	95%	8
2.0	90%	6

Notes:

1. Desired Range = desired confidence interval (CI) divided by standard deviation (S)
2. For example, if the standard deviation in the delay is 1.5 s and the desired confidence interval is 3.0 s at a 95-percent confidence level, then it will take eight repetitions to estimate the mean delay to within ± 1.5 s.

Appendix C: Estimation of the Simulation Initialization Period

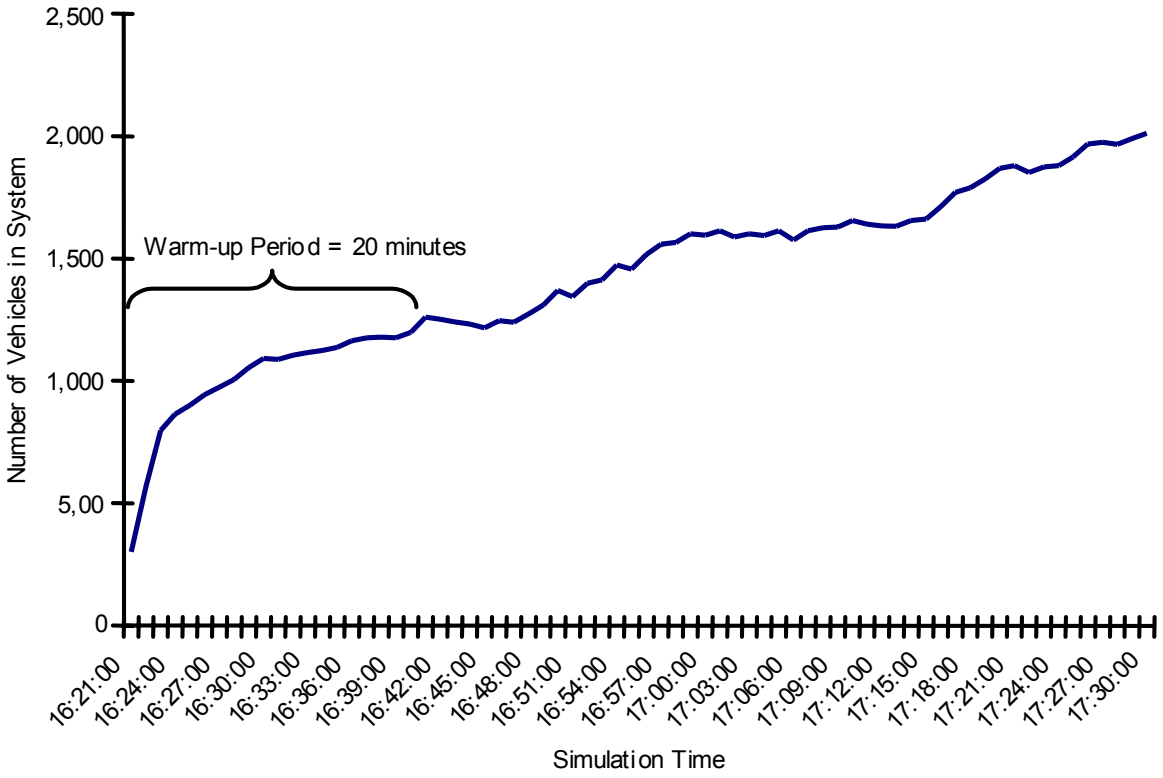
Simulation model runs usually start with zero vehicles on the network. If the simulation output is being compared to field measurements (as in calibration), then the artificial period where the simulation model starts out with zero vehicles (the warmup period) must be excluded from the reported statistics for system performance. Some software programs will do this automatically. For others, the warmup period must be computed offline by the analyst. This section explains how to identify the warmup period.

The number of vehicles present at any one time on the network is used to determine whether the model has reached equilibrium and, therefore, can start tallying performance statistics for the network. Once the number of vehicles present on the network ceases to increase by a minimum specified amount, then the warmup period is deemed to have been concluded (see figure 15).

If the number of vehicles and the mean speed do not level off within the first 15 min, it could be that the demand coded by the analyst for the system is greater than the system capacity. In this case, congestion will never level off. This will result in less accurate congestion statistics since the system never clears the congestion. The analyst should consider extending the starting and end times of the simulation to incorporate lower demand periods before and after the peak period.

If it is not feasible to extend the simulation period to uncongested time periods, the analyst should choose a warmup period that is equal to at least twice the estimated travel time at free-flow conditions to traverse the length of the network. For example, if the freeway being modeled is 8.1 km (5 mi) long, it takes roughly 5 min to traverse its length at the free-flow speed, so the warmup period is set at 10 min.

Note that in this example, the number of vehicles in the system is generally steadily increasing. The system never reaches equilibrium. Initialization is achieved when the number of vehicles entering the system is approximately equal to the number leaving the system (even though this slight decline is later superceded by greater increases). The generally increasing trend in the number of vehicles present in the system suggests that the simulation period should be extended earlier and later to incorporate lower demand periods at the beginning and the end of the peak period.



Source: I-680/I-580 Interchange Microsimulation Model, Dowling Associates, 2002

Figure 15. Illustration of warmup period.

Appendix D: Simple Search Algorithms for Calibration

Since simulation models are complex, it is not typically 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. The calibration problem is a nonlinear least-squares optimization problem.

There are many software programs available for identifying the optimal combination of calibration parameters for minimizing the squared error between the field observations and the simulation model. The Argonne National Laboratory (www-fp.mcs.anl.gov/otc/Guide/SoftwareGuide/Categories/nonlinleastsq.html) lists several software programs for nonlinear least-squares parameter estimation.

The sections below illustrate some simple approaches available for single-parameter estimation and dual-parameter estimation when working with a stand-alone simulation model. Estimation of three or more parameters, however, would require the use of a software program.

D.1 Single-Parameter Search Algorithm (Golden Section)[1]

Several methods are available for finding the value of a single parameter that minimizes the squared error between the model and the observations. These methods include Newton's Method, Secant Method, Quadratic Approximation Methods, and the Golden Section Method. The Golden Section Method is illustrated in the example below.

D.1.1 Example of Golden Section Method

Objective: Find the global value of the mean headway between vehicles that minimizes the squared error between the field counts of traffic volumes and the model estimates.

Approach: Use the Golden Section Method to identify the optimal mean headway.

[1] References: Hillier, F.S., and G.J. Lieberman, *Introduction to Operations Research*, Sixth Edition, McGraw-Hill, New York, 1995; Taha, H.A., *Operations Research, An Introduction*, Seventh Edition, Prentice-Hall, New York, 2003 (if the seventh edition is not available, look for the sixth edition published in 1996).

Step 1: Identify the maximum and minimum acceptable values for the parameter to be optimized.

This step brackets the possible range of the parameter in which the optimal solution is presumed to lie. The user can select the appropriate range. For this example, we will set the minimum acceptable value for the mean headway at 0.5 s and the maximum at 2.0 s. The larger the range, the more robust the search. However, it will take longer to find the optimum value. The smaller the range of acceptable values, the greater the likelihood that the best solution lies outside of the range.

Step 2: Compute the squared error for maximum and minimum values.

The simulation model is run to determine the volumes predicted when the maximum acceptable mean headway is input and again for the minimum acceptable headway. Either randomization should be turned off or the model should be run several times with each mean headway and the results averaged. The squared error produced by the model for each mean headway is computed.

Step 3: Identify two interior parameter values for testing.

The two interior points (x_1 and x_2) are selected according to specific ratios of the total range that preserve these ratios as the search range is narrowed in subsequent iterations. The formulas for selecting the two interior mean delays for testing are the following:

$$x_1 = \min + 0.382 \bullet (\max - \min) \tag{14}$$

$$x_2 = \min + 0.618 \bullet (\max - \min) \tag{15}$$

where:

x_1 = lower interior point value for the mean delay to be tested

x_2 = upper interior point value for the mean delay to be tested

min = lower end of the search range for the mean delay

max = upper end of the search range for the mean delay

The minimum and maximum ends of the search range are initially set by the user (based on the acceptable range set in step 1); however, the search range is then gradually narrowed as each iteration is completed.

Step 4: Compute squared error for two interior parameter values.

The simulation model is run for the new values of mean delay (x_1 and x_2) and the squared errors are computed.

Step 5: Identify the three parameter values that appear to bracket the optimum.

This step narrows the search range. The parameter value (x_1 or x_2) that produces the lowest squared error is identified. (If either the minimum or the maximum parameter values produce the least-squared error, the search range should be reconsidered.) The parameter values to the left (lower) and right (higher) of that point become the new minimum and maximum values for the search range. For instance, in figure 16, parameter value x_2 produces the lowest squared error.

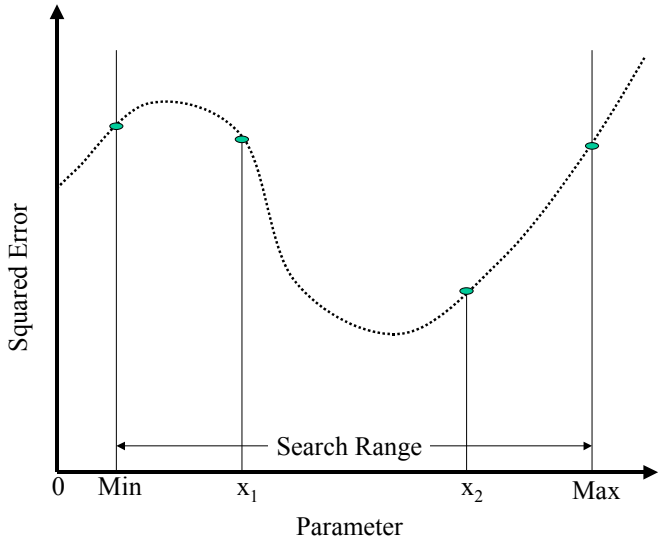


Figure 16. Golden Section Method.

Step 6: Return to step 3 and repeat until the uncertainty in the location of the optimal parameter value is satisfactory.

The Golden Section search is repeated until the range of values in which the optimum parameter value lies is small enough to satisfy the user's requirements. After about 10 iterations, the uncertainty in the optimal value of the parameter is reduced by a factor of 100. Therefore, if an initial range of 0.5 to 2.5 s is specified for the mean headway (a range of 2.0 s), this range will be reduced to 0.2 s after 10 iterations of the Golden Section Method. The user should obtain the optimal value of the mean headway to within ± 0.1 s.

D.2 Simple Two-Parameter Search Algorithm

In the case where two model parameters are to be optimized, a contour plot approach to identifying the optimal values of the parameters can be used. One first identifies the acceptable ranges of the two parameters and then exercises the model for pairs of values of the parameters. The squared error is computed for each pair of parameter values and is plotted in a contour plot to identify the value pairs that result in the lowest squared error. An example of this approach is shown in figure 17 for optimizing the mean headway and mean reaction time parameters. The search starts with the default values, blankets the region with a series of tests of different values, and then focuses in more detail on the solution area.

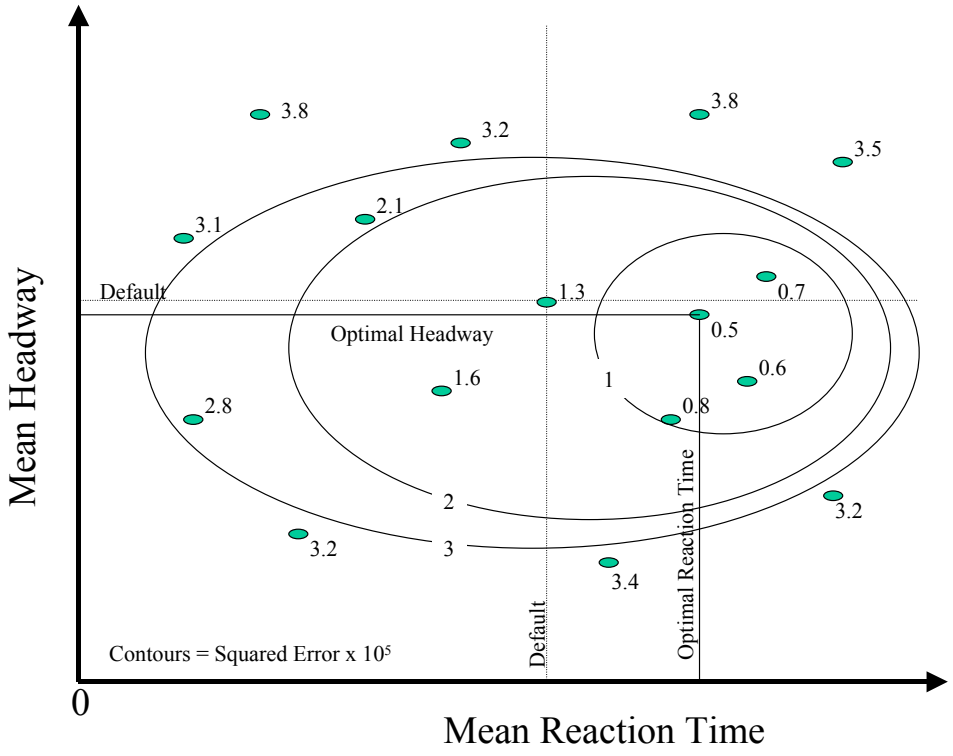


Figure 17. Example contour plot of the squared error.

D.3 Dual-Objective, Two-Parameter Search Algorithm[2]

This section addresses the case where the analyst wishes to consider two model calibration objectives separately, rather than as the weighted sum of the squared errors

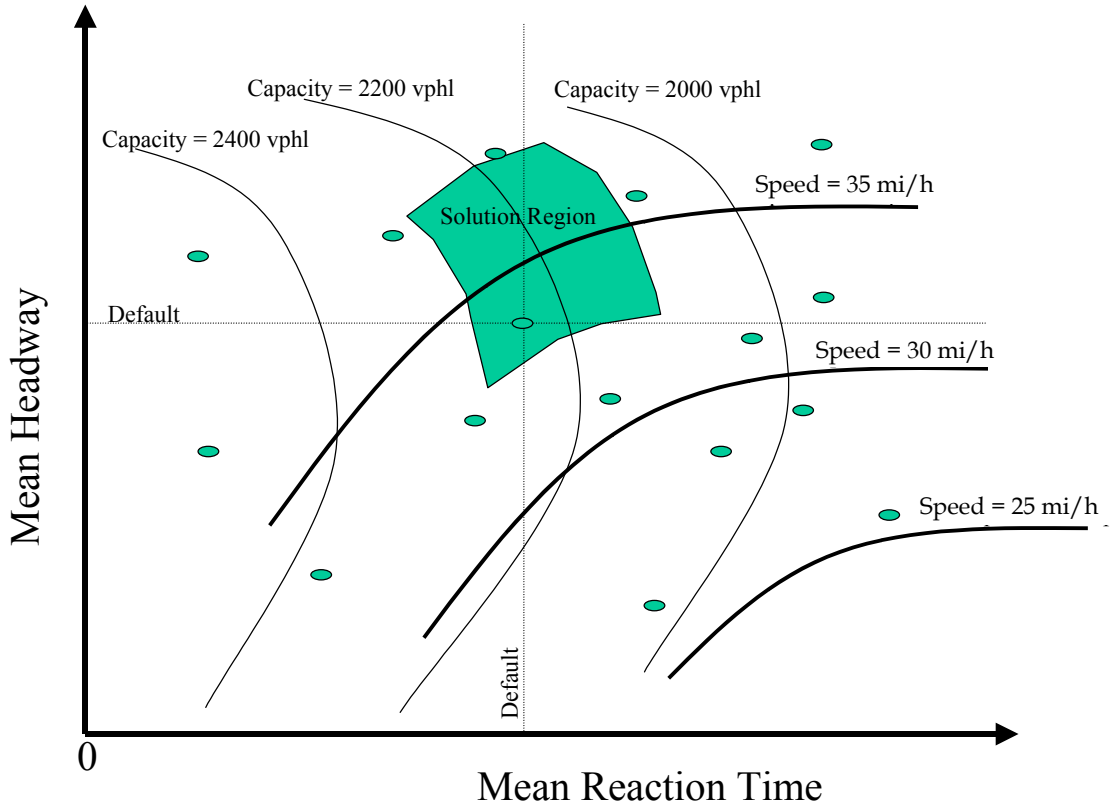
[2] Adapted from Gardes, Y., A.D. May, J. Dahlgren, and A. Skabardonis, *Bay Area Simulation and Ramp Metering Study*, California PATH Research Report UCB-ITS-PRR-2002-6, University of California, Berkeley, February 2002.

(single objective) discussed above. The analyst in this case can use a variation of the two-dimensional search algorithm identified above with up to two parameters. Instead of plotting a single set of contour lines for the squared error, the analyst plots two sets of contour lines (one set for each objective) for each pair of parameters. The analyst then visually identifies the acceptable solution range where both objectives are satisfied to the extent feasible. The following example illustrates this algorithm.

The analyst has identified the following two objectives for the model calibration:

- Model should predict a mean speed for the freeway of 56 km/h (35 mi/h) (± 3 km/h (2 mi/h)).
- Model should predict a maximum observed flow rate for the bottleneck section of 2200 veh/h per lane (± 100 veh/h).

Figure 18 shows how changing the mean headway and the mean reaction time changes the values of these two objectives (speed and capacity). The solution region shows the pairs of values of mean headway and mean reaction time that meet both objectives.



vphl = vehicles per hour per lane 1 mi/h = 1.61 km/h

Figure 18. Dual-objective, dual-parameter search.

Appendix E: Hypothesis Testing of Alternatives

When the microsimulation model is run several times for each alternative, the analyst may find that the variance in the results for each alternative is close to the difference in the mean results for each alternative. How is the analyst to determine if the alternatives are significantly different? To what degree of confidence can the analyst claim that the observed differences in the simulation results are caused by the differences in the alternatives and not just the result of using different random number seeds? This is the purpose of statistical hypothesis testing. Hypothesis testing determines if the analyst has performed an adequate number of repetitions for each alternative to tell the alternatives apart at the analyst's desired level of confidence.

E.1 Estimation of the Required Number of Model Repetitions

This section identifies how to estimate the minimum number of model run repetitions that would be required to determine if two alternatives with results a given distance apart are significantly different. This estimate, however, requires a preliminary estimate of the standard deviation of the model run results for the alternatives, which, in turn, require some preliminary runs to estimate the standard deviation.

The procedure involves the following steps:

1. Perform a preliminary set of model run repetitions for each alternative.

2. Estimate the standard deviation and the mean difference between the alternatives from the preliminary runs and then compute the required number of runs using the equations in this subsection.

3. If the required number of runs is greater than the preliminary number of runs, the analyst should perform the additional repetitions for each alternative and recompute the mean difference and standard deviation using the augmented set of model run repetitions.

E.1.1 Estimation of Pooled Standard Deviation

The analyst should perform about six to eight model repetitions of each alternative to estimate the pooled standard deviation for all alternatives according to the following equation:

$$s_p^2 = \frac{s_x^2 + s_y^2}{2} \quad (16)$$

where:

s_x = standard deviation of model run results for alternative x

s_y = standard deviation of model run results for alternative y

The preliminary model repetitions used to estimate the pooled estimate of the standard deviation of the model run results can also be used to estimate the likely difference of the means for each alternative.

E.1.2 Selection of Desired Confidence Level

A 95-percent confidence level is often selected for hypothesis testing. This means that there is a 5-percent chance (often called "alpha" in the textbooks) that the analyst will mistakenly reject the null hypothesis when it is really true (type I error). If a higher confidence level is desirable, it comes at the cost of increasing the likelihood of making a type II error (accepting the null hypothesis when it is really false) (table 9).

Table 9. Null hypothesis.

	True	False
Accept Hypothesis	OK	Type II Error
Reject Hypothesis	Type I Error	OK

The study objective may determine the desired confidence level. For example, if the objective is to design an alternative with a 95-percent probability that it will provide significant improvements over the current facility, then this is the appropriate confidence level for the determination of the number of model repetitions required.

E.1.3 Selection of Minimal Difference in the Means

The likely minimal difference in the means between the alternatives should be identified by the analyst. This is the target sensitivity of the simulation tests of the alternatives. Alternatives with mean results farther apart than this minimal difference will obviously be different. Alternatives with mean results closer together than this minimal difference will be considered to be indistinguishable.

The study objectives have some bearing on the selection of the minimal difference to be detected by the simulation tests. If the study objective is to design a highway improvement that reduces mean delay by at least 10 percent, then the tests should be designed to detect if the alternatives are at least 10 percent apart.

The preliminary model repetitions used to estimate the pooled estimate of the standard deviation of the model run results can also be used to estimate the likely difference of the means for each alternative. The smallest observed difference in these preliminary runs would be the selected minimal difference of the means to be used in determining the required number of repetitions.

E.1.4 Computation of Minimum Repetitions

Assuming that the analyst wishes to reject the null hypothesis that the means of the two most similar alternatives are equal with only an alpha percent chance of error against the counter hypothesis that the mean of alternative x is different than y, then the number of repetitions required can be computed according to the following equation:[3]

$$\left|\bar{x} - \bar{y}\right| > t_{(1-\alpha/2);2n-2} \bullet s_p \sqrt{\frac{2}{n}} \tag{17}$$

where:

$\left|\bar{x} - \bar{y}\right| =$ absolute value of the estimated difference between the mean values for the two most similar alternatives x and y

s_p = pooled estimate of the standard deviation of the model run results for each alternative

n = number of model repetitions required for each alternative

t = t statistic for a confidence level of 1-alpha and 2n-2 degrees of freedom[4]

Note that when solving this equation for N, it is necessary to iterate until the estimated number of repetitions matches the number of repetitions assumed when looking up the t statistic. Table 10 provides solutions for this equation for various target mean difference ranges and levels of confidence.

---

[3] If the analyst intends to perform hypothesis tests on only a few pairs of alternatives, then the equation provided should be sufficiently accurate. However, if the analyst plans to perform hypothesis testing of all possible pairs of alternatives, then this equation will underestimate the required number of repetitions needed to achieve the desired confidence level. The analyst should consult Lane, D.M., *Hyperstat OnLine, An Introductory Statistics Book and Online Tutorial for Help in Statistics*, Chapter 12: Introduction to Between-Subjects ANOVA ("All pairwise comparisons among means…"), Rice University (www.davidmlane.com/hyperstat).

[4] Note that this is a two-sided t-test for the null hypothesis that the means are equal versus the hypothesis that they are different.

Table 10. Minimum repetitions for distinguishing alternatives.

Minimum Difference of Means	Desired Confidence	Minimum Repetitions per Alternative
0.5	99%	65
0.5	95%	42
0.5	90%	32
1.0	99%	18
1.0	95%	12
1.0	90%	9
1.5	99%	9
1.5	95%	6
1.5	90%	5
2.0	99%	6
2.0	95%	4
2.0	90%	4

Notes:

1. The minimum difference in the means is expressed in units of the pooled standard deviation: $\frac{|\bar{x} - \bar{y}|}{S_p}$.
2. For example, if the pooled standard deviation in the delay for two alternatives is 1.5 s, and the desired minimum detectable difference in the means is a 3.0-s delay at a 95-percent confidence level, then it will take four repetitions of each alternative to reject the hypothesis that the observed differences in the simulation results for the two alternatives could be the result of random chance.

E.2 Hypothesis Testing for Two Alternatives

To determine whether simulation output provides sufficient evidence that one alternative is better than the other (e.g., build project versus no-build), it is necessary to perform a statistical hypothesis test of the difference of the mean results for each alternative. A null hypothesis is specified: "The model-predicted difference in VHT for the two alternatives occurred by random chance. There really is no significant difference in the mean travel time between the alternatives." A statistic is computed for a selected level of confidence, and if the difference between the two means is less than that statistic, then the null hypothesis is accepted and it is concluded that there is insufficient evidence to prove that the one alternative is better than the other. The analyst either makes more model repetitions for each alternative (to improve the sensitivity of the test) or relaxes his or her standards (confidence level) for rejecting the null hypothesis.

The specification of the problem is:

Null Hypothesis:

$$H_0 : \mu_x - \mu_y = 0 \tag{18}$$

against

$$H_1 : \mu_x - \mu_y \neq 0 \tag{19}$$

where:

μ_x = mean VHT (or some other measure) for alternative x

μ_y = mean for alternative y

This is a two-sided t-test with the following optimal rejection region for a given alpha (acceptable type I error):

$$\bar{x} - \bar{y} > t_{(1-\alpha/2);(n+m-2)} \bullet s_p \sqrt{\tfrac{1}{n} + \tfrac{1}{m}} \tag{20}$$

where:

$\bar{x} - \bar{y} =$ absolute value of the difference in the mean results for alternative x and alternative y

s_p = pooled standard deviation

t = Student's t-distribution for a level of confidence of 1-alpha and n+m-2 degrees of freedom

n = sample size for alternative x

m = sample size for alternative y

$$s_p^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{(m+n-2)} \tag{21}$$

where:

s_p = pooled standard deviation

s_x = standard deviation of the results for alternative x

s_y = standard deviation of the results for alternative y

n = sample size for alternative x

m = sample size for alternative y

The probability of mistakenly accepting the null hypothesis is alpha (usually set to 5 percent to get a 95-percent confidence level test). This is a type I error.

There is also a chance of mistakenly rejecting the null hypothesis. This is called a type II error and it varies with the difference between the sample means, their standard deviation, and the sample size.[5]

E.3 Hypothesis Testing for Multiple Alternatives

When performing hypothesis testing on more than one pair of alternatives, it is most efficient to first determine if *any* of the alternatives are significantly different from the others. An analysis of variance (ANOVA) test is performed to determine if the mean results for any of the alternatives are significantly different from the others:

- If the answer is **yes**, the analyst goes on to test specific pairs of alternatives.
- If the answer is **no**, the analysis is complete, or the analyst runs more model repetitions for each alternative to improve the ability of the ANOVA test to discriminate among the alternatives.

E.3.1 Analysis of Variance (ANOVA) Test

The ANOVA test has three basic requirements:

- Independence of samples (random samples).
- Normal sampling distribution of means.
- Equal variances of groups.

Levine's test of heteroscedasticity can be used for testing whether or not the variances of the model run results for each alternative are similar. Less powerful nonparametric tests, such as the Kruskal-Wallis test (K-W statistic) can be performed if the requirements of the ANOVA test cannot be met.

However, the ANOVA test is tolerant of modest violations of these requirements and may still be a useful test under these conditions. The ANOVA test will tend to be conservative if its requirements are not completely met (less likely to have a type I error with a lower power of the test to correctly reject the null hypothesis).

[5] Analysts should consult standard statistical textbooks for tables on the type II errors associated with different confidence intervals and sample sizes.

To perform the ANOVA test, first compute the test statistic:

$$F = \frac{MSB}{MSW} \tag{22}$$

where:

F = test statistic

MSB = mean square error between the alternatives (formula provided below)

MSW = mean square error among the model results for the same alternative (within alternatives)

The formulas below show how to compute MSB and MSW:

$$MSB = \frac{\sum_{i=1}^{g} n_i \bullet (\bar{x}_i - \bar{x})^2}{g - 1} \tag{23}$$

where:

MSB = mean square error between the alternatives (i = 1 to g)

n_i = number of model runs with different random number seeds for alternative i

$\bar{x}_i =$ mean value for alternative i

$\bar{x} =$ mean value averaged across all alternatives and runs

g = number of alternatives

and

$$MSW = \frac{\sum_{i=1}^{g} (n_i - 1) \bullet s_i^2}{N - g} \tag{24}$$

where:

MSB = mean square error between the alternatives (i = 1 to g)

n_i = number of model runs with different random number seeds for alternative i

s_i^2 = variance of the model run results for alternative i

N = total number of model runs summed over all alternatives

g = number of alternatives

The null hypothesis of equal means is rejected if:

$$F > F_{1\text{-}alpha,g\text{-}1,N\text{-}g} \tag{25}$$

where:

$F_{1\text{-}alpha,g\text{-}1,N\text{-}g}$ = F statistic for a type I error of alpha (alpha is usually set at 5 percent for a 95-percent confidence level) and g-1 and N-g degrees of freedom. N is the total number of model runs summed over all alternatives; g is the number of alternatives

If the null hypothesis cannot be rejected, then the analysis is either complete (there is no statistically significant difference between any of the alternatives at the 95-percent confidence level) or the analyst should consider reducing the level of confidence to below 95 percent or implementing more model runs per alternative to improve the sensitivity of the ANOVA test.

E.3.2 Pairwise Tests of Some Pairs of Alternatives

If performing hypothesis tests for only a few of the potential pairs of alternatives, the standard t-test can be used for comparing a single pair of alternatives:

$$\bar{x} - \bar{y} > t_{(1-\alpha/2);(n+m-2)} \bullet s_p \sqrt{\tfrac{1}{n} + \tfrac{1}{m}} \tag{26}$$

$\bar{x} - \bar{y} =$ absolute value of the difference in the mean results for alternative x and alternative y

s_p = pooled standard deviation

t = t distribution for a level of confidence of 1-alpha and n+m-2 degrees of freedom

n = sample size for alternative x

m = sample size for alternative y

$$s_p^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{(m+n-2)} \tag{27}$$

where:

s_p = pooled standard deviation

s_x = standard deviation of the results for alternative x

s_y = standard deviation of the results for alternative y

n = sample size for alternative x

m = sample size for alternative y

If one merely wishes to test that the best alternative is truly superior to the next best alternative, then the test needs to be performed only once.

If one wishes to test other possible pairs of alternatives (such as second best versus third best), it is possible to still use the same t-test; however, the analyst should be cautioned that the level of confidence diminishes each time the test is actually performed (even if the analyst retains the same nominal 95-percent confidence level in the computation, the mere fact of repeating the computation reduces its confidence level). For example, a 95-percent confidence level test repeated twice would have a net confidence level for both tests of $(0.95)^2$, or 90 percent.

Some experts, however, have argued that the standard t-test is still appropriate for multiple paired tests, even at its reduced confidence level.

E.3.3 Pairwise Tests of All Pairs of Alternatives

To preserve a high confidence level for all possible paired tests of alternatives, the more conservative John Tukey "Honestly Significantly Different" (HSD) test should be used to determine if the null hypothesis (that the two means are equal) can be rejected.[6]

The critical statistic is:

$$t_s = \frac{\left|\bar{x}_i - \bar{x}_k\right|}{\sqrt{\frac{MSE}{n}}} \tag{28}$$

where:

t_s = Studentized t-statistic

$\bar{x}_i$ = mean value for alternative i

MSE = mean square error = MSB + MSW

N = number of model runs with different random number seeds for each alternative (if the number of runs for each alternative is different, then use the harmonic mean of the number of runs for each alternative)

Reject the null hypothesis that the mean result for alternative i is equal to that for alternative k if:

$$t_s > t_{1\text{-}alpha,g\text{-}1} \tag{29}$$

---

[6] Adapted from Lane, D.M., *Hyperstat OnLine, An Introductory Statistics Book and Online Tutorial for Help in Statistics*, Rice University (www.davidmlane.com/hyperstat).

where:

$t_{1\text{-alpha},g\text{-}1}$ = t statistic for a desired type I error of alpha (alpha is usually set at 5 percent to obtain a 95-percent confidence level) and g-1 degrees of freedom, with g equal to the total number of alternatives tested, not just the two being compared in each test

Some experts consider the HSD test to be too conservative, failing to reject the null hypothesis of equal means when it should be rejected. The price of retaining a high confidence level (the same as retaining a low probability of a type I error) is a significantly increased probability of making a type II error (accepting the null hypothesis when it is really false).

E.4 What To Do If the Null Hypothesis Cannot Be Rejected

If the null hypothesis of no significant difference in the mean results for the alternatives cannot be rejected, then the analyst has the following options:

- Increase the number of model run repetitions per alternative until the simulation performance of one alternative can be distinguished from the other.
- Reduce the confidence level from 95 percent to a lower level where the two alternatives are significantly different and report the lower confidence level in the results.
- Accept the result that the two alternatives are not significantly different.

Appendix F: Demand Constraints

Microsimulation results are highly sensitive to the amount by which the demand exceeds the capacity of the facility, so it is vital that realistic demand forecasts be used in the analysis. The following steps outline a procedure for manually reducing the forecasted demands in the study area to better match the capacity of the facilities feeding the study area.

Step 1: Identify Gateway Bottlenecks

The analyst should first identify the critical bottlenecks on the facilities feeding the traffic to the boundaries of the microsimulation study area. Bottlenecks are sections of the facilities feeding the model study area that either have capacities less than other sections of the freeway or demands greater than the other sections. These are the locations that will probably be the first ones to experience congested conditions as traffic grows.

These bottlenecks may be located on the boundary of the microsimulation study area, in which case they are identical to the gateway zones on the boundary of the microsimulation model study area. Bottlenecks within the microsimulation model study area can be disregarded since they will be taken into account by the microsimulation model.

Inbound bottlenecks are congested sections feeding traffic to the microsimulation model area. Outbound bottlenecks are congested sections affecting traffic leaving the microsimulation model area. If an outbound bottleneck will probably create future queues that will back up into the microsimulation model study area, then the model study area should be extended outward to include the outbound bottleneck. If the future outbound queues will not back up into the model study area, then these bottlenecks can be safely disregarded.

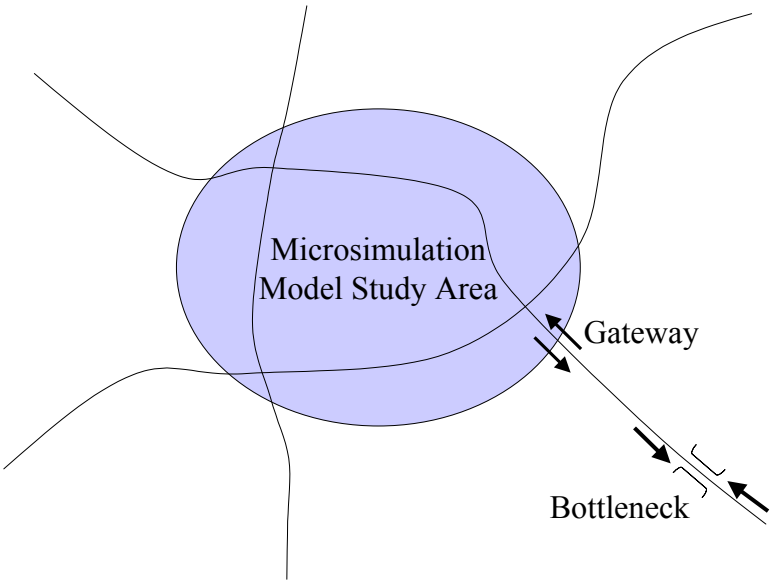


Figure 19. Bottleneck, gateway, and study area.

Step 2: Estimate Excess Demand at Inbound Bottlenecks

If the forecasted hourly demand at a bottleneck (in the inbound direction toward the model) exceeds its capacity, the proportion of the demand that is in excess of the available hourly capacity should be computed:

$$X = \frac{D - C}{C} \tag{30}$$

where:

X = proportion of excess demand

D = forecasted demand (veh/h)

C = estimated capacity (veh/h)

Step 3: Reduce Forecasted Demand Inbound at Gateways

The forecasted hourly demands for the off-ramps between the bottleneck and the gateway entering the microsimulation study area should be reduced in proportion to the amount by which the forecasted bottleneck demand exceeds its capacity:

$$D_{const} = D_{unconst} * (1 - X) \tag{31}$$

where:

D_{const} = constrained demand (veh/h) for a downstream off-ramp or exit point

$D_{unconst}$ = unconstrained demand forecast (veh/h)

X = proportion of excess demand

It is suggested that the off-ramp demand be reduced in proportion to the reduction in demand that can get through the bottleneck, assuming that the amount of reduction in the downstream flows is proportional to the reduction in demand at the bottleneck. If the analyst has superior information (such as an O-D table), then the assumption of proportionality can be overridden by the superior information. The constrained downstream gateway demand is then obtained by summing the constrained bottleneck, off-ramp, and on-ramp volumes between the bottleneck and the gateway to the study area.

Figure 20 illustrates how the proportional reduction procedure would be applied for a single inbound bottleneck that reduces the peak-hour demand that can get through from 5000 veh/h to 4000 veh/h. Since there is an interchange between the bottleneck and the entry gate to the microsimulation study area, the actual reduction is somewhat less (800 veh/h) at the gate.

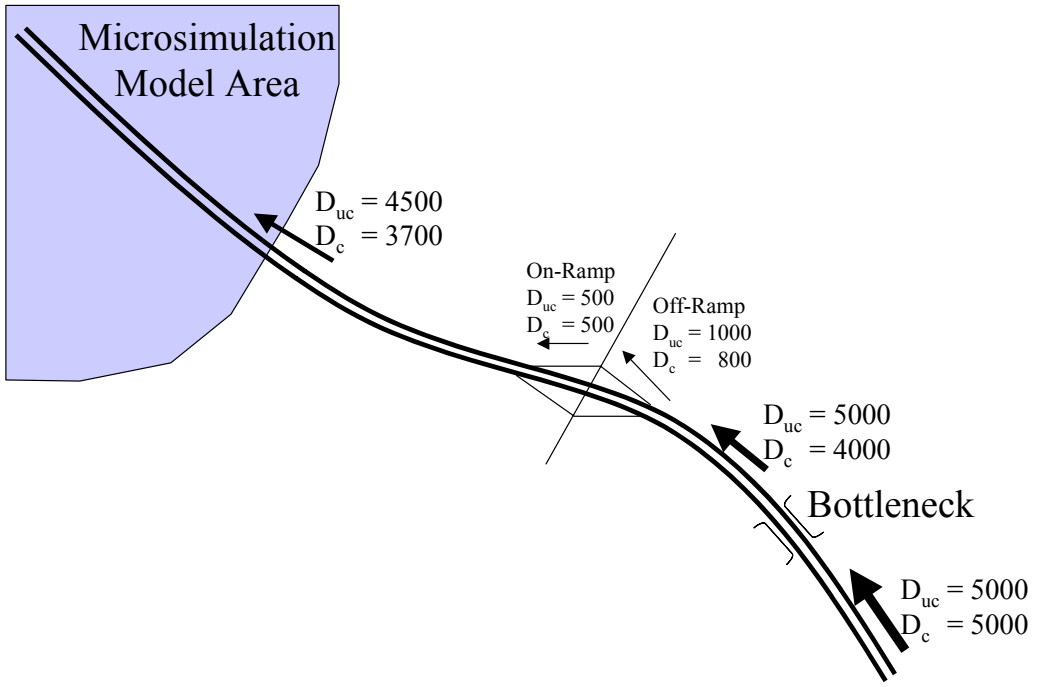


Figure 20. Example proportional reduction of demand for capacity constraint.

Starting upstream of the bottleneck, there is an unconstrained demand for 5000 veh/h. Since the bottleneck has a capacity of 4000 veh/h, the downstream capacity constrained demand is reduced from the unconstrained level of 5000 veh/h to 4000 veh/h. Thus, 1000 vehicles are stored at the bottleneck during the peak hour. Since it is assumed that the stored vehicles are intended for downstream destinations in proportion to the exiting volumes at each off-ramp and freeway mainline, the downstream volumes are reduced the same percentage as the percentage reduction at the bottleneck (20 percent). A 20-percent reduction of the off-ramp volume results in a constrained demand of 800 veh/h. The on-ramp volume is unaffected by the upstream bottleneck, so its unconstrained

demand is unchanged at 500 veh/h. The demand that enters the microsimulation study area is equal to the constrained demand of 4000 veh/h leaving the bottleneck, minus the 800 veh/h leaving the freeway on the off-ramp, plus 500 veh/h entering the freeway at the on-ramp, which results in a constrained demand of 3700 veh/h.

Appendix G: Bibliography

G.1 Traffic Microsimulation Fundamentals

Revised Monograph on Traffic Flow Theory, N.H. Gartner, C.J. Messer, and A. Ruthi (editors), FHWA, 1994 (www.tfhrc.gov/its/tft).

May, A.D., *Traffic Flow Fundamentals*, Prentice Hall, Englewood Cliffs, NJ, 1990.

Ahmed, K.I., *Modeling Drivers' Acceleration and Lane Changing Behavior*, Ph.D. Dissertation, Massachusetts Institute of Technology, Cambridge, MA, February 1999.

Elefteriadou, L., G. List, J. Leonard, H. Lieu, M. Thomas, R. Giguere, R. Brewish, and G. Johnson, "Beyond the Highway Capacity Manual: A Framework for Selecting Simulation Models in Traffic Operational Analyses," *Transportation Research Record 1678*, National Academy Press, 1999, pp. 96-106.

G.2 Calibration and Validation of Microsimulation Models

Abdullai, B., J. 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.

Ben-Akiva, M., A. Davol, T. Toledo, H.N. Koutsopoulos, W. Burghout, I. Andréasson, T. Johansson, and C. Lundin, "Calibration and Evaluation of MITSIMLab in Stockholm," Paper presented at the Annual Meeting, TRB, Washington, DC, 2002.

Fellendorf, M., and P. Vortisch, "Validation of the Microscopic Traffic Flow Model VISSIM in Different Real-World Situations," Paper presented at the Annual Meeting, TRB, Washington, DC, 2001 (www.itc-world.com/library).

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

Gardes, Y., A.D. May, J. Dahlgren, and A. Skabardonis, *Bay Area Simulation and Ramp Metering*, California PATH Program Research Report.

Hellinga, B.R., *Requirements for the Calibration of Traffic Simulation Models*, Canadian Society of Civil Engineering (CSCE), University of Waterloo, Ontario, Canada, 1998.

Jayakrishnan, R., J. Oh, and A.B. Sahraoui, *Calibration and Path Dynamics Issues in Microscopic Simulation for Advanced Traffic Management and Information Systems*, UCI-ITS-WP-00-22, Institute of Transportation Studies, University of California, Irvine, December 2000.

Lee, D-H., X. Yang, and P. Chandrasekar, "Parameter Calibration for Paramics Using Genetic Algorithm," Paper No. 01-2339, Annual Meeting, TRB, Washington, DC, 2001.

Ma, T., and B. Abdulhai, "A Genetic Algorithm-Based Optimization Approach and Generic Tool for the Calibration of Traffic Microscopic Simulation Parameters," Paper presented at the Annual Meeting, TRB, Washington, DC, 2002.

Owen, L.E., G. Brock, S. Mastbrook, R. Pavlik, L. Rao, C. Stallard, S. Sunkari, and Y. Zhang, *A Compendium of Traffic Model Validation Documentation and Recommendations*, FHWA Contract DTFH61-95-C-00125, FHWA, 1996.

Project Suite Calibration and Validation Procedures, Quadstone, Ltd., Edinburgh, United Kingdom, February 2000 (www.paramics-online.com/tech_support/ customer/cust_doc.htm (password required)).

Rakha, H., B. Hellinga, M. Van Aerde, and W. Perez, "Systematic Verification, Validation and Calibration of Traffic Simulation Models," Paper presented at the Annual Meeting, TRB, Washington, DC, 1996.

Rao, L., L. Owen, and D. Goldsman, "Development and Application of a Validation Framework for Traffic Simulation Models," *Proceedings of the 1998 Winter Simulation Conference*, D.J. Medeiros, E.F. Watson, J.S. Carson, and M.S. Manivannan (editors).

Skabardonis, A., *Simulation of Freeway Weaving Areas*, Preprint No. 02-3732, Annual Meeting, TRB, Washington, DC, 2002.

Traffic Appraisal in Urban Areas: Highways Agency, Manual for Roads and Bridges, Volume 12, Department for Transport, London, England, United Kingdom, May 1996 (www.archive.official-documents.co.uk/document/ha/dmrb/index.htm).

Wall, Z., R. Sanchez, and D.J. Dailey, *A General Automata Calibrated With Roadway Data for Traffic Prediction*, Paper No. 01-3034, Annual Meeting, TRB, Washington, DC, 2001.

G.3 Analysis of Results

Benekohal, R.F., Y.M. Elzohairy, and J.E. Saak, *A Comparison of Delay From HCS, Synchro, Passer II, Passer IV, and CORSIM for an Urban Arterial*, Preprint, Annual Meeting, TRB, Washington, DC, 2002.

Joshi, S.S., and A.K. Rathi, "Statistical Analysis and Validation of Multi-Population Traffic Simulation Experiments," *Transportation Research Record 1510*, TRB, Washington, DC, 1995.

Prassas, E.S., *A Comparison of Delay on Arterials by Simulation and HCM Computations*, Preprint No. 02-2652, Annual Meeting, TRB, Washington, DC, 2002.

Tian, Z.Z., T. Urbanik II, R. Engelbrecht, and K. Balke, *Variations in Capacity and Delay Estimates From Microscopic Traffic Simulation Models*, Preprint, Annual Meeting, TRB, Washington, DC, 2002.

Trueblood, M., "CORSIM…SimTraffic: What's the Difference?," *PC-TRANS*, Winter/Spring 2001.

Gibra, I.N., *Probability and Statistical Inference for Scientists and Engineers*, Prentice-Hall, Englewood Cliffs, NJ, 1973.

Lane, D.M., *Hyperstat OnLine, An Introductory Statistics Book and Online Tutorial for Help in Statistics*, Rice University, Houston, TX (www.davidmlane.com/hyperstat).

Gerstman, B., and M. Innovera, *StatPrimer, Version 5.1*, California State University, San Jose (www.sjsu.edu/faculty/gerstman/StatPrimer).

HRDO-03/07-04(WEB)E