The Secrets to HCM Consistency Using Simulation Models

Ronald T. Milam, AICP
David Stanek, PE
Chris Breiland

Fehr & Peers
2990 Lava Ridge Court, Suite 200
Roseville, CA 95661
r.milam@fehrandpeers.com
(916) 773-2015 fax

ABSTRACT

Traffic operations analysis has evolved to a point where many applications require the use of micro-simulation software programs. Major contributing factors to this need include an increase in traffic congestion and the development of multi-modal transportation systems. When traffic demand exceeds capacity, applying the analytical methods in the Highway Capacity Manual (HCM), Transportation Research Board, 2000 are not always appropriate and may underestimate delay and congestion. Further, the HCM does not contain a robust procedure or methodology for analyzing multiple travel modes in a system where the interactions between modes are fully considered and the performance measures can be extracted by mode. As a result, practitioners faced with problems involving these conditions often turn to the use of micro-simulation traffic software programs to create models that can provide this information.

Using micro-simulation programs to create models to represent the conditions noted above is becoming more common as the software becomes more user-friendly and the models require less time to set up and operate. However, most of these programs come with little or no direction on how to generate performance measures that are consistent with the HCM. This is an important limitation of these programs because many public agencies responsible for conducting or reviewing traffic operations analysis still require the use of HCM-consistent performance measures. Consistency with the HCM is desirable because of the wide-spread acceptance of using LOS to describe operating conditions. Unfortunately, current micro-simulation programs do not offer much guidance on how to correctly extract the quantitative performance measures such as delay or density that are used to determine LOS.

In response to this problem, this paper will identify three specific HCM consistency problems that occur when applying simulation models and propose solutions (share some of the secrets) to these problems based on over a decade of experience developing and applying micro-simulation models for a wide variety of projects in a range of different environments.
INTRODUCTION

Traffic operations analysis has evolved to a point where many applications require the use of micro-simulation software programs. Major contributing factors to this need include an increase in traffic congestion and the development of multi-modal transportation systems. When traffic demand exceeds capacity, applying the analytical methods in the *Highway Capacity Manual* (HCM), Transportation Research Board, 2000 are not always appropriate and may underestimate delay and congestion. Recognition of this condition is noted throughout the HCM as indicated by the following quotes.

“… the HCM methods are generally not appropriate…for the evaluation of queues that are building over both time and space.” – Page 9-1, HCM

“Certain freeway traffic conditions cannot easily be analyzed by the methodology. Multiple overlapping bottlenecks are an example. Therefore, other tools may be more appropriate…” - Page 22-1, HCM

The problem of how to analyze traffic operations under congested conditions or in a multi-modal environment is expanding as urban areas struggle with growing travel demand and a transportation system that has been constrained by lack of investment.

The HCM does provide some direction for handling the above circumstances such as advising the use of simulation models. However, this direction tells the user “when” to use simulation and not “how”. While traffic operations micro-simulation programs are becoming more user-friendly, most of the programs come with little or no direction on how to generate performance measures that are consistent with the HCM. This is an important limitation of these programs because many public agencies responsible for conducting or reviewing traffic operations analysis still require the use of HCM-consistent performance measures.

Consistency with the HCM is desirable because of the wide-spread acceptance of using LOS to describe operating conditions. Unfortunately, current micro-simulation programs do not offer much guidance on how to correctly extract the quantitative performance measures such as delay or density that are used to determine LOS. Contacting software vendors did not provide complete answers to solving this problem but the well documented original research that supports the HCM did.

The original research contains the information about how data were collected to support the HCM calculations. The flexibility of micro-simulation programs allows the user to create models that allow for performance data to be collected in a manner that replicates the original data collection. This approach allowed for performance calculations to closely approximate those in the HCM and also revealed important observations that should be shared with the profession.
HCM PERFORMANCE MEASURES

To understand whether micro-simulation software programs generate HCM consistent performance measures, the HCM performance measure definitions had to be clearly established. Table 1 contains a summary of the HCM performance measures for multiple roadway facilities.

<table>
<thead>
<tr>
<th>FACILITY</th>
<th>PERFORMANCE MEASURE</th>
<th>DEFINITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic Freeway Segment</td>
<td>Vehicle Density</td>
<td>Passenger cars/mile/lane (must exclude influence areas of on- and off-ramps)</td>
</tr>
<tr>
<td>Weaving Sections</td>
<td>Vehicle Density</td>
<td>Passenger cars/mile/lane (must be less than 2,500 feet)</td>
</tr>
<tr>
<td>Ramp Junctions</td>
<td>Vehicle Density</td>
<td>Passenger cars/mile/lane (within 1,500-foot influence area of ramp)</td>
</tr>
<tr>
<td>Arterials</td>
<td>Average Speed</td>
<td>Miles per hour</td>
</tr>
<tr>
<td>Signalized Intersections</td>
<td>Control Delay</td>
<td>Seconds/vehicle (includes initial deceleration delay, queue move-up time, stopped delay, and final acceleration delay)</td>
</tr>
<tr>
<td>Stop-Sign Controlled Intersections</td>
<td>Control Delay</td>
<td>Seconds/vehicle (includes initial deceleration delay, queue move-up time, stopped delay, and final acceleration delay)</td>
</tr>
<tr>
<td>Roundabouts</td>
<td>Capacity and Volume-to-Capacity (V/C) Ratio</td>
<td>Vehicles/hour</td>
</tr>
</tbody>
</table>

Notes:
(1) All performance measures are calculated for the peak 15 minutes of the peak hour.

The definitions in Table 1 contain important clarifying information about the specific details of each performance measure and how performance measure data must be collected from a simulation model. For example, most simulation programs generate total delay for intersections and not control delay. They may also report density in vehicles/mile/lane instead of passenger car equivalents. Further, because simulation programs have multiple methods of collecting and
Summarizing data, being specific about defining the limits of a weaving section or a ramp junction is essential input information that must be accounted for correctly when building the model.

**SIMULATION MODEL SPECIFICATION**

When developing a traffic operations model using micro-simulation software programs such as CORSIM, Paramics, SimTraffic, or VISSIM, the specification of the model network and output data is essential in creating HCM consistent performance measures. Network development is particularly important for many simulation programs because the network set up includes the identification of points, nodes, or links where output data are to be measured or accumulated. During model setup, the user is also instructing the program about which data to output and, if possible, developing user-defined variables.

For the most part, micro-simulation programs provide sufficient data to accurately calculate the “average speed” and “density” performance measures noted in Table 1. However, the default output often measures density in terms of vehicles and not passenger car equivalents. An adjustment to account for trucks and other heavy vehicles is necessary to convert the output to passenger car equivalents. Control delay is available in some programs such as CORSIM, but other programs only provide total delay. While total delay is similar to control delay, it is approximately 10 percent higher. According to the Guidelines for Applying Traffic Microsimulation Modeling Software (FHWA, August 2003), this difference is small enough to allow the use of total delay in place of control delay. Nevertheless, use of total delay should be acknowledged by practitioners when reporting intersection analysis results.

Unfortunately for practitioners, micro-simulation software programs do not all use the same or even similar network coding provisions. As such, any guidance about network development must be software specific. For the purposes of this paper, three software program examples are provided involving three different types of network coding issues as listed below.

- Peak hour factor – Paramics
- Signalized intersection – SimTraffic
- Freeway ramp junction – VISSIM

For each of these examples, typical network coding issues are identified that directly relate to the ability to obtain accurate performance measure estimates. For each coding issue, a recommended practice is provided for HCM consistency.

**Peak Hour Factor**

HCM analysis is generally based on a 15-minute peak period (see graphic below). However, practitioners often rely on hourly inputs of volume and then use peak hour factors to adjust the hourly volumes. This common practice is not likely to provide accurate simulation results in many programs, especially when simulating large networks or congested conditions. In large or
congested networks the peak 15-minute volume may not correspond temporally with the peak 15-minute delay at a study intersection or roadway segment. This is because bottlenecks, traffic signals, and driveways cause a lag in the time when vehicles enter the study area and when they arrive at the study intersection or roadway segment. Therefore a realistic demand profile should be entered into the micro-simulation software to accurately measure HCM performance results.

To accurately model the dynamics of vehicle flow, demand data must be entered in a way that better reflects real-world conditions. Traffic count data collected manually or electronically are often aggregated into 5 or 15-minute increments. By entering traffic count data in these smaller time intervals, queuing, congestion, and vehicle speeds can be more accurately modeled over the entire study period (for example, the PM peak hour). In contrast, entering hourly count data with a peak hour factor captures only the variation of a single 15-minute interval.

While all of the major micro-simulation programs allow the user to enter demand data in discrete intervals, the example below shows how to enter demand data into the Paramics program. Paramics allows the user to input demand data in user defined intervals by editing the “profile” file. The graphics below shows six 15-minute demand intervals defined for a PM peak-period.
PARAMICS DEMAND PROFILE

The Paramics profile file enables the user to define a unique origin-destination matrix for each user-defined period. In the example above, 15-minute traffic count data can be directly entered for each of the six demand periods.

PARAMICS DEMAND MATRIX
With an accurate demand profile set up, the Paramics output can be analyzed in a spreadsheet to calculate the peak 15-minute delay or density values for each study facility independent of when the peak 15-minute demand occurred. Provided the study interval is long enough to allow congestion to move through the study area, HCM consistent performance measures can be extracted at all study locations.

**Signalized Intersection**

To code signalized intersections correctly in a program such as SimTraffic requires a thorough understanding of how the program accumulates delay for use in output summaries. Consider the network example below as being a typical arterial represented in SimTraffic.

**SIMTRAFFIC ARTERIAL CODING EXAMPLE**

In this example, queuing from intersection 2 extends through intersection 6. Intersection 6 is a local shopping center driveway with stop control on the minor street approaches. When SimTraffic estimates delay associated with the signal control at intersection 2, the delay estimates are limited to its approach links and does not extend beyond any upstream nodes such as intersection 6.

**SIMTRAFFIC PERFORMANCE REPORT FOR INTERSECTION 2**

<table>
<thead>
<tr>
<th>Movement</th>
<th>EBL</th>
<th>EBT</th>
<th>EBR</th>
<th>WBL</th>
<th>WBT</th>
<th>WBR</th>
<th>NBL</th>
<th>NBT</th>
<th>NBR</th>
<th>SBL</th>
<th>SBT</th>
<th>SBR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delay (s)</td>
<td>78.8</td>
<td>94.0</td>
<td>92.4</td>
<td>30.2</td>
<td>21.9</td>
<td>10.9</td>
<td>27.7</td>
<td>11.6</td>
<td>3.1</td>
<td>71.2</td>
<td>64.1</td>
<td>11.5</td>
</tr>
<tr>
<td>Vehicles Exited</td>
<td>22</td>
<td>9</td>
<td>101</td>
<td>43</td>
<td>8</td>
<td>33</td>
<td>93</td>
<td>43</td>
<td>75</td>
<td>18</td>
<td>70</td>
<td>10</td>
</tr>
<tr>
<td>Hourly Exit Rate</td>
<td>88</td>
<td>36</td>
<td>404</td>
<td>172</td>
<td>32</td>
<td>132</td>
<td>372</td>
<td>172</td>
<td>30</td>
<td>72</td>
<td>280</td>
<td>40</td>
</tr>
</tbody>
</table>
This is appropriate for a series of nodes that represent signalized intersections and may be theoretically correct for a combination of signalized and unsignalized intersections such as that shown in the example above. However, the queue from intersection 2 extends through intersection 6 and if this problem is to be accurately described, the SimTraffic model should recognize the source of the delay (i.e., intersection 2) and measure its full extent (i.e., through intersection 6) as shown on the following page.

The delay estimates and LOS are different if based on link 1 alone versus the combination of links 1 and 2 as shown in Table 2 below.

<table>
<thead>
<tr>
<th>Facility</th>
<th>Link 1 Only</th>
<th>Links 1 &amp; 2 Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Delay</td>
<td>LOS</td>
</tr>
<tr>
<td>Southbound Approach</td>
<td>59 seconds</td>
<td>E</td>
</tr>
<tr>
<td>Overall Intersection</td>
<td>40 seconds</td>
<td>D</td>
</tr>
</tbody>
</table>

Without including link 2, the severity of the problem may have not been recognized. This same type of problem occurs when ‘bend nodes’ are added in SimTraffic. Bend nodes are used to help add curvature to the roadway network or to better match the actual geometric alignment of roadways. However, the user must view the simulation to determine whether queues extend beyond these nodes.

**Freeway Ramp Junction**

The HCM defines the freeway ramp junction influence area for a merge location as the two shoulder lanes plus adjacent auxiliary lane(s) within 1,500 feet downstream of the ramp gore.
The figure below shows a typical freeway ramp merge as coded in VISSIM. The four-lane freeway segment has a one-lane on-ramp that merges with a 450-foot auxiliary lane.

VISSIM will report link statistics including density, which is used to set the level of service, by link and by lane using the Link Evaluation function. Additionally, the user can specify the segment length along a link in which to report these statistics.

To calculate the density only within the ramp influence area, the density for each of the three right lanes (the auxiliary lane and two adjacent through lanes) should be collected for link 190. For link 191, the density for the right two lanes should be collected by lane, but only for the first 1,050 ft of the link (that is, 450 + 1,050 = 1,500 ft). The density of the ramp influence area is estimated by first calculating the volume-weighted average density across links for each lane, and then by calculating the volume-weighted average density across lanes.
VISSIM LINK EVALUATION RESULTS

The density estimates and LOS are different if based on link 190 or 191 alone versus the ramp influence area as shown in Table 3 below.

<table>
<thead>
<tr>
<th>Location</th>
<th>Length</th>
<th>Density</th>
<th>LOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Link 190</td>
<td>450 ft</td>
<td>45.8 vehicles/lane-mile</td>
<td>F</td>
</tr>
<tr>
<td>Link 191</td>
<td>2,930 ft</td>
<td>37.3 vehicles/lane-mile</td>
<td>D</td>
</tr>
<tr>
<td>Ramp Influence Area</td>
<td>1,500 ft</td>
<td>42.1 vehicles/lane-mile</td>
<td>E</td>
</tr>
</tbody>
</table>

If the ramp junction LOS were based on the 450-foot link 190, then the severity of traffic operations would be overstated.

CONCLUSIONS

The examples above demonstrate the importance of correctly specifying simulation model networks and output variables to generate HCM consistent performance measures and analysis results. This information is not available through user manuals or software help instructions.

While the HCM does a good job describing when to use simulation and the FHWA Guidelines for Applying Traffic Microsimulation Modeling Software provide useful information on how to perform simulation modeling, insufficient information is available to practitioners when applying specific software programs if achieving HCM consistency is desired. The profession would benefit from additional federal efforts such as the Next Generation Simulation (NGSIM) program.
that relate directly to simulation model output and analysis. Absent this type of information, software vendors could provide more complete information about the ability of the software programs to generate HCM consistent performance measures.