Paper No. 02-2296

Variations in Capacity and Delay Estimates from Microscopic Traffic Simulation Models

Zong Z. Tian, Thomas Urbanik II, Roelof Engelbrecht, and Kevin Balke

One of the issues involved in using microscopic simulation models is the variation in the simulation results. This study examined some of the more popular microscopic traffic simulation models, CORSIM, SimTraffic, and VISSIM, and investigated the variations in the performance measures generated by these models. The study focused on the capacity and delay estimates at a signalized intersection. The effects of link length, speed, and vehicle headway generation distribution were also investigated. With regard to variations in performance measures, the study found that CORSIM yields the lowest variations, whereas SimTraffic yields the highest. The highest variation in each simulation model normally occurs when the traffic demand approaches capacity. It was also found that delays are affected by the link length and speed in simulation models. Such an impact on delays is closely related to the range of speed variations. In general, shorter links and higher link speeds result in lower delays. There is no strong evidence that the headway distribution used to generate vehicles in the simulated network has any effect on capacity and delay estimates. Multiple simulation runs are necessary to achieve an accurate estimate on the true system performance measures. With a 10% error range in estimated delay, two to five runs may be enough for under-capacity conditions, but more than 40 multiple runs may be necessary to accurately estimate delay at, near, or over capacity.

With advances in computing technology and the ever-increasing power of personal computers, many sophisticated stochastic microscopic simulation models have been developed in the area of transportation engineering. Improved user interfaces have significantly reduced the effort needed to code and interpret the results of these simulation models. As a result, more traffic engineers are relying on microscopic simulation models to analyze complex transportation problems when analytical methods cannot provide satisfactory solutions. However, the easy-to-use features of many simulation models also present a number of challenges, one of which is that inexperienced users are not aware of the potential variations in the output of stochastic simulation models. Although some users may have realized the importance of reporting results from multiple runs, time and budget often constrain them from doing so. An important aspect of stochastic simulation modeling is that each simulation run produces different results. Therefore, it is important to understand what factors contribute to the variations. It is the purpose here to examine three commonly used traffic simulation models and to investigate the range of results that these models produce and the conditions affecting the variability.

A number of studies (*1–3*) have compared the results of various traffic simulation models with those of analytical models, such as the models in the *Highway Capacity Manual* (HCM) (*4*). These studies usually focus on comparing the results produced by different models and make recommendations on which simulation model better replicates the results of either the HCM or field data. The results from simulation models are usually based on the average values from an arbitrary number of runs and simulation times, for example, 10 runs with a 1-h simulation time. No studies, however, have investigated how the performance measures vary among different simulation models. Traffic engineers and researchers need to know when it is necessary to perform multiple runs and how many runs are needed in order to yield reasonable estimates in the performance measures.

The scope of this paper is limited to signalized intersection operations. First, a simple case with a single-lane approach and 100% through traffic is presented. Another case considers a single-lane approach with a right-turn pocket, where the variations depend not only on general driver and vehicle characteristics but also on the stochastic process of queue blocking occurrences. Three of the most popular simulation models currently being used in the United States are examined: CORSIM (*5*), SimTraffic (*6*), and VISSIM (*7*). Finally, a summary and some conclusions are provided, and recommendations are then made for general applications of simulation models in practice.

SIMULATION MODEL VERSUS ANALYTICAL MODEL

Delay is one of the major performance measures for transportation systems. However, different methodologies are usually used to calculate delay in simulation models and analytical models such as those of the HCM. Detailed results and discussions on this issue can be found in several studies (*1, 2*). Although the focus of this paper is on the variations in performance measures, it might be useful to illustrate what results are expected when the default parameters of each simulation model are used. The differences between the delays produced from simulation and those calculated using the HCM methodology for signalized intersections are summarized below:

- The HCM reports an average control delay, which includes the deceleration, queue moving time, stopped time, and the acceleration. However, the length of an approach as well as the speed that may contribute to the acceleration and deceleration portions of control delay are not specifically considered. For example, higher speeds may require longer deceleration and acceleration times.
- Most simulation models report average total delay, which is measured as the difference in travel time at a lower speed compared

---

Z. Z. Tian, R. Engelbrecht, and K. Balke, Texas Transportation Institute, Texas A&M University System, 3135 TAMU, College Station, TX 77843-3135. T. Urbanik II, Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, TN 37996-2010.

with the travel time at the free-flow speed. Therefore, total delay includes control delay and other delays resulting from conditions such as normal congestion and car following.

- CORSIM (Version 4.32 and earlier) and SimTraffic report total delay on a link basis. Therefore, the delay associated with accelerating to free-flow speed, which typically occurs on the downstream link, is not accounted for in the delay computations. In VISSIM, however, the user can define segments on which to collect delay statistics. Therefore, delay from VISSIM can be collected to account for the acceleration by defining the correct travel time segment. A newly released version of CORSIM (5.0) incorporates a methodology to take into account the acceleration so that it is consistent with the HCM's delay definition (*1*). Since CORSIM 5.0 was still under beta testing at the time of conducting simulation runs in this research, CORSIM 4.32 was used instead. However, the major conclusions of this study remain valid.
- The HCM reports delay for the vehicles arriving during the analysis period, whereas simulation models only report delay for the vehicles departing during the analysis period. In undersaturated conditions, the input demand equals the throughput flow, and the difference is usually negligible when a relatively long simulation period (e.g., 15 min) is used. However, the difference can be significant if oversaturated conditions exist.
- Simulation models automatically take into account the residual queues from the previous time period. Although the HCM provides guidelines on how to consider the residual queue effect, most HCM-based analytical software packages, for example, the Highway Capacity Software, do not compute the delay associated with residual queues.

Unlike deterministic analytical models, simulation models are driven by samples of random variables from probability distributions. These random variables may have large variances. As a result, these estimates could, in a particular simulation run with a particular random seed, differ greatly among different runs of the model. The net effect is that there could be a significant probability of making erroneous inferences about the system under study if this variability is not taken into account. Several factors may affect the range of variations. In general, the variation can be reduced by using a longer simulation period or increasing the number of multiple runs. On the other hand, some simulation models, such as CORSIM, may have adopted variance reduction techniques (VRTs) in the model structure (*8*). The purpose of VRTs in simulation is to somehow reduce the variance of an output random variable of interest, such as delay, without disturbing its expectation (the average). Such techniques, if properly used, can normally obtain greater precision, for example, smaller confidence intervals, for the same amount of simulation effort.

STUDY CASES

Case 1: Single-Lane Approach with 100% Through Traffic

The first case studied involves a single-lane approach at a signalized intersection. The following assumptions are made for further analysis:

- The length of the approach is 760 m (2,500 ft);
- The free-flow speed is 73 km/h (45 mph);
- The traffic flow consists of through vehicles only, all of which are passenger cars with 6.1-m (20-ft) bumper-to-bumper spacing;
- The signal timing has a 90-s cycle and a 50-s effective green; and
- All other default parameters of each simulation model are used, including driver characteristics and car-following characteristics.

In microscopic simulation models, vehicles are usually generated on the basis of a certain headway distribution. CORSIM provides three types of vehicle generation distributions: uniform distribution, normal distribution, and Erlang distribution. Both VISSIM and SimTraffic use a negative exponential distribution, a special case of the Erlang distribution. The default distribution in CORSIM is uniform distribution and is used for the analysis in Case 1. The effect of different vehicle distributions is addressed in a later section of the paper.

A range of traffic flow conditions is investigated, and the HCM methodology is used to derive the delay and capacity results for each condition, with the HCM default saturation flow rate of 1,900 passenger cars per hour of green. These traffic flow conditions are included in Table 1. The capacity of the approach is obtained as follows:

TABLE 1 Traffic Flow Conditions

Traffic Demand, veh/h	HCM v/c Ratio	Uniform Delay, d_1, s/veh	Incremental Delay, d_2, s/veh	Average Control Delay, d, s/veh
106	0.1	9.4	0.2	9.6
211	0.2	10.0	0.4	10.4
317	0.3	10.7	0.7	11.4
422	0.4	11.4	1.1	12.6
528	0.5	12.3	1.7	14.0
633	0.6	13.3	2.5	15.9
739	0.7	14.5	3.9	18.4
844	0.8	16.0	6.4	22.4
950	0.9	17.8	12.1	29.9
1056	1.0	20.0	27.7	47.7
1108	1.05	20.0	41.8	61.8
1161	1.1	20.0	59.2	79.2
1214	1.15	20.0	78.7	98.7
1267	1.2	20.0	99.3	119.3

$$c = (g/C) \times s = (50/90) \times 1900 = 1056\,(vph) \qquad (1)$$

where

c = approach capacity (vph),
g = length of effective green (s),
C = cycle length (s), and
s = saturation flow rate (vph).

The traffic conditions defined in Table 1 are simulated using CORSIM, SimTraffic, and VISSIM. A total of 30 multiple runs are conducted for each traffic flow condition. A 30-min total simulation time is used for each run, and the results from the second 15-min interval are reported. Using a 15-min interval in the simulation is consistent with the 15-min analysis period used in the HCM. The same simulation settings are applied for all the study scenarios throughout this study. Both the throughput flow rates and the delays from the simulation models are analyzed.

Figures 1 through 3 illustrate the throughput flow rates from the three simulation models based on different levels of traffic demand (simulation input). The results from each individual run, the average of 30 runs, and the range of variations are depicted. The average throughput flow rates should match the traffic demand input for under-capacity conditions. Each simulation model produces a maximum throughput flow rate, which can be used as an estimate of the capacity condition for that particular model. It should be noted that each model produces a different capacity value than does the HCM since the default model settings were used in this study. In other words, no effort was made to calibrate the models to the capacity predicted by the HCM.

Different levels of variation can be observed for the three models. CORSIM produces the lowest variations, whereas SimTraffic produces the highest variations. The highest variations for each model are observed when traffic demand is close to the capacity condition. For example, at the HCM demand level of $(v/c)_{HCM} = 1.1$ (1,162 vph),

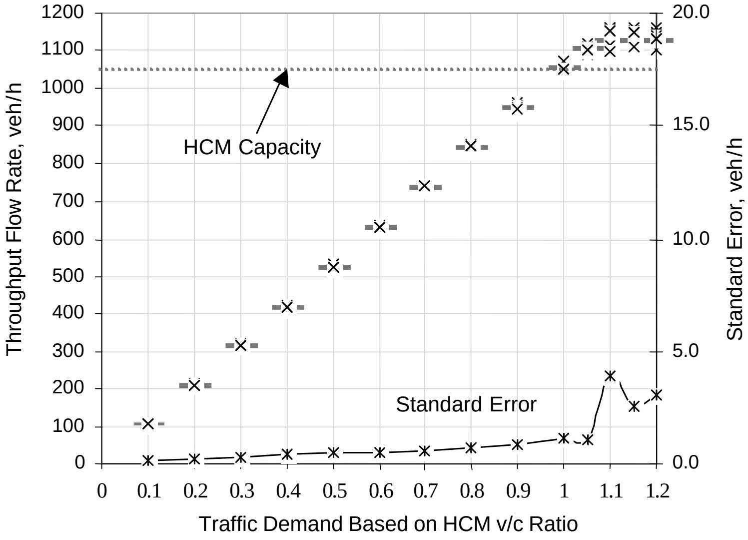

FIGURE 1 Throughput flow rates from CORSIM.

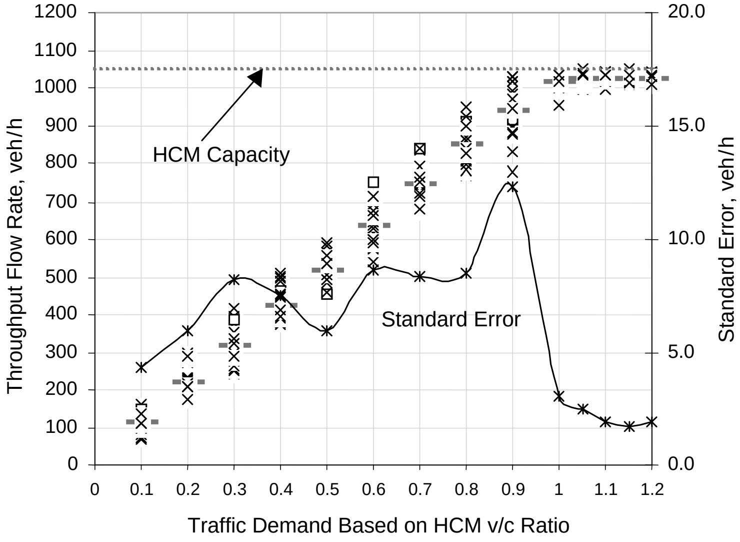

FIGURE 2 Throughput flow rates from SimTraffic.

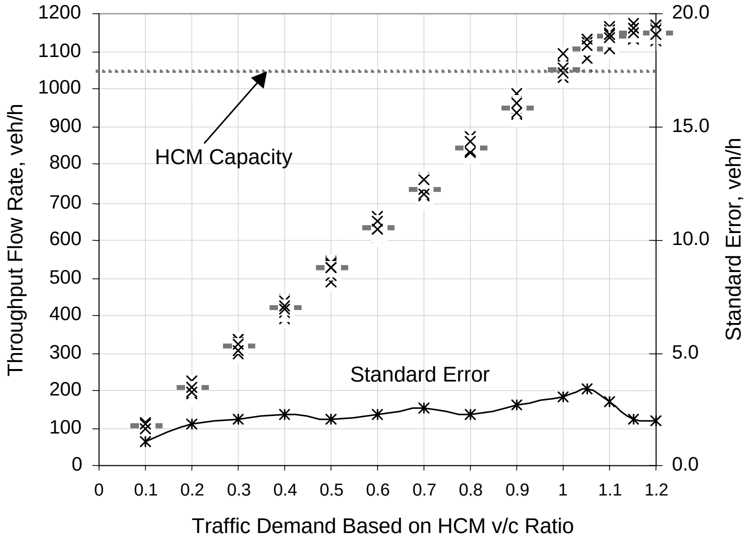

FIGURE 3 **Throughput flow rates from VISSIM.**

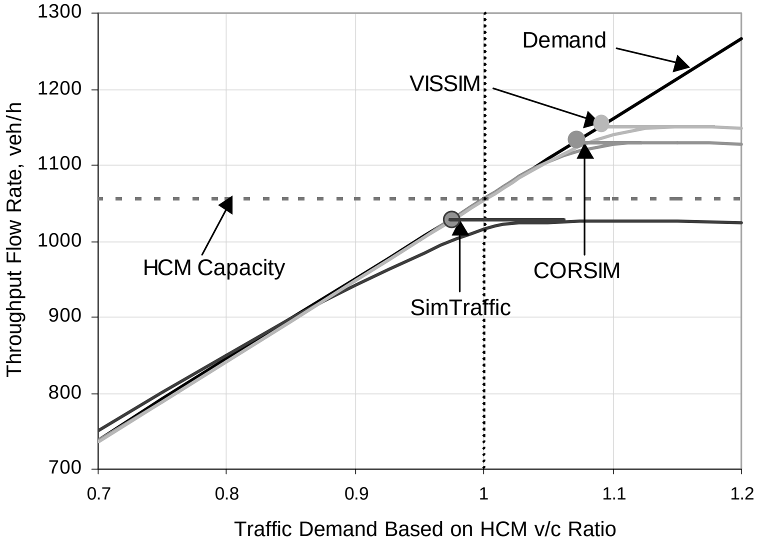

FIGURE 4 **Average throughput flow rates.**

the throughput flow rates in CORSIM range between 1,088 and 1,160 vph. For SimTraffic, the highest variation occurs at $(v/c)_{\text{HCM}}$ = 0.9 (950 vph), where the throughput flow rates range between 780 and 1,028 vph. In VISSIM the highest variation occurs at $(v/c)_{\text{HCM}}$ = 1.05 (1,109 vph), where the throughput flow rates range between 1,060 and 1,132 vph. The variations tend to decrease once the traffic demand is significantly higher than the capacity.

Figure 4 illustrates the average of the throughput flow rates from the simulation models along with the input traffic demand curve. For under-capacity conditions, the average throughput flow rates from the simulation models all match the input traffic demands closely. The capacity conditions that each simulation model produces are also demonstrated, which vary from the HCM values depending on the model. SimTraffic yields the lowest capacity value at approximately 1,025 vph, and VISSIM yields the highest capacity value at approximately 1,145 vph. The estimated capacity from CORSIM is about 1,130 vph.

Similarly, the delay results from each individual run are illustrated in Figures 5 through 7. A similar pattern can be observed in the variations from each simulation model. Again, CORSIM yields the lowest variation in delays, and SimTraffic yields the highest variation in most of the simulation runs. The highest variations occur at the capacity level and slightly beyond. For example, the delays in CORSIM range between 23.4 and 71.5 s/veh close to its capacity condition [at $(v/c)_{\text{HCM}}$ = 1.1, or 1,162 vph], and the delays in SimTraffic range between 27.9 and 248.4 s/veh at the $(v/c)_{\text{HCM}}$ ratio of 1.0 (1,056 vph). For VISSIM the delays range from 20.8 to 79.7 s/veh at the $(v/c)_{\text{HCM}}$ ratio of 1.1 (1,162 vph).

It is suspected that the lowest variations in CORSIM may be due to the use of the uniform distribution in generating vehicles. How-

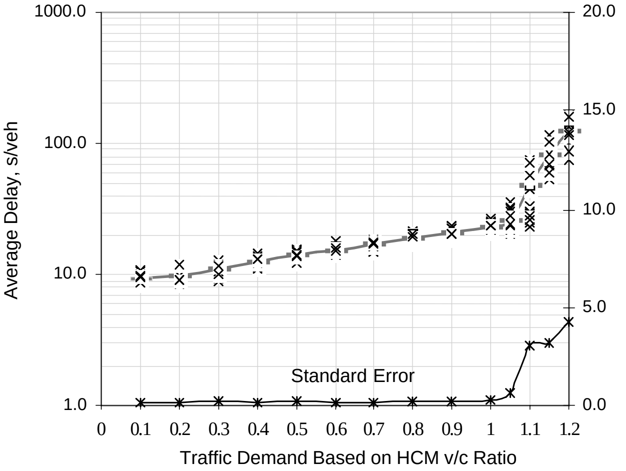

FIGURE 5 Delay results from CORSIM.

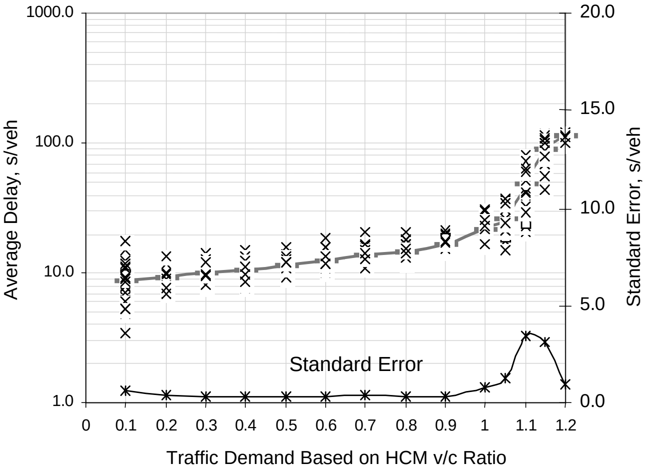

FIGURE 7 Delay results from VISSIM.

ever, a further investigation to be conducted later in this section indicates that the same level of variation is achieved in CORSIM when other types of vehicle distributions are used.

Figure 8 illustrates the average delay from the three models along with the results from the HCM. Recognizing the differences among the delays produced by different models and the specific conditions assumed in the simulation models (link length of 760 m and speed of 73 km/h), the results indicate that the delays from CORSIM match closely the HCM results for under-capacity conditions. SimTraffic results in slightly higher delays and VISSIM results in slightly lower delays than those in the HCM. A dramatic increase in delay can be observed once the capacity condition is reached. It should be kept in mind that each model produces slightly different capacity results when the default parameters are used.

Figure 8 indicates that the delays reported from the simulation models deviate somewhat from the HCM results, especially when the capacity condition is reached. Besides the differences in delay calculations in each model discussed earlier, the results from simulation may be specific to the conditions assumed in the models. The length of the link and the travel speed are the two elements that are not specifically considered in the HCM methodology. With this in mind, another set of simulation runs was conducted using a different link length and travel speed. Figure 9 shows the results from these runs with the CORSIM model.

As may be observed from Figure 9, link length and travel speed do affect the delay results. For the cases studied, lower delays result from using shorter link length, and higher delays result from using lower speed. However, the capacity values appear to remain the same. Similar results are also found in SimTraffic. However, the delays are identical in VISSIM when different speed and link length are used. There are two possible reasons for these results. The first is that a portion of the delay occurs while a vehicle is traveling on the link following a slower leading vehicle. Longer links provide more opportunity to accrue additional, noncontrol delay in this manner. The second reason is associated with the speed distribution assumed in each model. Both CORSIM and SimTraffic assume 10 different driver types, each with a specified speed relative to the link's free-flow speed. For example, Driver Type 1 has a speed that is 75% of the free-flow speed, and Driver Type 10 has a speed that is 127% of the free-flow speed. In VISSIM, however, there is no such a default

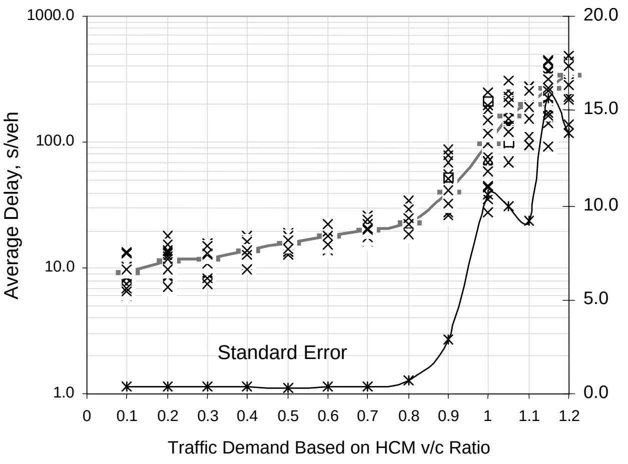

FIGURE 6 Delay results from SimTraffic.

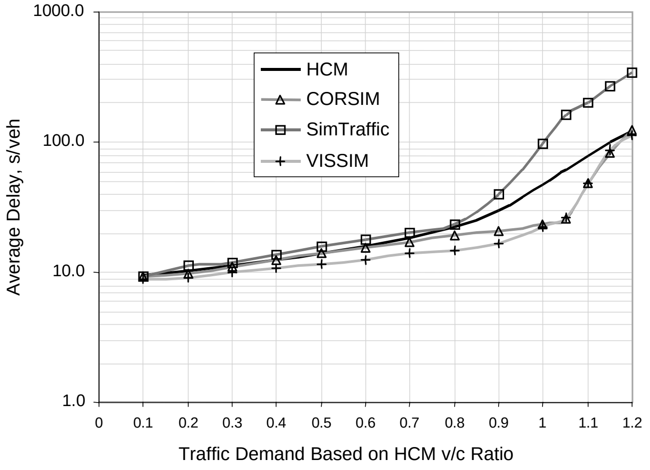

FIGURE 8 Average delay from different models.

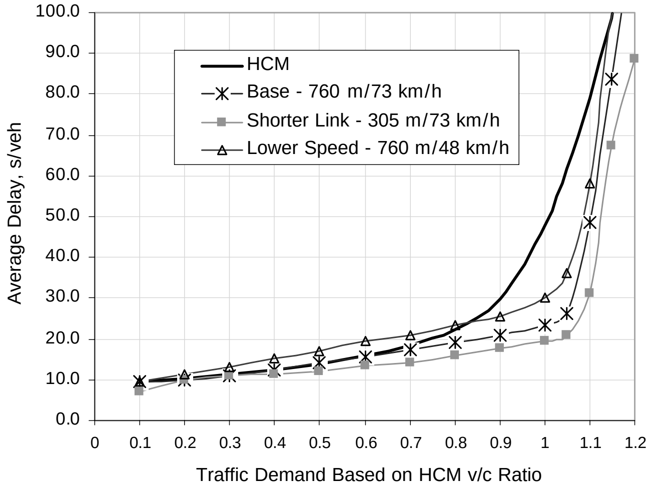

FIGURE 9 Effect of link length and speed on delay.

speed distribution, and the user has to define a speed distribution profile. In the current study, the speed distribution is defined in such a way that the 85th-percentile speed is taken as the free-flow speed, and the minimum and maximum speeds are obtained by subtracting from and adding 10 km/h to the 85th-percentile speed. As a result, both CORSIM and SimTraffic would show higher speed variations than would VISSIM. Larger speed variations result in larger potential delays to following vehicles. Since CORSIM and SimTraffic define speed variation as a percentage of free-flow speed, higher free-flow speed leads to larger speed variations and a larger amount of noncontrol delay accrued on a link.

The effect of different headway distributions is further investigated using CORSIM, where the three types of headway distributions are applied to Case 1, presented earlier. Figure 10 illustrates the delay results based on the three headway distributions: uniform, normal, and negative exponential. The standard errors of the delays are also shown. In the study, the default saturation headway of 1.8 s is replaced by 2.0 s to achieve better correlation with the HCM results. It can be seen that different headway distributions yield negligible differences in the delays. The capacity values are also almost identical. The same low level of delay variation is obtained under low traffic volume conditions. When traffic demand approaches capacity, the exponential distribution does result in different levels of variations compared with the normal and uniform distributions, but no consistent conclusions can be drawn. It is also interesting to note that a nonmonotonic trend is observed in the variations for all three models tested. It is suspected that when a link is highly saturated, drivers tend to behave more uniformly, resulting in less variation in the capacity and delays.

Case 2: Single-Lane Approach with Right-Turn Pocket

Case 2 is derived from a real-world application. At a signalized intersection, the projected traffic volume is 1,150 vph on the eastbound approach with 230 right-turn vehicles and 920 through vehicles (all passenger cars). The current HCM procedure only considers a shared-lane situation or an exclusive-lane situation; the effect of queue blocking on the right-turn pocket is not modeled specifically. On the basis of the HCM methodology, the capacities and level of service for the approach with a single shared lane and with an exclusive right-turn lane can be obtained. Using a 90-s cycle length and a 50-s effective green (identical to that in Case 1), the approach would operate at level-of-service (LOS) F with a *v/c* ratio of 1.25 if the approach has a single shared-lane configuration. With an exclusive right-turn lane, the through movement, the critical movement of the approach, would operate at LOS C with a *v/c* ratio of 0.87. The result indicates that a right-turn pocket is needed to handle the projected traffic demand. Because of the topography and the right-of-way constraints, constructing a longer-than-necessary right-turn pocket can be costly. The purpose of the investigation is to determine the minimum length of the right-turn pocket to accommodate the projected traffic demand.

Traffic operations based on different lengths of right-turn pocket (long enough to store *N* vehicles) are analyzed using CORSIM. Figure 11 illustrates the estimated capacity values for the eastbound

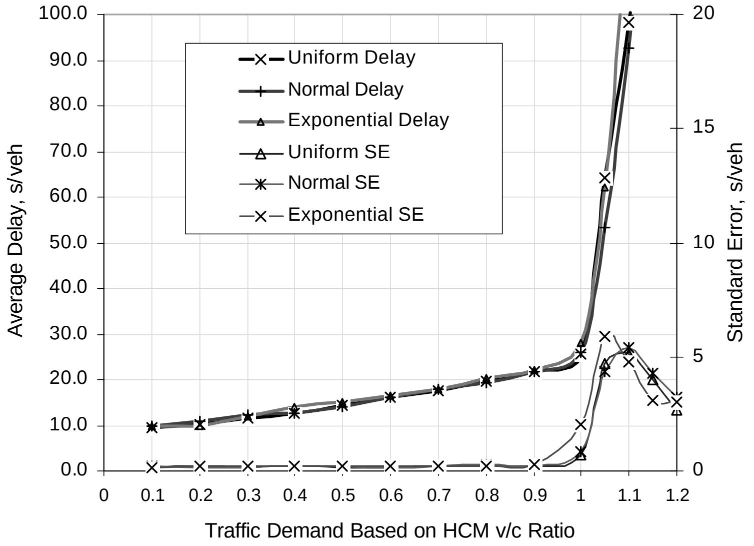

FIGURE 10 Effect of vehicle-generating distribution, showing delay results and their standard error (SE).

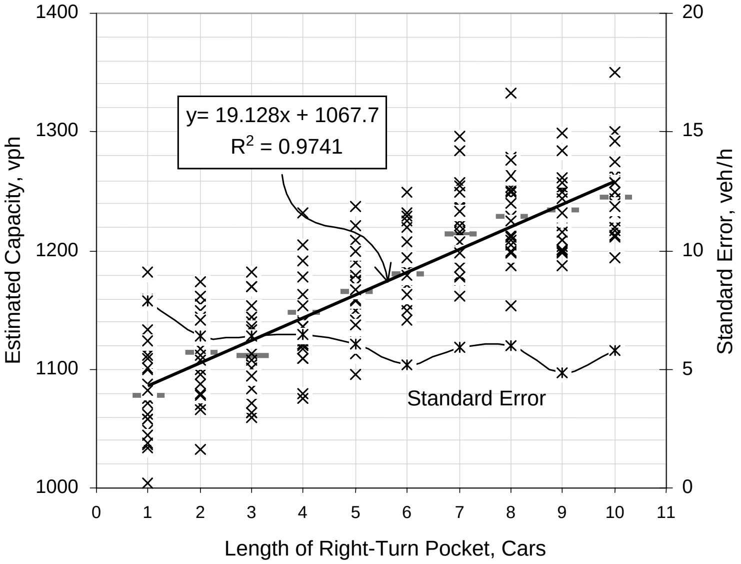

FIGURE 11 Capacity estimates with regard to length of right-turn pocket.

approach based on various lengths of right-turn pockets. The capacities in this case are estimated on the basis of the throughput flow rates, in which the input traffic demands in the simulation are set high enough to block the right-turn pocket in every cycle at the beginning of the green phase. It can be seen that the standard error ranges between 5 and 8 veh/h, which is higher than that in Case 1. Such an increase in the variations is primarily due to the added random elements in traffic flow. It can be observed that the capacity for the eastbound approach is reached at approximately $N = 4$ cars (the throughput flow rate is approximately 1,150 vph), although each individual run may still vary significantly.

Figure 12 illustrates the delay results associated with the length of the right-turn pocket. It can also be observed that high variations exist for at- or over-capacity conditions (with N fewer than four cars). Once N reaches five cars or longer, the approach will function under capacity, and the variation in delay is significantly reduced. Although over-capacity conditions also exist with N between 2 and 4, the variation in delay is not as high as that in Case 1, and the average delays are still within the range between LOS C and D. From the results of the analysis, it may be recommended that the right-turn pocket be designed with a length of two or three vehicles to minimize the construction cost while still maintaining acceptable operations most of the time. However, occasional breakdowns may be expected with this design.

NUMBER OF SIMULATION RUNS

Because of the variation in results of microsimulation models, multiple runs must be conducted in order to provide an accurate estimate of the true performance measures. From the theory of probability and

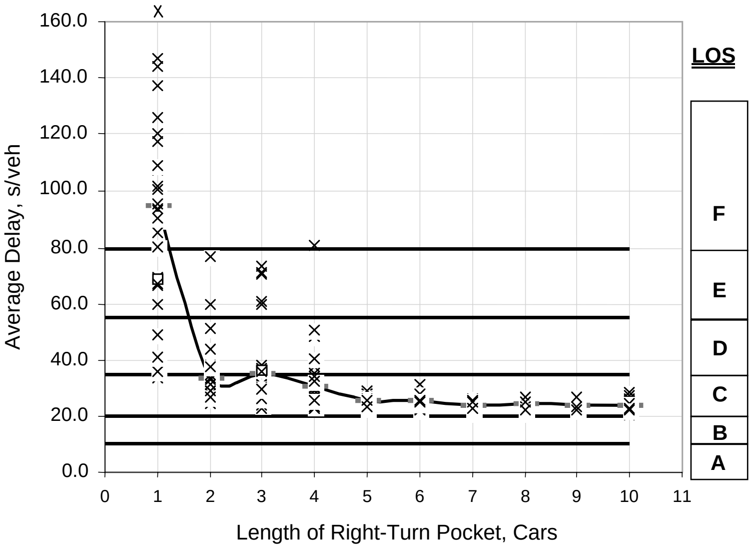

FIGURE 12 Delay estimates with regard to length of right-turn pocket.

statistics, Equation 2 can be used to estimate the required number of runs to provide an estimate of the mean with a specified confidence interval and an error range:

$$n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2 \tag{2}$$

where

n = required number of simulation runs;
σ = sample standard deviation (based on 30 multiple runs from this study);
$z_{\alpha/2}$ = threshold value for $100(1-\alpha)\%$ confidence interval; with 95% confidence interval, $z_{\alpha/2} = 1.96$; and
E = allowed error range.

Using Equation 2 and the delay results presented in Case 1, the required number of runs is calculated with a 95% confidence interval and a 10% error range, as presented in Table 2. A 10% error range is considered acceptable for general practice since it provides sufficient accuracy to predict the correct level of service about 60% of the time over a delay range between 0 and 80 s. For the remaining 40% of the time, the predicted level of service would be within one level of service of the actual level of service.

As the results in Table 2 indicate, the highest number of runs is required for at- and over-capacity conditions because of the high variations. For example, a total of 45 runs is necessary for CORSIM at the level of $(v/c)_{\mathrm{HCM}} = 1.1$. An even higher number of runs is necessary for SimTraffic and VISSIM because of the high variation in the resulting delay measures. Both SimTraffic and VISSIM also indicate a higher number of runs at the lower traffic demand levels. Because the delay values are usually small for low traffic demand conditions, a biased estimate based on a fewer number of runs may not be a concern in general practice. In general, two or five multiple runs may be sufficient for under-capacity conditions. However, when the traffic demand is at or over the capacity level, at least 40 runs may be necessary, depending on the model to be used. The number of runs may be reduced by using a longer simulation time, as shown by the VISSIM 15-min and 1-h results in Table 2. However, only a 15-min period would be consistent with the HCM procedures. A dilemma situation occurs in practical applications, however, in that the capacity condition is usually an unknown parameter. As a general guideline, it is recommended that at least two runs be conducted under any condition. Whenever a capacity condition is suspected (a high delay is obtained from an individual run), more runs are necessary to achieve more accurate estimates.

SUMMARY AND CONCLUSIONS

The results of an investigation into the variation produced by three commonly used microscopic traffic simulation models—CORSIM, SimTraffic, and VISSIM—are presented. Variations in both the delay and the throughput traffic flow rates are examined on the basis of different traffic flow conditions. The following conclusions are reached this study:

- For the three models tested, CORSIM produces the lowest variation in both delay and throughput flow rates, whereas SimTraffic produces the highest variations. The highest variations occur when capacity conditions are reached. The lower variations in CORSIM may due to the variance reduction techniques applied in the CORSIM model. When the average values are considered from multiple runs, the throughput flow rates from all three models match the input demand closely for under-capacity conditions.
- The three simulation models tested produce different results when the default traffic flow parameters from each simulation model are used. In general, VISSIM produces the highest capacity and lowest delay estimates, and SimTraffic produces the lowest capacity and highest delay estimates.
- All three simulation models tested in this study illustrate a nonmonotonic trend in the variations of the performance measures. It is suspected that when a link is highly saturated, drivers tend to behave more uniformly, resulting in reduced variations in the performance measures.

TABLE 2 Required Number of Simulation Runs

Traffic Demand (Based on HCM v/c Ratio)	CORSIM	SimTraffic	VISSIM-15 min	VISSIM-1 h
0.1	2	18	56	6
0.2	3	15	18	2
0.3	4	13	10	2
0.4	2	8	9	1
0.5	2	4	7	1
0.6	1	6	8	2
0.7	1	5	7	1
0.8	1	11	6	1
0.9	1	58	4	1
1.0	2	137	13	2
1.1	45	25	57	18
1.2	14	18	1	1

NOTE: Traffic demand is determined based on the v/c ratio from the HCM. Each simulation has a different capacity condition. The HCM v/c ratio at 1.0 does not necessarily mean the capacity condition for a particular model.

- Although link length and speed are not factors in the HCM methodology for delay calculation, the study found that link length and link speed directly affect the delays in microsimulation models. In general, the simulation models would produce lower delays with shorter links or higher speeds. Such an impact is mainly due to the speed variability and car-following behavior. The headway distribution used to generate the vehicles does not show a significant effect on delay and capacity estimates as well as on the variations, on the basis of a test performed with CORSIM.
- Variations can be reduced by either using a longer simulation period or conducting more runs. It is also important to realize the differences between simulation and analytical method when the performance measures, such as delay, are reported. Although longer runs are desired to reduce the variations, only a 15-min period would be consistent with the HCM procedure, especially when varied traffic demand is modeled over time.
- To produce a reasonable estimate of the average conditions, it is recommended that at least two runs be conducted for low to medium traffic demand level. For near- and over-capacity conditions, at least 20 runs are recommended. This recommendation is based on the simplest case tested in this study. Higher variations are expected for most practical applications; therefore, more simulation runs may be necessary. Simulation models should include automated utilities to perform multiple runs and aggregate the results. Alternatively, simulation models should be designed to allow users to develop such utilities.

REFERENCES

1. Zhang, Y.L., L. Zhang, and L. Owen. Control Delay Definition, Modeling, and Calculation in CORSIM. Presented at the 80th Annual Meeting of the Transportation Research Board, Washington, D.C., Jan. 2001.
2. Milam, R. T., and F. R. T. Choa. HCM and CORSIM—Resolving the Differences. In *Compendium of Papers,* District 6 Annual Meeting, San Diego, Calif., Institute of Transportation Engineers, Washington, D.C., 2000.
3. Bloomberg, L., and J. Dale. Comparison of VISSIM and CORSIM Traffic Simulation Models on a Congested Network. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1727,* TRB, National Research Council, Washington, D.C., 2000, pp. 52–60.
4. *Highway Capacity Manual.* TRB, National Research Council, Washington, D.C., 2000.
5. *CORSIM User's Manual, Version 4.32.* Office of Safety and Traffic Operations R&D, Intelligent Systems and Technology Division, FHWA, U.S. Department of Transportation, June 1999.
6. *Synchro 4, Traffic Signal Coordination Software.* Trafficware Corporation, Albany, Calif., 1999.
7. *User Manual, VISSIM 3.50.* PTV Planung Transport Verkehr AG, Karlsruhe, Germany, 2001.
8. Rathi, A. K., and M. M. Venigalla. Variance Reduction Applied to Urban Network Traffic Simulation. In *Transportation Research Record 1365,* TRB, National Research Council, Washington, D.C., 1992, pp. 133–146.

Publication of this paper sponsored by Committee on Highway Capacity and Quality of Service.