

1
2
3
4
5
6
7
8 **Modelling Day-to-Day Variability of Intersection Performance using Micro-Simulation**
9

10
11
12 Zeeshan Abdy
13 PhD Candidate
14 Department of Civil and Environmental Engineering,
15 University of Waterloo
16 Waterloo, Ontario, N2L 3G1, Canada
17 Tel: (519) 888-4567 ext. 6596
18 Fax: (519) 888-4349
19 E-mail: zrabdy@uwaterloo.ca
20

21
22
23 Bruce Hellinga, PhD, PEng
24 Associate Professor
25 (Corresponding Author)
26 Department of Civil and Environmental Engineering,
27 University of Waterloo
28 Waterloo, Ontario, N2L 3G1, Canada
29 Tel: (519) 888-4567 ext. 32630
30 Fax: (519) 888-4349
31 E-mail: bhellinga@uwaterloo.ca
32
33
34
35
36

37 Accepted for presentation at TRB 2008 and
38 under review for publication in Transportation Research Record
39

40 Nov. 14, 2007
41
42
43
44
45
46
47
48
49
50

ABSTRACT

The performance of signalized intersections in the field exhibit significant variation. Even when considering a specific time period (i.e. PM peak period) on a weekday, the average delay experienced by vehicles varies from day to day. This variation arises from a number of sources including variations in peak period traffic demands.

The use of micro simulation tools to analyze and/or predict the performance of intersections, and more generally road sub-networks, is now common practice among traffic engineers. Popular micro simulation tools, such as Paramics, VISSIM, Integration, Aimsun, and NetSim are considered “stochastic” models in that they use pseudo random numbers to control random processes within the simulation, such as lane changing decisions, desired speeds, etc. As a result, traffic engineers and simulation model users typically carry out several model runs, each with a different random number seed, for each set of traffic and control conditions. Often, the results from the replications are averaged in the hope that the mean of the multiple runs is a reliable predictor of the average conditions that would occur in the field.

In this study we seek to address two key questions, namely:

(1) Does the current practice of conducting multiple simulation runs, each with a different random seed but with the same traffic demands, adequately replicate the day-to-day variability typically observed in the field?

(2) What is the best method by which to use simulation models to reflect day-to-day variability in intersection performance?

Three methods of modelling day-to-day variability of intersection performance (in terms of delay) are examined. Method 1, reflecting the typical current practice, introduces variability through the use of different random number seeds but traffic demands are held constant. Method 2 consisted of using only a single random number seed but traffic demands are randomly selected from a distribution fit to field data. Method 3 consisted of using different random seeds and randomly selecting traffic demands from the field calibrated distribution. For each method, eleven traffic demand scenarios were developed encompassing intersection volume to capacity (v/c) ratios ranging from 0.6 to 1.10.

The results show that the current method of incorporation variability in simulation runs (i.e. Method 1) does not adequately capture the day-to-day variation observed in field peak hour approach volumes. More specifically,

1. The use of Method 1 results in intersection peak hour approach volumes that exhibit only about half the variation observed in the field. Methods 2 and 3 exhibited variations in peak hour approach volumes that were very similar to those observed in the field.
2. The average intersection delays obtained from the three methods differ by as much as 47% suggesting that the method used to simulate day-to-day variability in intersection delays has a significant influence on the results.
3. The differences in average intersection delays are sufficiently small at v/c less than about 0.8 that they are not likely to be of practical significance.

KEYWORDS: traffic signals, variability, intersection delay, simulation, peak hour volume.

1.0 INTRODUCTION

Increasingly micro simulation models are being used by traffic engineers to analyze the expected performance of road networks operating under proposed physical or operational changes. Most of the popular micro simulation models, such as Paramics, VISSIM, Aimsun, Integration, and NetSim, are discrete event simulation models that simulate the movement of individual vehicles as they travel through the road network. Conceptually, each of these models can be considered to be composed of various sub-models, each of which defines the logic associated with a specific behavioural attribute, such as car-following, lane selection, routing, vehicle generation, etc. The micro simulation models tend to differ in the details of these behavioural sub-models.

These micro simulation models are considered “stochastic” in that many of the behavioural sub-models contain random distributions. Specific values are selected from the distribution using a pseudo random number selected from a sequence of random numbers. A random number *seed* is used to generate the sequence of random numbers. Each time the micro simulation model is executed with the same random number seed the same sequence of random numbers is generated and, if all model inputs remain unchanged, the model produces the same outputs.

However, if the random number seed is changed and all other model inputs remain unchanged, then model outputs vary due to the different values selected from the distributions within the behavioural sub-models. For example, vehicles are generated at the entry nodes (origins) based on the input volumes and an assumed headway distribution. When a vehicle is generated the simulation model assigns driver and vehicle characteristics such as vehicle type (car, bus, truck, etc.), vehicle length, width, maximum acceleration and deceleration, maximum speed, maximum turn radius, etc. For each driver, values are assigned for driver aggressiveness, reaction time, desired speed, critical gaps (for lane changing, merging, crossing), destination (route), etc. Most of these attributes are selected from distributions on the basis of the pseudo random numbers.

As a result of the randomness of model results, traffic engineers and simulation model users typically carry out several model runs for each set of traffic and control conditions (each run with a different random number seed) in order to imitate the randomness in field observations. Often, the results from the replications are averaged in the hope that the mean of the multiple runs is a reliable predictor of the average conditions that would occur in the field.

This paper presents the finding of a study that has been carried out to determine the extent to which the use of multiple runs each with a different random number seed captures the degree of variability typically present in real networks.

This paper seeks to answer the following specific questions that begin to address these issues with respect to modelling arterial networks;

1. What level of variability is introduced into the intersection approach volumes by the random number seeds?
2. How does this level of variability compare to what is observed in the field?
3. How does the variability in approach volumes vary with the volume to capacity ratio?
4. What is the distribution of intersection delay resulting from different random seeds at different volume to capacity ratio?
5. What implications do these results have in terms of appropriate methods for using micro simulation models to evaluate alternatives?

The following section summarizes findings from previous research that is relevant to this paper. Section 3 examines the randomness in peak hour approach volumes that is created by the VISSIM simulation model using each of the three methods for representing day-to-day variability. These results are compared to the variation in peak hour approach volumes observed in field data. Section 4 explores the variation in intersection performance (in terms of delay) corresponding to each of the three methods of representing day-to-day variability. Finally, Section 5 provides conclusions and recommendations.

2.0 LITERATURE REVIEW

2.1 REPRESENTING VARIABILITY IN MICRO-SIMULATION MODELS

There is a common consensus among traffic engineers and microscopic simulation model users on the need for performing multiple runs when using any stochastic simulation model. When evaluating the expected impact of one or more possible future treatments (e.g. new traffic control devices, operating strategies, policies, etc.) the scenarios simulated typically consist of the base traffic conditions (e.g. peak and off-peak traffic demands) modelled without the treatment(s) and then modelled with the treatment(s). In current practice, it is typical to simulate each with and without treatment scenario multiple times, each time with a different random number seed but holding all other simulation model inputs constant. The median or the average results from multiple simulation runs using different seeds are assumed to reflect the average traffic condition of a specific scenario.

There are ample examples in the literature in which this approach has been taken for various models including SimTraffic (Shaaban et al., 2005), PARAMICS (Chu et al., 2004), VISSIM (Sayed et al., 2004), Integration (Dion et al., 2005) and TRAF-Netsim (Garrow et al., 1997).

However, this is not the only method of modelling stochastic variability. In this paper we identify three methods of modelling day-to-day variability of transportation system performance (Table 1). Method 1, reflecting the typical current practice, introduces variability through the use of different random number seeds but other model inputs, including traffic demands are held constant. In general, many of the model inputs that are held constant in Method 1 are in fact subject to stochastic variation and ought to be represented by a distribution. Peak hour approach volumes are particularly important with respect to the performance of signalized intersection. Other factors that may influence intersection performance and are subject to stochastic variations, such as the peak hour factor, saturation flow rate, etc. are considered to have a smaller degree of variability (Hellings and Abdy, 2007) and therefore the variability of these other factors is not separately modelled. Consequently, Method 2 consists of using only a single random number seed but traffic demands are randomly selected from a distribution that is calibrated to field data. Method 3 consists of using different random seeds and randomly selecting traffic demands from the field calibrated distribution. In each method, multiple runs of the simulation are conducted in order to estimate both the average (mean) and variance of the transportation measures of performance that are of interest.

In the literature there is substantial variation in the methods used to decide on how many simulation runs need to be conducted. Some researchers and practitioners arbitrarily decide on the required numbers of runs. The literature shows that there is a large variation in the number of runs considered to be necessary such as 2 replications (Garrow et al., 1997), 3 replications (Smith et al., 2006), 5 replications (Sayed et al., 2004) and 30 replications (Dion et al., 2004).

Rather than arbitrarily selecting the number of runs, it is possible to determine the number of simulation runs required to achieve a specified accuracy (Hellings et al., 2007; Birst et al., 2007; Chu et al., 2004; and Shaaban et al., 2005). An initial set of simulation runs is executed first and the mean

and standard deviation of the performance measure is calculated. Using this mean and standard deviation, the required number of simulation runs is calculated by

$$N = \left(t_{\alpha/2} \frac{\delta}{\mu \varepsilon} \right)^2 \quad - 1$$

Where μ and δ are the mean and standard deviation of the performance measure based on the initial set of simulations runs; ε is the allowable error specified as a fraction of the mean μ ; and $t_{\alpha/2}$ is the critical value of the t-distribution at significance level α .

In a study by Birst et al., (2007) CORSIM, SimTraffic and VISSIM models were compared using various levels of traffic congestion under pre timed signalized control. Using equation 1, the authors determined the number of required runs at different volume to capacity ratio. It was found that more runs are needed as the volume reached capacity due to increased variation. The increase in variation is in line with the study by Hellinga et al., (2007) and Sullivan et al., (2006).

However the use of Equation 1 does not assist the model user in determining whether or not the variation arising from the simulation model (from the use of different random seeds) appropriately imitates the variation observed in the field.

In an effort to examine the implications of using Method 1, 2, or 3 for modelling variability, we focus on signalized intersection performance as measured by average delay during a peak hour. In the next section we examine the day-to-day variability that exists in arterial peak hour approach volumes and use these field data to calibrate a distribution that is required for Methods 2 and 3.

2.2 VARIABILITY OF PEAK HOUR VOLUMES IN THE FIELD

There appears to be relatively little work reported in the literature examining the variability of intersection performance in terms of day-to-day variations. Sullivan et al., (2006) conducted an analysis to examine the impact of day to day variations in urban traffic peak hour volumes on intersection service levels. Using weekday data from 22 directional continuous traffic counting stations in the City of Milwaukee, the authors found that the coefficient of variation¹ (COV) of peak hour traffic volume ranged from 5% to 16% with a mean of 8.9%.

Hellinga et al., (2007) conducted a similar analysis. Using weekday data from 16 continuous traffic counting stations located mid-block on major arterial roadways in the City of Waterloo, the authors found that the COV ranged from 5.4% to a maximum of 13.1% and on average is equal to 8.7%.

Hellinga et al., (2007) also developed a linear regression model showing that the COV of peak hour volume decreases as the mean peak hour approach volume increases. Although the regression intercept and coefficient are statistically significant at the 95% level, the regression explains only a small portion of the variance within the data (adjusted $R^2 = 0.15$) and therefore must be viewed with scepticism.

In this study we have also acquired traffic counts for a 12 month period from 9 permanent count stations located in the City of Toronto. The COV of peak hour volume was computed for these locations using PM peak hour week-day non-holiday data and found to vary from 3.1% to 9.5% with a mean of 7.0%.

¹ Coefficient of variation = standard deviation / mean

For all three data sets (Figure 1), the peak hour volumes were found to follow a Normal distribution. Consequently, on the basis of these data, we conclude that a Normal distribution with a coefficient of variation of 0.084 is suitable to model the variation in day-to-day peak hour approach volumes on arterial roadways.

In the next section we simulate a hypothetical signalized intersection approach using the VISSIM model and introduce variability using each of the three methods identified in Section 2.1. We examine the performance of these methods in terms of the resulting distribution of peak hour approach volumes.

3.0 VARIATION OF PEAK HOUR VOLUME

3.1 HYPOTHETICAL INTERSECTION

A hypothetical signalized intersection approach was simulated. The approach consisted of three exclusive through lanes. All lane widths, grade, curb radii, etc. were considered to be ideal with no on-street parking, no transit vehicles, and adequate storage and discharge space. The intersection geometry was developed using links and connectors and modelled in VISSIM.

The intersection approach was controlled by a two-phase signal timing plan with a cycle length of 80s; 38s effective green for the modelled approach; 34s effective green for the cross street phase; and 4 seconds of inter green between each phase. Right-turn on red was not permitted and no turning movements were modelled.

Six traffic demand scenarios were developed such that mean intersection volume to capacity (v/c) ratios ranged from 0.8 to 1.10. Each demand scenario was replicated 50 times using each of the three methods. For Method 1, each of the 50 replications used a different random number seed but all approach volumes remained constant. For Method 2, a single random number seed was used for all replications, but for each replication the peak hour approach volumes were randomly selected from a Normal distribution having a mean equal to the mean volume for the v/c demand scenario being simulated and a standard deviation equal to 0.084 times the mean (i.e. COV = 0.084). For each replication of Method 3, a different random number seed and a random selected peak hour approach volume (from the same distribution as used in Method 2) was applied. For all cases, the traffic stream was assumed to consist of only passenger cars.

Vehicles were generated within the simulation model for 60 minutes. An additional 30 minutes were used to ensure that all generated vehicles were able to complete their trips. For all simulations, the signal timing plan and all other inputs except the approach volumes and random number seed remained unchanged.

The VISSIM model was calibrated to have a base saturation flow rate of 1900 pcphpl.

3.2 RESULTS

The volume of vehicles generated by the simulation was recorded for each of the replications conducted for each of the five traffic demand scenarios. The mean, standard deviation and COV of the generated volumes were also computed.

Figure 2 shows the resulting mean approach peak hour volume generated by the simulation model for each of the three methods as a function of the v/c ratio as well as the expected average peak hour volume. The results show that all three methods produce average approach volumes that are very similar to the target average. These results suggest that the distribution of approach volumes that results from the use of random number seeds is symmetrical.

Figure 3 shows the coefficient of variation of the generated peak hour volumes for each of the three methods. The average coefficient of variation of peak hour approach volumes observed in the field data (i.e. $COV=0.084$) is also shown in the Figure as a benchmark against which the results from the three methods can be evaluated.

It is evident from the results that the COV of peak hour volume resulting from Method 1 is substantially smaller than the variation exhibited in the field data. The COV from Method 1 ranged from approximately 4.4% to 3.7% with an average of 4.1% (only 49% of the COV associated with the field data).

Method 2 resulted in an average COV of peak hour volumes of 9.4% which is very similar to the COV of the field data. This result is not surprising since the simulation runs were set up with specified traffic demands that were drawn from a distribution having a $COV = 8.4\%$.

Method 3 resulted in an average COV of peak hour volumes of 9.6%. This result suggests that the use of random seeds with random volumes does not substantially increase the resulting variability in the generated volumes.

The COV can be used to characterise the variability within the approach volume distribution, however we are also interested to determine the shape of the distribution. The resulting peak hour volumes for each scenario were tested to determine the shape of distribution. For this task, the Kolmogorov-Smirnov (KS) test was used to determine if each distribution could be adequately described by the Normal distribution at the 99% level of confidence. It was found that the distributions of generated peak hour volume for all demand scenarios for all three methods follow the Normal distribution.

These results indicate that though Method 1 introduces variability into the peak hour approach volumes generated by the simulation model this level of variability is substantially less than that typically observed in the field. However, in most cases model users are interested in measures of performance (i.e. intersection delay) rather than the approach volumes. Consequently, the next section examines the impact of method used to simulate variability (i.e. Methods 1, 2, or 3) on intersection delay.

4.0 IMPACT ON INTERSECTION PERFORMANCE

Ideally, the distributions of intersection delay resulting from the use of Methods 1, 2, and 3 would be compared to distributions of delay obtained from the field (similar to that done for the variation of volumes in Section 3). Unfortunately, intersection delay can not be computed from permanent count station data and direct measurement of delay is a resource intensive effort that is sensitive to measurement errors (Teply and Evans, 1989; Teply, 1989; Colyar and Rouphail, 2003). Not surprisingly then, we were unable to locate an existing database or collect sufficient data to create our own database of measured intersection delays that would contain a sufficient number of observations for a range of intersections experiencing a range of v/c conditions. Consequently, we opted to consider a hypothetical intersection and to explore the impact that Methods 1, 2 and 3 have on estimates of mean and variance of delay.

4.1 HYPOTHETICAL INTERSECTION

A hypothetical 4-leg intersection with a single exclusive lane on each approach was created. The intersection was controlled by a 2-phase fixed time signal running the same signal timings as used in Section 3. All geometry, queuing space, etc. was considered to be ideal.

For each of the three methods described in Table 1 and Section 2.1, eleven traffic demand scenarios were developed encompassing intersection volume to capacity (v/c) ratios ranging from 0.6 to 1.10. Each demand scenario was simulated 100 times (replications). For all cases, the traffic stream was assumed to consist of only passenger cars. Intersection volume was generated for 15 minutes. Adequate link lengths were provided so that even at the highest volume to capacity ratio queues did not spill off of the network. An additional 30 minutes were provided to ensure that adequate time is available to capture delay of vehicles still in the system. The total delay including the car-following delay was recorded. For all simulations, the signal timing plan, turning movement proportions and all other inputs, except the average approach volumes and random number seed remained unchanged.

4.2 RESULTS

Figure 4 shows the average intersection delay (seconds/vehicle) as a function of v/c ratio for all three methods. In this figure, average delay represents the peak hour intersection delay averaged over many days².

Figure 5 shows the standard deviation of intersection delay as a function of v/c ratio for all three methods.

The results depicted in Figures 4 and 5 suggest that Method 1, which produces the smallest variability in peak hour approach volumes, also produces the lowest estimate of the mean delay and the smallest variation in delay. In contrast, Method 3 provides the highest mean delay and the largest variation in delay despite the fact that Method 3 produces almost the same mean and variance of approach volumes as Method 2.

The F-test was used to determine if the variances of peak hour delay exhibited by the three methods were statistically different. A 2-tailed test at the 95% confidence limit was used (Table 2). The test results suggest that the three methods produce statistically different variances of estimated peak hour intersection delay over almost all v/c ratios examined.

The t-test was used to determine if the mean delays estimated by the three methods are statistically different from each other. A 2-tailed test at the 95% confidence limit was used assuming unequal variances (as suggested by the results in Figure 5 and Table 2). The test results (Table 3) indicate that the average delays resulting from the three methods are statistically different from each for almost v/c scenarios examined.

The relative differences in the mean delay estimated by the three methods was quantified in terms of the Relative Error (RE)

$$RE = \frac{d_i - d_1}{d_1} \quad - 2$$

Where

d_i = Average intersection delay estimated using Method i ($i = 2$ or 3)

d_1 = Average intersection delay estimated using Method 1.

Figure 6 shows the relative error for different v/c ratios. The results indicate that the estimates of average peak hour delay obtained from Method 2 may be as much as 15% larger than the

² In this study each simulation replication represents a different simulated "day"

corresponding average delay obtained using Method 1 and the estimates from Method 3 may be as much as 50% larger. These findings have several implications:

1. Intersection evaluation and design is typically based primarily on mean intersection delay (though other intersection characteristics may also be important such as queue lengths). The intersection delay estimated from a microscopic simulation model is clearly dependent on the method used to reflect variability (i.e. Method 1, 2, or 3). Depending on the method used, different estimates of mean intersection delay may result.
2. In practical terms, the absolute magnitude of the differences (and relative differences) in average delays estimated by the three methods is very small for v/c less than about 0.80. Consequently, if only the mean delays are being used as measures of performance, then the current practice (i.e. Method 1) of simulating multiple runs each with a different random seed while holding traffic demands constant, is likely adequate.
3. Though there is no opportunity for a direct comparison between the mean delays estimated by the simulation model and field delays, it seems reasonable to expect that since Method 1 significantly under represents the degree of variation of peak hour approach volumes observed in the field, and the mean delays estimated from Method 1 are significantly lower than the means delays estimated by Method 2 and 3, that the mean delays from Method 1 are also lower than those that would be observed in the field.

Increasingly, traffic engineers are interested in more than just average performance. They are also interested in variability, which is often thought of in terms of fraction of time that the intersection performs more poorly than some specified level of service. For this type of analysis, the variability of simulation results is particularly important. To illustrate, consider the fraction of days for which the intersection is expected to experience peak hour delays greater than some threshold – in this case computed as 1.4 times the mean intersection delay³. Figure 7 shows the fraction of the 100 simulated “days” (replications) for which the estimated peak hour delay exceeds the threshold delay for Methods 1 and 2. The figure illustrates that if an analyst is using a simulation model to determine how often a design is likely to fail (i.e. delay exceeds some design threshold) then the use of Method 1 (with smaller variance) tends to predict many fewer ‘failures’ than does Method 2.

From a statistical perspective, the variance of delay is also important for testing whether or not any change in the mean delay due to the evaluated treatment (e.g. traffic control strategy, technology, or policy) is statistically significant. Given that each of the three Methods provides different estimates of the mean delay and of the variation in delay, it is quite possibly, and even likely, that under some conditions, conclusions about the significance of a treatment option may vary depending on the Method used to model variability.

This possibility is particularly troubling given that the current practice is to model variability using Method 1 and the evidence provided in this study suggests that Method 1 significantly under-represents the day-to-day variability of intersection approach volumes observed in the field. This also suggests that Method 1 significantly under-estimates the mean intersection delay and the variability of intersection delay.

³ Since each method produces different mean delays (as per Figure 4) the threshold delay is computed separately for each Method and for each v/c ratio.

5.0 CONCLUSIONS AND RECOMMENDATIONS

The primary purpose of this study was to initiate an exploration (and spur discussion) of the most appropriate ways to reflect variability that occurs in the field within micro-simulation studies when these studies are used to establish the effect of a treatment option (e.g. traffic control strategy, technology, policy, etc.). On the basis of the results obtained from this study, the following conclusions can be made:

1. There appears to be at least three methods of reflecting variability within microscopic traffic simulation models. The most commonly used method in practice is to hold all simulation parameters and traffic demands constant and to conduct a set number of model runs (replications), each replication having a different random number seed.
2. This method of introducing variability significantly under-represents (approximately half) the day-to-day variability that is observed within peak hour intersection approach volumes.
3. The under-representation of variability in intersection approach volumes suggests that Method 1 also under estimates the mean delays that occur in the field.
4. Differences in mean delay appear to be very small for v/c ratios of less than about 0.8 and suggest that for these conditions, there is little practical difference between using Method 1, 2, or 3.
5. The use of Method 1 produces significantly smaller variation in peak hour intersection delays than Methods 2 and 3. As a result of this smaller variation, Method 1 also predicts many fewer 'failures' (i.e. peak hour intersection delay exceeding some threshold) than may actually be experienced.

The evidence obtained from this study seems to suggest that Method 1 is not an adequate means of capturing the day-to-day variation in peak hour approach volumes that is observed in the field. However, the study also raises a number of unanswered questions, including the following:

1. The COV of peak hour approach volumes generated by the simulation model under Methods 1, 2, and 3 seem to decline as mean peak hour approach volume increases. This observation has also been made from field data (Hellinga and Abdy, 2007; Sullivan et al., 2006). However, the sensitivity of this trend needs to be established for a wider range of mean volumes.
2. In practice, how can/should model users implement either Method 2 or Method 3 to simulate day-to-day variability? How much field data must be collected to adequately define the distribution of traffic demands? How many simulation runs must be performed?
3. This study has used only the VISSIM simulation model. Are the issues raised and results obtained in this study applicable to other commonly used microscopic traffic simulation model?

The answers to these questions will require additional research.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the following agencies/organization that provided financial support for this project:

- Transport Canada, Urban Transportation Showcase Program
- Regional Municipality of Waterloo, and
- Natural Science and Engineering Research Council of Canada.

REFERENCES

“Highway Capacity Manual” (2000) Published by the Transportation Research Board of the National Academies Washington DC.

Birst S., Baker J., and Shouman K., (2007) “Comparison of Traffic Simulation Models to HCM 2000 Using Various Traffic Levels under Pre-timed Signal Control”, Transportation Research Board Annual Meeting CD ROM.

Colyar, J.D. and N.M. Rouphail (2003). “Measured Distributions of Control Delay on Signalized Arterials.” Transportation Research Record 1852, TRB, National Research Council, Washington D. C., pp. 1-9.

Chu L., Liu H., Oh J., and Recker W., (2004) “A Calibration Procedure for Microscopic Traffic Simulation”, Transportation Research Board Annual Meeting CD ROM.

Dion F., and Rakha H., (2004) “Integration of Transit Signal Priority within Adaptive Traffic Signal Control Systems” Transportation Research Board Annual Meeting CD ROM.

Fu, L. and B. Hellinga (2000) “Delay variability at signalized intersections”, Transportation Research Record: Journal of the Transportation Research Board No 1710 pp. 215-221.

Garrow M and Machemehl R. (1997) “Development and Evaluation of Transit Signal Priority Strategies” Southwestern Region University Transportation Center, Center for Transportation Research, The University of Texas, Austin, Texas 78712.

Hellinga B.R. and Abdy, Z.R. (2007) “Signalized Intersection performance considering day-to-day variability in peak hour volumes”. Presented at the Annual Meeting of the Transportation Research Board, Washington DC, 2007.

Levinson, H.S., Roark, A.L., and Guhin, J.S. (1979) “Sampling Vehicle Kilometers of Travel” in Transportation Research Record 728, Washington, DC, 1979, pp 65-70.

Sayed T and Abdelfatah A. (2004) “Impacts of Various Parameters on Transit Signal Priority Effectiveness”, Journal of Public Transportation, Vol. 7, No. 3, 2004.

Shaaban S.K and Radwan E. (2005) “A Calibration of Validation Procedure for Microscopic Simulation Model: A Case Study of SimTraffic for Arterial Streets”, Transportation Research Board Annual Meeting CD ROM.

Smith M. and Sadek A., (2006) “Challenges Calibrating a Large-scale Microscopic Simulation Model of a Diverse Urban, Suburban & Rural Network: A Practical Guide”, Transportation Research Board Annual Meeting CD ROM.

Sullivan, D., Levinson, H.S., Bryson, R.W. (2006) “Effects of Urban Traffic Volume Variations on Service Levels”. Paper # 06-2381 presented at the 85th annual meeting of the TRB.

Teply, S., and G. D. Evans (1989). “Evaluation of the Quality of Signal Progression by Delay Distributions.” Transportation Research Record 1225, TRB, National Research Council, Washington, D.C., pp. 1-7.

Teply, S. (1989). “Accuracy of Delay Surveys at Signalized Intersections” Transportation Research Record 1225, TRB, National Research Council, Washington, D.C., pp. 24-32.

LIST OF TABLES:

Table 1: Description of three methods for simulating day-to-day variability of intersection performance

Table 2: F-test results from comparing estimates of variance of delay

Table 3: t-test results from comparing estimates of mean delay

Table 1: Description of three methods for simulating day-to-day variability of intersection performance

Method	Random Number Seeds	Input Volumes
Method 1	Random	Constant
Method 2	Constant	Random
Method 3	Random	Random

Table 2: F-test results from comparing variance of delay

		V/C										
		0.6	0.7	0.8	0.85	0.90	0.925	0.95	0.975	1.00	1.05	1.10
Method 1 vs 2	P-Value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Conclusion	Dif	Dif	Dif	Dif	Dif	Dif	Dif	Dif	Dif	Dif	Dif
Method 2 vs 3	P-Value	0.331	0.028	0.018	0.000	0.000	0.000	0.000	0.001	0.004	0.015	0.036
	Conclusion	Same	Same	Dif	Dif	Dif	Dif	Dif	Dif	Dif	Dif	Same
Method 1 vs 3	P-Value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Conclusion	Dif	Dif	Dif	Dif	Dif	Dif	Dif	Dif	Dif	Dif	Dif

Dif = Means are statistically different.

Table 3: t-test results from comparing estimates of mean delay

		V/C										
		0.6	0.7	0.8	0.85	0.90	0.925	0.95	0.975	1.00	1.05	1.10
Method 1 vs 2	P-Value	0.124	0.373	0.023	0.000	0.000	0.001	0.001	0.003	0.006	0.016	0.054
	Conclusion	Same	Same	Dif	Dif	Dif	Dif	Dif	Dif	Dif	Dif	Same
Method 2 vs 3	P-Value	0.000	0.112	0.075	0.023	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Conclusion	Dif	Same	Same	Dif	Dif	Dif	Dif	Dif	Dif	Dif	Dif
Method 1 vs 3	P-Value	0.006	0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Conclusion	Dif	Dif	Dif	Dif	Dif	Dif	Dif	Dif	Dif	Dif	Dif

Dif = Means are statistically different.

LIST OF FIGURES:

Figure 1: COV of peak hour volumes observed in the field

Figure 2: Average simulated peak hour volumes

Figure 3: COV of simulated peak hour volumes

Figure 4: Mean intersection delay

Figure 5: Standard deviation of intersection delay

Figure 6: Relative error of estimated intersection delay

Figure 7: Fraction of days exceeding intersection delay threshold

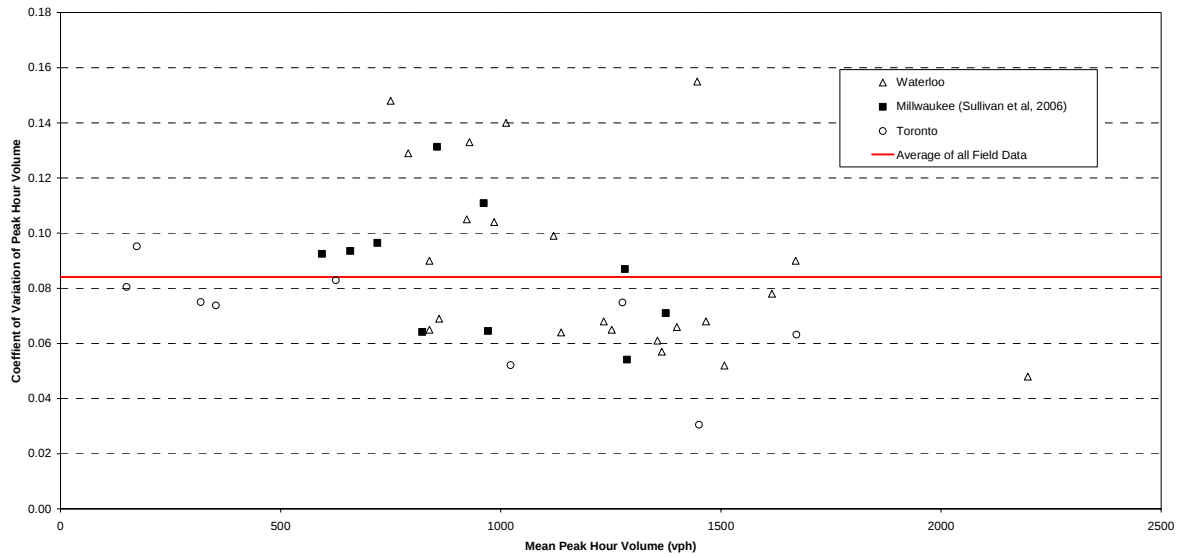


Figure 1: COV of peak hour volumes observed in the field

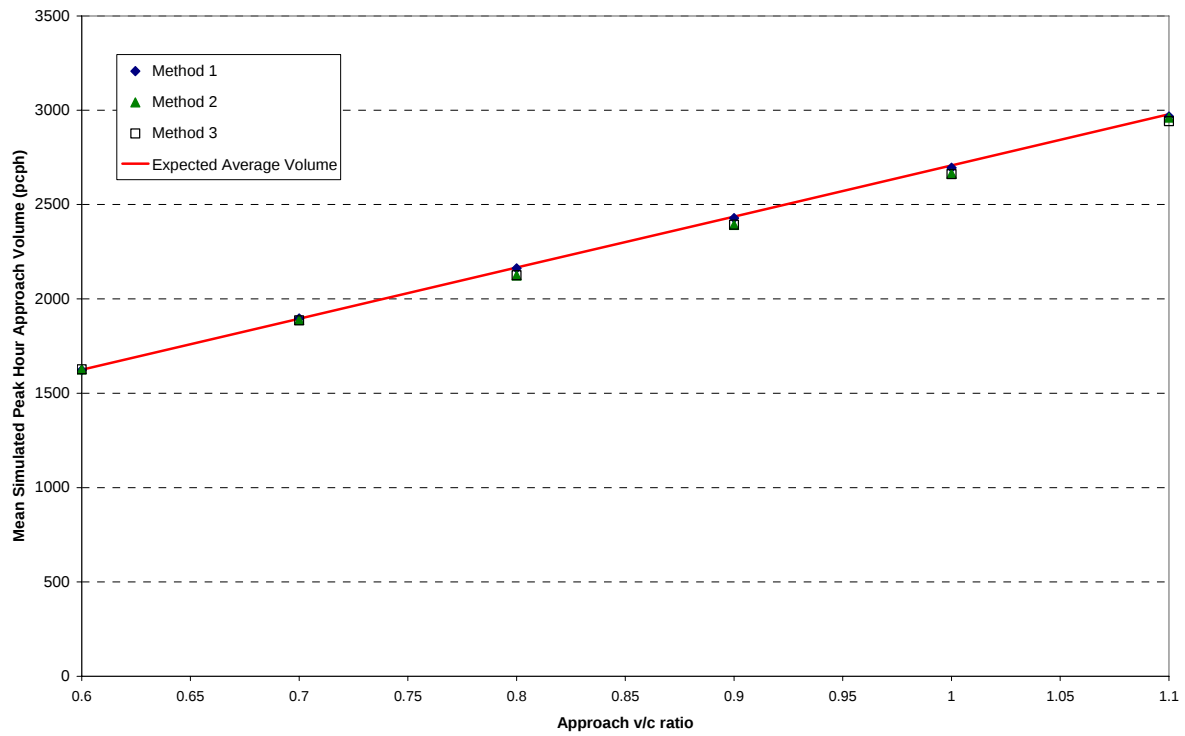


Figure 2: Average simulated peak hour volumes

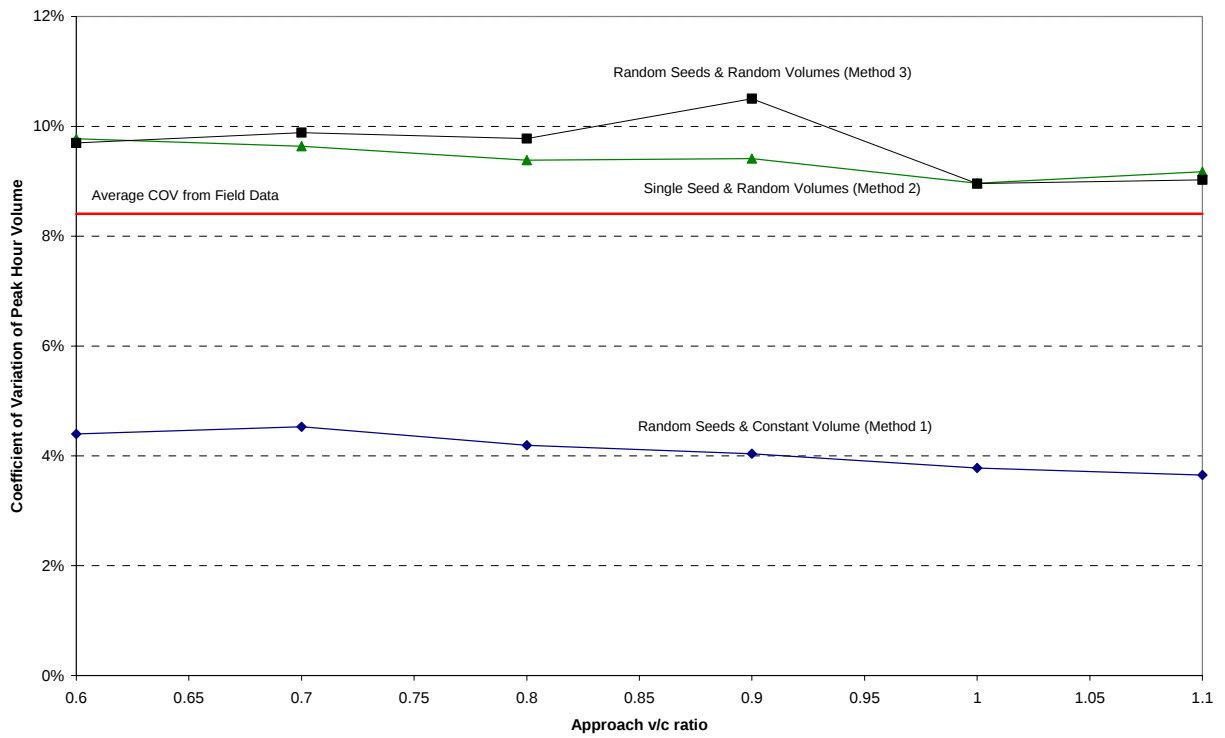


Figure 3: COV of simulated peak hour volumes

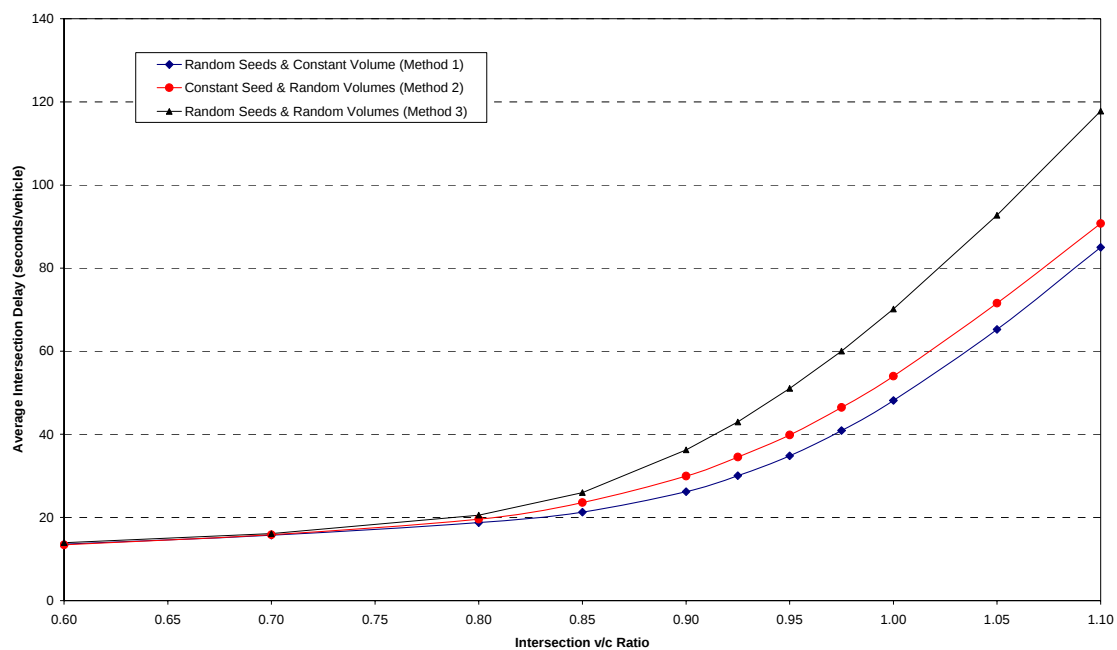


Figure 4: Mean intersection delay

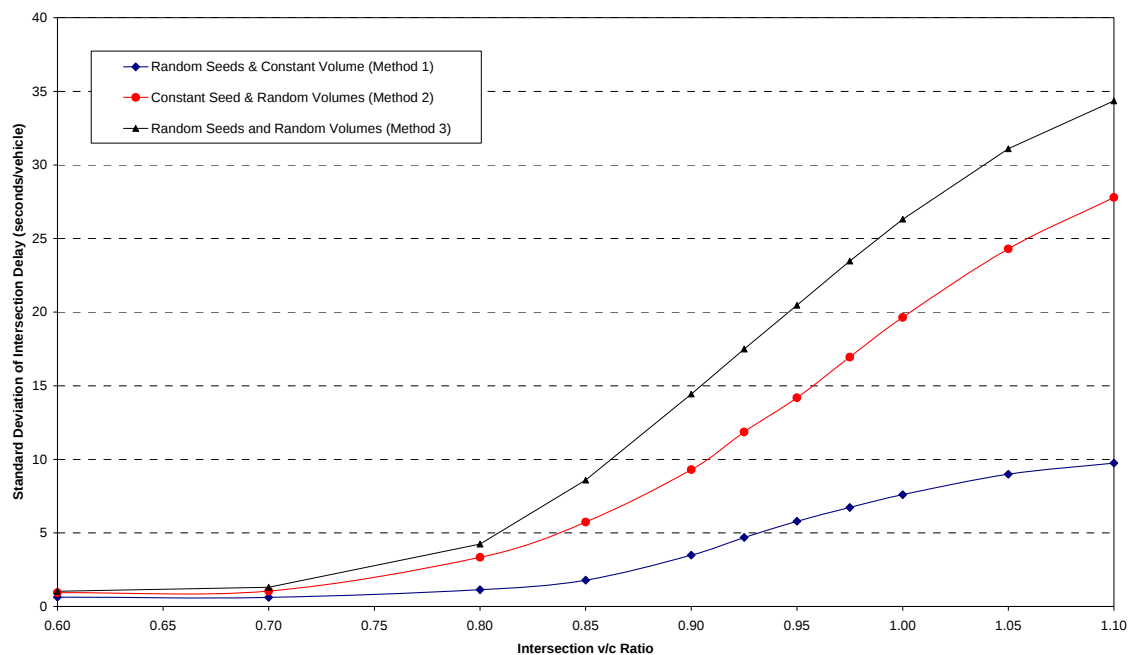


Figure 5: Standard deviation of intersection delay

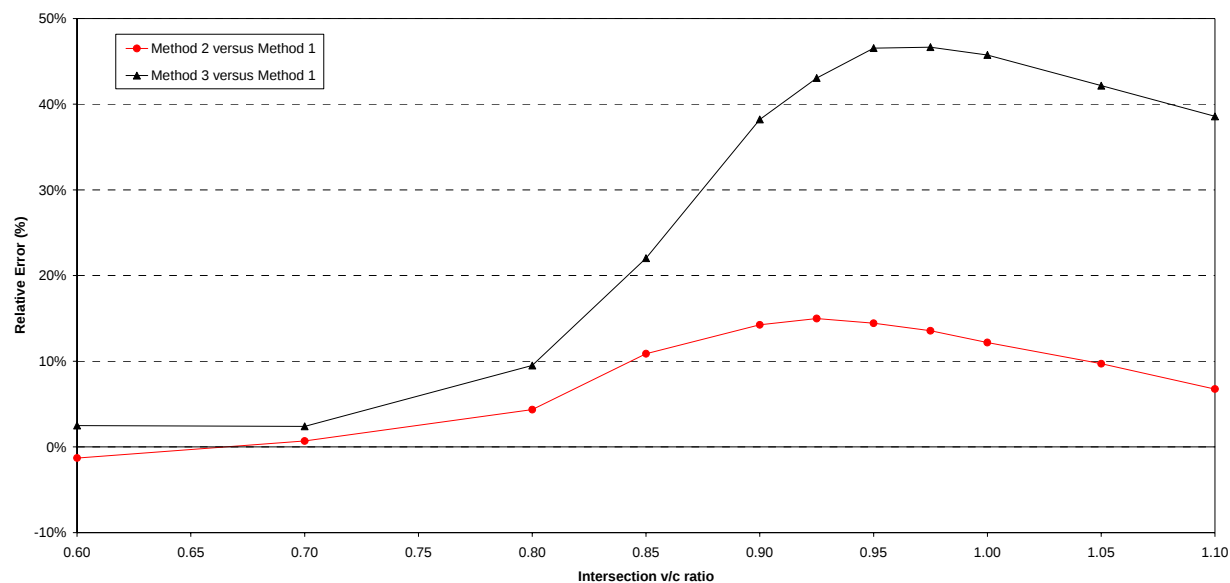


Figure 6: Relative error of estimated intersection delay

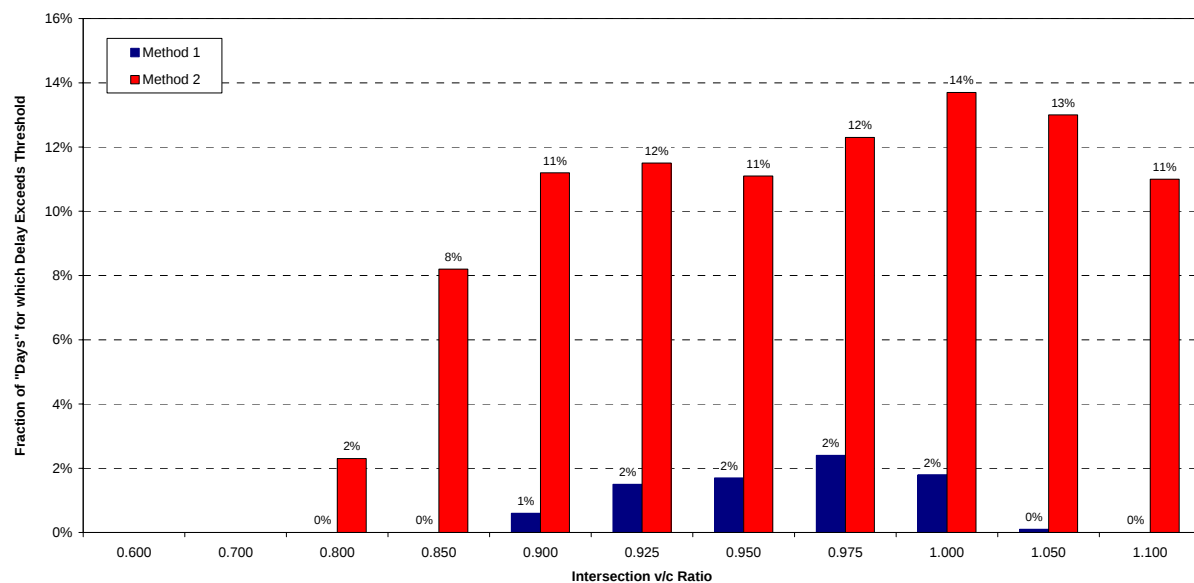


Figure 7: Fraction of days exceeding intersection delay threshold