DELAY VARIABILITY AT SIGNALIZED INTERSECTIONS

- Draft for Review -

Liping Fu

Assistant Professor, Department of Civil Engineering

University of Waterloo, Waterloo, ON

Canada, N2L 3G1

Phone: 519-885-1211 ext. 3984, Fax: 519-888-6197, E-mail: lfu@uwaterloo.ca

and

Bruce Hellinga

Assistant Professor, Department of Civil Engineering

University of Waterloo, Waterloo, ON

Canada, N2L 3G1

Phone: 519-885-1211 ext. 2630, Fax: 519-888-6197, E-mail: bhellinga@uwaterloo.ca

Submitted for

Presentation at the TRB Annual Meeting and Publication in the Transportation Research Record

DELAY VARIABILITY AT SIGNALIZED INTERSECTIONS

Liping Fu and Bruce Hellinga
Assistant Professor, Department of Civil Engineering
University of Waterloo, Waterloo, ON
Canada, N2L 3G1

Abstract

Delays that individual vehicles may experience at a signalized intersection are usually subject to large variation due to randomness of traffic arrivals and interruption caused by traffic signal control. While such variation may have important implications for the planning, design and analysis of signal controls, currently there is no analytical model available to quantify it. This paper describes the development of an analytical model for predicting the variance of overall delay. The model is constructed on the basis of the delay evolution patterns under two extreme traffic conditions: highly undersaturated and highly oversaturated conditions. A discrete cycle-by-cycle simulation model is used to generate data for calibrating and validating the proposed model. The practical implications of the model are demonstrated through its use in analyzing factors influencing delay variability, in determining optimal cycle times with respect to delay variability, and in assessing level of service according to the percentiles of overall delay.

INTRODUCTION

The ability to accurately quantify vehicle delays at signalized intersections is a critical component for the planning, design and analysis of signal controls. As a result of random fluctuations in traffic flow and interruptions caused by traffic controls, delays that individual vehicles experience at a signalized approach are often subject to highly stochastic and time-dependant variation. It has been increasingly recognized that the estimate of the variability of delays is also of importance for many applications (Olszewski 1995; Rouphail, 1995; Fu and Teply, 1996). For example, having knowledge of the variability of delays makes it possible to estimate the confidence limits about the mean delays and thus provide a more informative comparison of alternative signal plans in identifying optimal signal settings. By considering the variability of delay, more reliable signal control strategies may be generated resulting in improved Level of Service (LOS) of signalized intersections.

The problems of estimating delays at signalized intersections have been extensively studied in the literature, however the vast majority of the work has focused on developing models for estimating the mean delay - a point estimate of stochastic delays. Detailed discussions of these average delay prediction models have been provided by Allsop (1972), Newell (1982) and Hurdle (1984). However, much less work has been done to quantify the variability of delay at a signalized approach. Teply and Evans (1989) analyzed the delay distribution at a signalized approach for evaluating signal progression quality. They observed that most of the delay distributions are bimodal and a point estimator is not adequate to describe these distributions. By considering the cyclic overflow delay over time as a Markov chain, Kimber and Hollis (1979), Cronje (1983) and recently Olszewski (1993, 1994) developed numerical methods to calculate the average delay and time-dependant distribution of average cyclic delay. This type of model, while capable of completely specifying the delay distribution, requires substantial computational resources for calculating and storing state and transition probabilities and therefore is not well suited for use in practical situations. The objective of this paper is to develop an analytical model for estimating the variability of delays at signalized intersections with specific focus on predicting the variance of delays of vehicles traversing a signalized approach during a given time interval.

The paper presents an approximate model for predicting the variance of delays. Section 2 outlines the methodologies applied to develop the approximate model. Section 3 presents the development of the approximate model. Section 4 describes the discrete cycle-by-cycle simulation model that was developed for calibrating and validating the proposed model. This simulation model is used in Section 5 to generate data for calibrating and validating the proposed models under a variety of signal operating conditions. Applications of the developed model are demonstrated in Section 6 through its use in a sensitivity analysis and in determining reliability-oriented optimal cycle times and levels of service. Finally, Section 7 presents conclusions and recommendations.

ASSUMPTIONS AND NOTATION

The delay that a particular vehicle experiences when it travels through a signalized intersection approach depends on a number of factors including the probabilistic distribution of arrival flow, signal timings and the time when the vehicle arrives at the approach. In a real application environment, many of these factors are random variables, which makes accurate estimation of this delay a very complicated process. As an initial research effort, this paper considers the following idealized road traffic and signal control conditions:

i) The intersection approach consists of a single through lane controlled by a fixed-time signal. The approach has unlimited space for queuing and has a constant saturation flow rate;

ii) The vehicle arrival at the approach is a random variable with a known probabilistic distribution. The rate of vehicle arrivals during the evaluation time is assumed to be constant. No initial queue is present at the beginning of the evaluation time. The flow rate increases abruptly from zero to the rate for the evaluation time. The traffic stream consists only of passenger car units (pcu);

Consider the cumulative arriving and departing of vehicles during the time interval [0,T] at the stopline of a signalized approach as illustrated in Figure 1. The delay for a particular vehicle arriving at time *t*, called *overall delay* and noted as *D*, is considered to include two components: uniform delay and overflow delay, as shown in Equation 1:

$$D = D_1 + D_2 \tag{1}$$

where the uniform delay component, D_1, is defined as the portion of delay that would be incurred by a vehicle when the approach is undersaturated and all vehicle arrives uniformly. The overflow delay component, D_2, represents the portion of delay that is caused by temporary overflow queues resulting from the random nature of arrivals and by continuous overflow when the arrival rate during the time period [0, T] exceeds the capacity.

The estimation of the overflow delay component in Equation (1) is complicated as a result of the complex time-dependent stochastic nature of the queuing process, and currently there is no theory available for use to develop a single analytical model suitable across all saturation levels. Past research has mainly focused on developing approximate models for estimating the average overall delay using simulation data as a mechanism to obtain data for calibration (Webster, 1958; Kimber and Hollis, 1979; Akcelik, 1981; Teply et al., 1995; Rouphail and Akcelik, 1990; Brilon and Wu, 1990). A number of similar delay models are available to provide estimates of this measure. For example, the Canadian delay model uses Equation (2) to estimate the average overall delay (Teply et al., 1984 and 1995, note that units of some parameters have been changed for use in this paper):

$$E[D] = k_f \frac{c_y (1-\lambda)^2}{2(1-\lambda x_1)} + 0.25T \left[(x-1) + \sqrt{(x-1)^2 + \frac{4x}{c_a T}} \right] \tag{2}$$

where

$E[D]$ = average overall delay (seconds)

T = evaluation time (seconds)

c_y = cycle time (seconds)

g_e = effective green interval duration (seconds)

λ = g_e/c_y

q = average arrival flow rate from time 0 to time T (pcu/second).

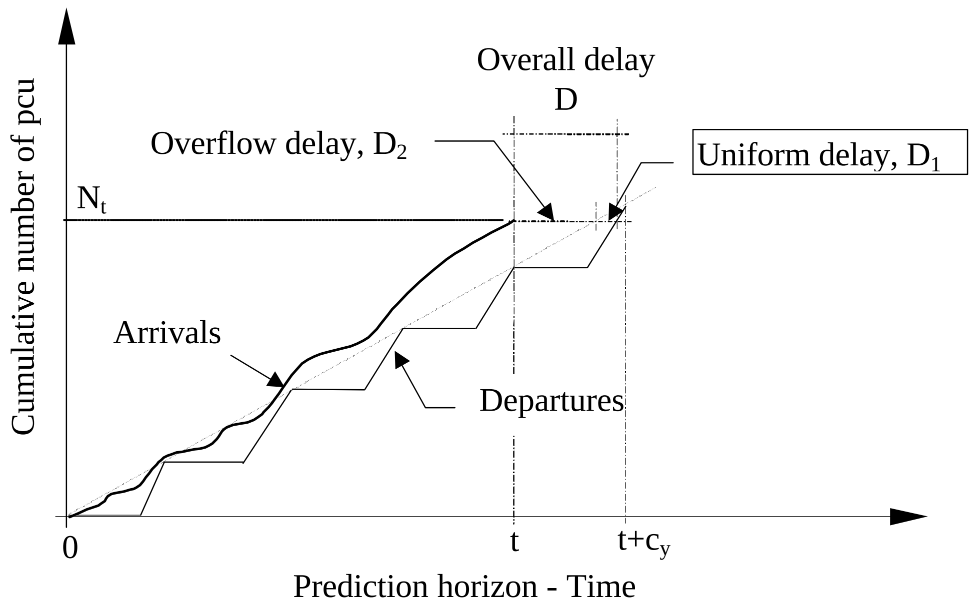

Figure 1. Queuing diagram illustrating the components of delays

k_f = adjustment factor for the effect of the quality of progression, defined as $k_f = (1-p)f_p/(1-l)$. p is the proportion of vehicles arriving during the green interval and f_p is a supplemental adjustment factor for platoon arrival type. Note that this study does not consider the effect of signal progression, i.e. $k_f = 1.0$

c_a = capacity (pcu/second), determined by $s\,l$, where s is the saturation flow rate (pcu/second)

x = degree of saturation, defined as q/c_a

x_1 = minimum of (1.0, x)

This paper considers the development of a model for the variance of overall delay, i.e., Var[D], which is *defined* as the summation of the variance of uniform delay and the variance of random delay (Equation 3):

$$Var[D] = Var[D_1] + Var[D_2] \qquad (3)$$

Where Var[D_1] is the variance of uniform delay defined as the variance of delay that would be experienced by vehicles when all vehicles arrive uniformly at a constant arrival rate; Var[D_2] is the variance of random delay, or the difference between the variance of overall delay and the variance of uniform delay. The variance of uniform delay can be deirved theoretically. The variance of overflow delay is directly calibrated from simulation data with its functional form constructed on the basis of an analysis of the variance models under two traffic extremes: highly undersaturated and highly oversaturated conditions. The following section provides a detailed description of the development of these models.

APPROXIMATE MODEL FOR THE VARIANCE OF OVERALL DELAY

The variance of uniform delay, $Var[D_1]$, represents the variation of uniform delay that would be experienced by vehicles arriving during the time interval $[0,T]$. This variation results from the uncertainty of the vehicle's arrival time during each cycle of the interval. The vehicle can arrive at any moment within a cycle and thus experience variable delays as a result of the signal control. An estimate of this variance component can be obtained theoretically on the basis of a deterministic queuing model with vehicles arriving uniformly during the cycle (Van Aerde et. al., 1993; Rouphail, 1995):

$$Var[D_1] = \frac{c_y^{\,2} \cdot (1-l)^3 \cdot (1+3l-4lx_1)}{12(1-lx_1)^2} \qquad (4)$$

In order to establish a model for the variance of delay caused by overflow queue, two extreme traffic conditions are first investigated: undersaturated conditions ($x<1.0$) and oversaturated conditions ($x>1.0$). For undersaturated conditions, overflow delay experienced by a vehicle arriving during the time interval $[0, T]$ is mainly caused by occasional overflows of traffic from each cycle. The relationship between the variance of this delay and the degree of saturation can be approximated from the well-known Pollaczek-Khintchine formula for an M/G/1 system (for the general formula and derivation, see e.g. Medhi,1991) by supposing that the signal is acting as a server with a constant service time $1/c_a$, as shown Equation 5.

$$Var[D_2] \approx \frac{x \cdot (4-x)}{12c_a^{\,2} \cdot (1-x)^2} \qquad (5)$$

It should be emphasized that the above model is merely an approximate estimate of the variance because a steady-state may not be reachable during time interval [*0,T*]. Nevertheless, the equation can be used to illustrate the qualitative relationship between the variance of delay and the degree of saturation. With this assumption, the variance is time-independent and an infinite variance would be predicted as the degree of saturation (x) approaches unity. In reality, at high degrees of saturation, the system is not likely to settle into a steady-state by time T. Consequently, it can be expected that Equation (5) provides a reasonable approximation of the variance only under light traffic conditions ($x<<1.0$).

If the intersection approach is highly oversaturated during the time period [*0, T*], there is a high probability that an overflow queue always exists during the period from time *0* to time T. Consider a vehicle arriving at time t during the time period [*0,T*]. The overflow queue for a vehicle arriving at time t, Q_t, can be determined as the total arrivals minus the total departures (Equation 6).

$$Q_t = N_t - c_a \cdot t \tag{6}$$

The number of arrivals, N_t, is a random variable with a mean equal to qt. The delay experienced by the vehicle can then be simply determined on the basis of the overflow queue as expressed in Equation 7.

$$D_2 = \frac{N_t - c_a \cdot t}{c_a} \tag{7}$$

Based on Equation (7), the variance of delay for vehicles arriving during time interval [*0,T*] can be obtained by assuming the arrival time t is a random variable with known distribution:

$$\begin{aligned} Var[D_2] &= \frac{E[Var[N_t - c_a t \mid t]] + Var[E[N_t - c_a t \mid t]]}{c_a^{\,2}} \\ &= \frac{E[Var[N_t \mid t] - 0] + Var[E[N_t \mid t] - c_a t]}{c_a^{\,2}} \\ &= \frac{E[Var[N_t \mid t]] + Var[qt - c_a t]}{c_a^{\,2}} \\ &= \frac{E[Var[N_t \mid t]] + (q - c_a)^2 Var[t]}{c_a^{\,2}} \end{aligned} \tag{8}$$

If the ratio of the variance-to-mean of the vehicle arrivals, denoted as I_a, is assumed to be constant during the time interval [$0,T$] and given, then

$$Var[N_t|t] = I_a E[N_t|t] = I_a q_t t \tag{9}$$

Note that if the vehicle arrivals follow a Poisson distribution, I_a is equal to 1. In this study, Poisson arrival is assumed but the parameter I_a is still used for the convenience of future extension. With Equation (9), Equation (8) can be further expressed as:

$$Var[D_2] = \frac{I_a q E[t] + (q - c_a)^2 Var[t]}{{c_a}^2} \tag{10}$$

If we assume the arrival time is uniformly distributed during the time interval *[0,T]*, Equation (10) can be further expressed as:

$$Var[D_2] = \frac{I_a T x}{2c_a} + \frac{T^2(1-x)^2}{12} \tag{11}$$

It must be emphasized that Equation (11) is valid only when there is an overflow queue present during the period from time 0 to time t. In reality, however, it is possible that no overflow queue exists at time t and consequently no overflow delay is experienced. Consequently, it can be concluded that Equation (11) represents an upper bound estimate of the variance of overflow delay. The actual variance would be lower than that predicted by Equation (11), but the prediction error should become smaller as the degree of saturation increases, and the associated likelihood of overflow queuing increases.

Figure 2 depicts the relationships between the variances of overflow delay as functions of the degree of saturation represented by Equation (5) and (11). Both curves are only appropriate within certain flow domains: either highly undersaturated or highly oversaturated traffic conditions. Consequently, it is hypothesized that the true relationship between the variance and the degree of saturation follows the dashed curve in Figure 2. It can be observed that it is difficult, if not impossible, to derive the functional relationship for the transitional curve directly from Equation (5) and (11) through the traditional coordinate transformation technique. Therefore, the non-linear function, expressed in Equation (12), is proposed to model the variance:

$$Var[D_2(t)] = \left\{ \frac{I_a T x}{2c_a} + \frac{T^2 (1-x_1)^2}{12} \right\} e^{-\left(\frac{x_0}{x}\right)^b} \tag{12}$$

Where $x_1 = max\{1, x\}$

The parameters x_0 and b determine the shape of the delay curve and their values need to be calibrated. It can be observed that the proposed function has two desired attributes. First, the function is asymptotic to the model for oversaturated conditions (Equation 11). Second, similar to the undersaturated model (Equation 5), the function goes to zero as x approaches zero. However, while these characteristics are necessary, they do not of themselves demonstrate that the proposed function is realistic. Therefore, data from a simulation model were used to calibrate appropriate values for x_0 and b and to validate the calibrated model, as discussed in Section 4.

Having developed expressions for the variances of uniform delay and overflow delay, the variance associated with the overall delay (Equation 7), can be expressed by Equation (13).

$$Var[D(t)] = \frac{c_y^{\ 2} \cdot (1-l)^3 \cdot (1+3l-4lx_1)}{12(1-lx_1)} + \left\{ \frac{I_a T x}{2c_a} + \frac{T^2 (1-x_1)^2}{12} \right\} e^{-\left(\frac{x_0}{x}\right)^b} \tag{13}$$

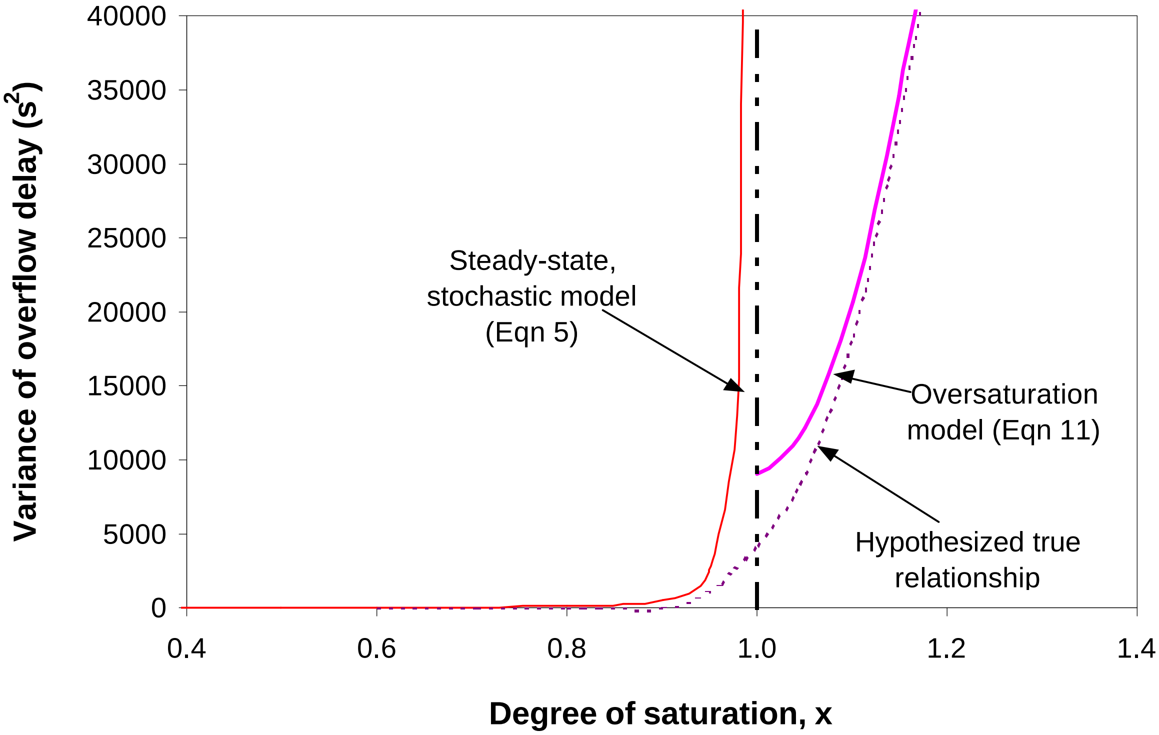

Figure 2. Models for the variance of overflow delay

SIMULATION MODEL

In order to obtain data to calibrate and validate the proposed models, a discrete cycle-by-cycle simulation system was developed. The following sections briefly describe the design and verification of the simulation model.

Logic of the Simulation Model

The simulation model explicitly models the delay that a vehicle experiences when traversing a signalized intersection approach. The approach is used exclusively for through traffic and controlled by a pre-timed traffic signal. The vehicle arrivals are randomly distributed with the vehicle headway following a negative exponential distribution with a minimum headway equal to one second.

The vehicle discharge pattern during the green interval depends on the queue status at the approach. If there is no queue present when a vehicle arrives, then the vehicle can immediately be discharged without any delay. Otherwise, the vehicle must wait until the queued vehicles ahead of it discharge. Vehicle discharge headway is determined based on saturation flow rate.

The simulation starts with no queue present and reset the queue size to zero whenever the elapsed clock time reaches a pre-specified evaluation time. The simulation terminates once the required total number of cycles has been simulated. The arrival time and delay associated with each vehicle are recorded for use in the analysis stage. Information such as the mean and variance of delays experienced by vehicles arriving during the evaluation time can then be derived.

Verification of the Simulation Model

Before the simulation model was used to generate data for calibrating and testing the proposed models, it was verified against results from other available models. Two comparisons were made. First, the average overall delays obtained from the simulation model for a given evaluation period under different saturation ratios were compared to the results from the Australian (Akcelik, 1981), Canadian (Teply et al., 1995), HCM (TRB, 1994) and Markov chain models (Olszewski, 1994). For convenience, the scenario used in this comparison is the same as that used by Olszewski (1994) for a similar purpose. The evaluation period duration is 15 minutes. The signal timing consists of a cycle time of 60 seconds, an effective green interval of 24 seconds and a saturation flow of 1800 pcu/hr. A total of 6000 cycles, corresponding to 100 hours of traffic flow, was simulated for each degree of saturation. It was estimated that this number of simulations would result in an estimation error of less than 0.5 seconds at a significance level of 95%.

Figure 3 illustrates the average overall delay obtained from the simulation model and the four other methods. It should be noted that the overall delays associated with the HCM model have been obtained by multiplying the stopped delays from the HCM formula by 1.3 to convert stopped delay to overall delay. The Markov chain model assumes Poisson arrivals and constant departure during the green interval. As it would be expected, the simulation results are almost identical to the Markov chain model. Among the three other models, the Australian model shows the best agreement with the simulation model under all levels of saturation and the Canadian model provides the best agreement with the simulation for oversaturated conditions. It should be noted that the differences between the HCM, Canadian and Australian delay equations are expected and have been addressed by Akcelik (1988).

The objective of the second comparison is to provide an indication of the validity of the simulation model in estimating the variance of delays. The simulation results are compared to those reported by Olszewski (1994) in which the exact means and variances of delays under various levels of saturation were obtained for a given case from a Markov chain model. The system parameters are the same as for the previous comparison except the evaluation time is 30 minutes, instead of 15 minutes. In this comparison, the number of cycles to be simulated was estimated on the basis of an analysis of the confidence interval for the variance. It was estimated that a total of 6000 cycles for each degree of saturation would yield an estimation error for the standard deviation of less than two seconds at a significance level of 95%. Figure 4 shows that the standard deviations of delay estimated by the simulation model and provided by Olszewski (1994) from the Markov chain model. It can be observed that the estimates of the standard deviation of delay from the simulation model are quite consistent with those obtained from the Markov chain model. The overestimation of the standard deviation of delay by the simulation model, especially in the range $x<1.0$, is expected because the Markov chain model does not consider the variation of travel time within the cycle as quantified by Equation 4.

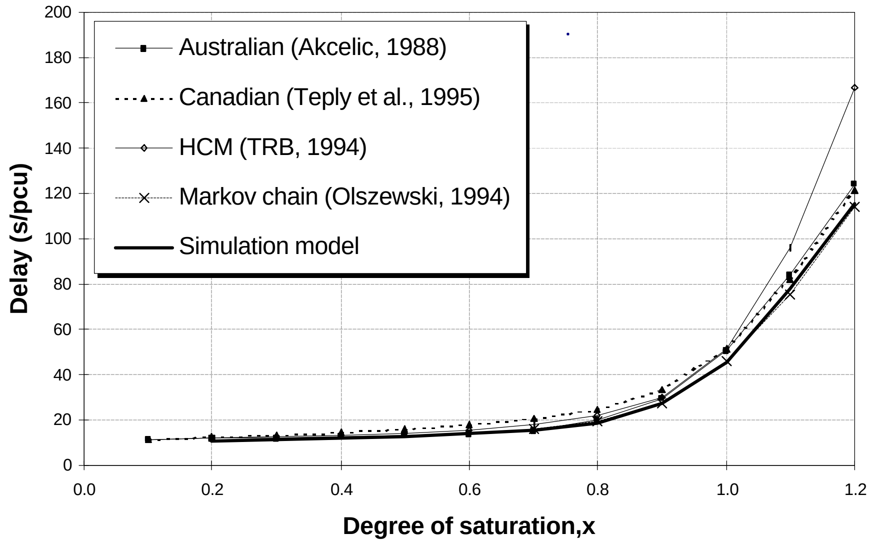

Figure 3. Average overall delay estimated by the Australian, Canadian, HCM, Markov chain and simulation models (*c_y = 60s, g_e = 24s, s = 1800 pcu/h and t = t_e = 15 min; simulated cycles = 6 000*)

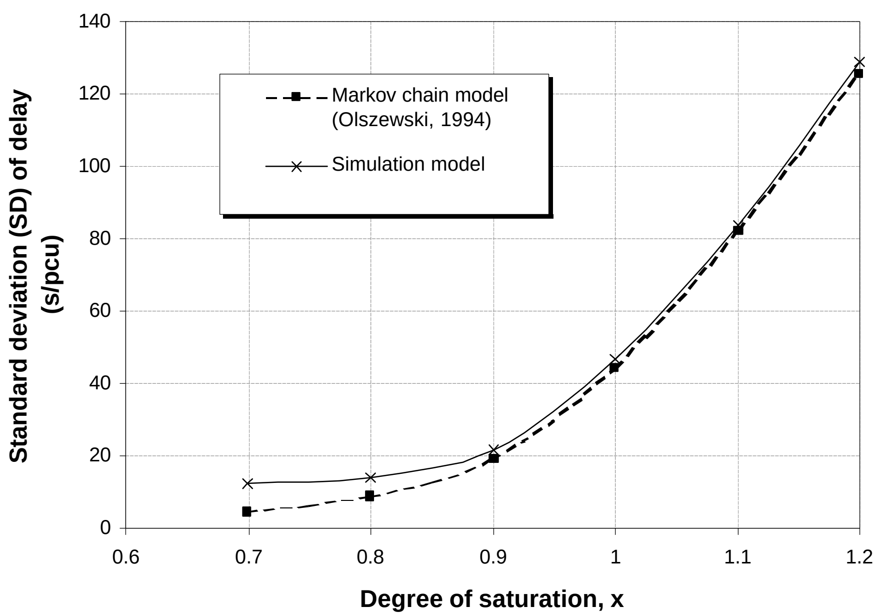

Figure 4. Standard deviation of delay estimated by the Markov chain model and simulation model (*c_y = 60s, g_e = 24s, s = 1800 pcu/h and t = t_e = 30 min; simulated cycles = 6 000)*

MODEL CALIBRATION AND VERIFICATION

Model Calibration

To determine the appropriate parameter values for the overflow delay variance model shown in Equation (3), a two-step sequential calibration procedure is performed. The first step is to find the x_0 and b values that would produce the best fit between the estimates of the variance of the overflow delay from Equation (12) and the estimates from the simulation model (representing the true values) for a given cycle time (c), effective green interval (g_e) and evaluation time (T). Following the definition of Equation 3, the variance of random delay from simulation is obtained as the difference between the variance of overall delay which is calculated from the delay of simulated vehicles and the variance of uniform delay from Equation 4.

Due to the non-linear relationship between the variables, a non-linear regression process is conducted by first transforming Equation (12) into an equivalent linear equation (Equation 14).

$$Y = a + b\,X \qquad (14)$$

Where:

$$Y = ln\left\{ ln\left(\{ \frac{I_a qx}{c_a} + d\,\frac{T^2(1-x)^2}{12} \} \right) - ln\left(Var\left[D_2\right]\right) \right\}$$

$X = \ln(x)$

$a = b\ln(x_0)$

$b = -\,b$

The simulation model is used to obtain the values of the variance of overflow delay ($Var[D_2]$), which is the difference between the variance of overall delay and the variance of uniform delay calculated based on Equation (4), under various combination of c_a, x and T. These data were transformed to X and Y values as in Equation (14). For a set of pre-specified x values, linear regressions were performed to determine the values of a and b, which were subsequently transformed back to values for x_0 and b. The data points used in regression were determined by simulation by fixing the values of x, g_e and T and varying the degree of saturation x from 0.8 to 1.2 with an increment of 0.05. Each data point results from a simulation of 15000 cycles. The regressed x_0 and b, together with (c_y, g_e, T), form a new data point (c_y, g_e, t, x_0, b). By changing the values of the parameter set (c_y, g_e, T) and repeating the regression analysis, a number of such data points can be obtained. In this study, a total of 18 points were generated with the following combinations of parameters: $c_y = \{70, 90, 120\}$; $l = \frac{g_e}{c_y} = \{0.2, 0.5, 0.8)$; $T = \{900, 3600\}$. It was found that the linear relationship shown in Equation (14) is statistically significant for each of the 18 combinations with a minimum R^2 of 0.95, which indicates that the proposed functional form is appropriate.

In the second step, a series of correlation analysis of the relationships between the parameters (x_0, b) and (c_y, g_e, T, l, T/c_a) were conducted and the following best fit equations were obtained:

$$x_0 = 0.947 + 1.330*10^{-6}\, T/c_a + 0.157\, l \qquad (15)$$

$$(R^2 = 0.87, t_1 = 13.07, t_2 = 3.84)$$

$$b = 8.294 + 6.080*10^{-4}\, T/c_a \qquad (16)$$

$$(R^2 = 0.93, t_1 = 13.07)$$

The obtained high R^2 values indicate that both equations explain a large portion of the variation in the simulated data. All t-values are greater than the critical t-value at the 5% level of significance, which indicates that the included parameters are statistically significant.

Model Evaluation

The simulation system is first used to estimate the variance of overall delay corresponding to various evaluation times and traffic conditions. A total of 210 combinations were simulated with the following combinations of parameters: $c_y = \{50, 60, 80, 100, 120\}$, $l = \{0.3, 0.5, 0.7)$, $t = \{900, 3600\}$ *and* $x = \{0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2\}$.

Figure 5 shows the correlation between the standard deviation of the delay obtained by the analytical model and the simulation results. Each point represents the result of simulation runs of 15,000 cycles. The approximate model exhibits no apparent bias and has a high correlation with the simulated estimates ($R^2 = 99.3\%$).

The calibrated model is further evaluated using results from Olszewski (1994) in which the exact variances of overflow delays under various levels of saturation were obtained for a given case from a Markov chain model. The values of the system parameters for the case as well as the results are shown in Figure 6. It can be observed that the estimates of the standard deviation of delay from the simulation model are very consistent with those obtained from the Markov chain model.

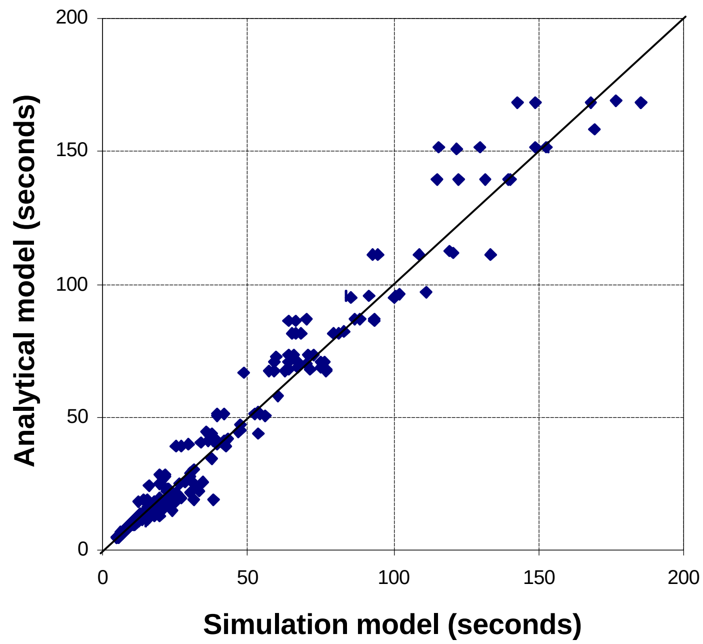

Figure 5. Correlation of the standard deviations of overall delay estimated by the analytical model with simulation results *(s = 1800 pcu/h; simulated cycles = 15000/combination*)

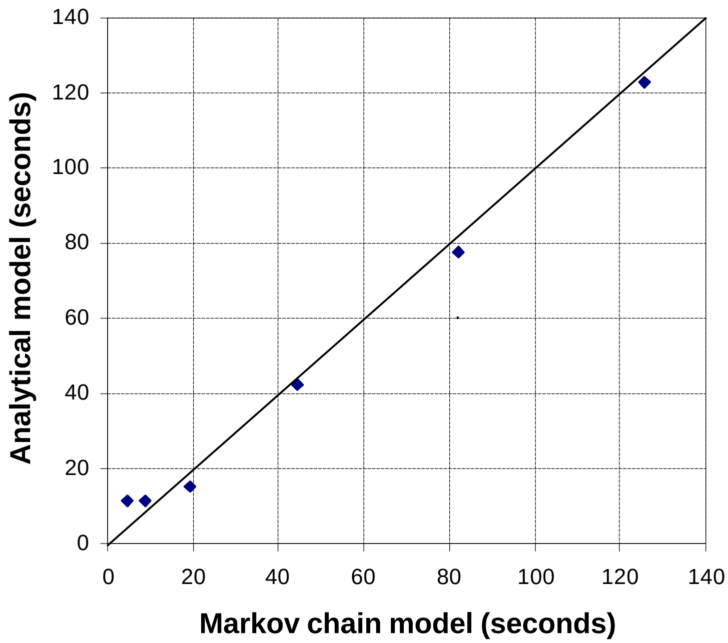

Figure 6. Standard deviation of overflow delay estimated by the Markov chain model (Olszewski, 1994) and simulation model (*c_y = 60s, g_e = 24s, s = 1800 pcu/h and T = 30 min*)

APPLICATIONS OF THE VARIANCE MODEL

Factors Influencing Delay Variability

The objective of this section is to to analyze the effects of various factors on the delay variability. The delay variability is represented by the standard deviation of overall delay. The influencing factors considered in this analysis were limited to the degree of saturation reflecting the level of traffic congestion (x), the green to cycle ratio representing the traffic signal setting (l), the evaluation time (T) and saturation flow (s).

Figure 7 shows the standard deviations of delay as a function of the degree of saturation under different values of the influencing factors. The following observations can be made:

1. In general, the more congested the traffic is, the larger the variance of delay becomes. However, it is interesting to note that the standard deviation of delay is almost

constant for the undersaturated traffic conditions ($x < 0.8$) with a value largely depending on the variance of uniform delay (Equation 4). This implies that the variance of uniform delay is insensitive to the level of traffic congestion. This also indicates that if only uniform delay is considered, as in Rouphail (1993), the variance would be significantly underestimated for saturated traffic conditions.

2. The evaluation time has a significant impact on the overflow component of the variance of overall delay, as shown in Figure 7-(a). The longer the evaluation time is, the larger the delay variability is, similar to the average overflow delay.
3. As shown in Figure 7-(b), the saturation flow rate makes different only under saturated traffic conditions (x>0.9) and the impact is quite consistent across the degree of saturation.
4. From Figure 7-(c), it can also be observed that the green proportion allocated to the approach has an important impact on both the variance of overflow delay and the variance of uniform delay. The smaller this proportion is, the larger the delay variability becomes.

Optimal Cycle Time

Average overall delay has been traditionally used as one criterion in determining optimal cycle times. This section examines if there is an optimal cycle time that minimizes the variability of delay that individual vehicles experience at a signal controlled intersection. An idealized two-phase, four approach intersection with equal flows on all approaches is considered. The saturation flow rate is 1800 pcu/hour and the lost time at each phase is 4 seconds. Figure 8 shows the relationship between the variance of overall delay and the average overall delay as functions of cycle time. It can be observed that the average overall delay and the variance of overall delay have similar trends with respect to the cycle length. Furthermore, the range of optimal cycle times with respect to average overall delay (q=600 pcy/h: optimal c_y = 40~60 s and q=800 pcu/h: optimal c_y = 80~100 s) overlaps with those determined on the basis of minimizing the variance of overall delay (q=800 pcy/h: optimal c_y = ~40 s and q=800 pcu/h: optimal c_y = 90~120 s). This indicates that, for the scenarios examined, the current practice of determining optimal cycle lengths on the basis of minimizing average overall delay is appropriate with respect to the objective of minimizing the variance in overall delays.

Variability of Level of Service

This section illustrates the possible use of delay variability in quantifying level of service for signalized intersections. In the current HCM (1997), level of service for signalized intersections is defined in terms of average overall stopped delay. With the ability to estimate the variance of overall delay, it is feasible to integrate the concept of reliability in design and analysis of signalized intersections. For example, delay at a certain percentile, instead of average value, can be used to define the level of service. A 95th percentile delay means that 95% of vehicles would experience delay less than or equal to this delay. The percentile value can be approximately estimated using $E[D] + z_a \sqrt{Var[D]}$ where z_a is a statistic for the normal distribution and can be determined based on the pre-specified reliability level. Figure 9 shows average overall delay and 90th percentile delay (with $z_a \approx 1.3$) under different degrees of saturation. It assumes that ranges of delay values used in defining each LOS in the HCM are also applicable to individual vehicles, as shown in Figure 9. It can be observed that, for the given case with a degree of saturation of 0.9, the average overall delay is 20 seconds, which would yield a LOS of C (point a). However, if the 90th percentile delay is used, the level of service would be D (point b). On the other hand, in order to guarantee 90% of vehicles going though the intersection approach experience LOS C or higher, the degree of saturation needs to be reduced to 0.8 (point c) by either increasing the capacity or decreasing the demand.

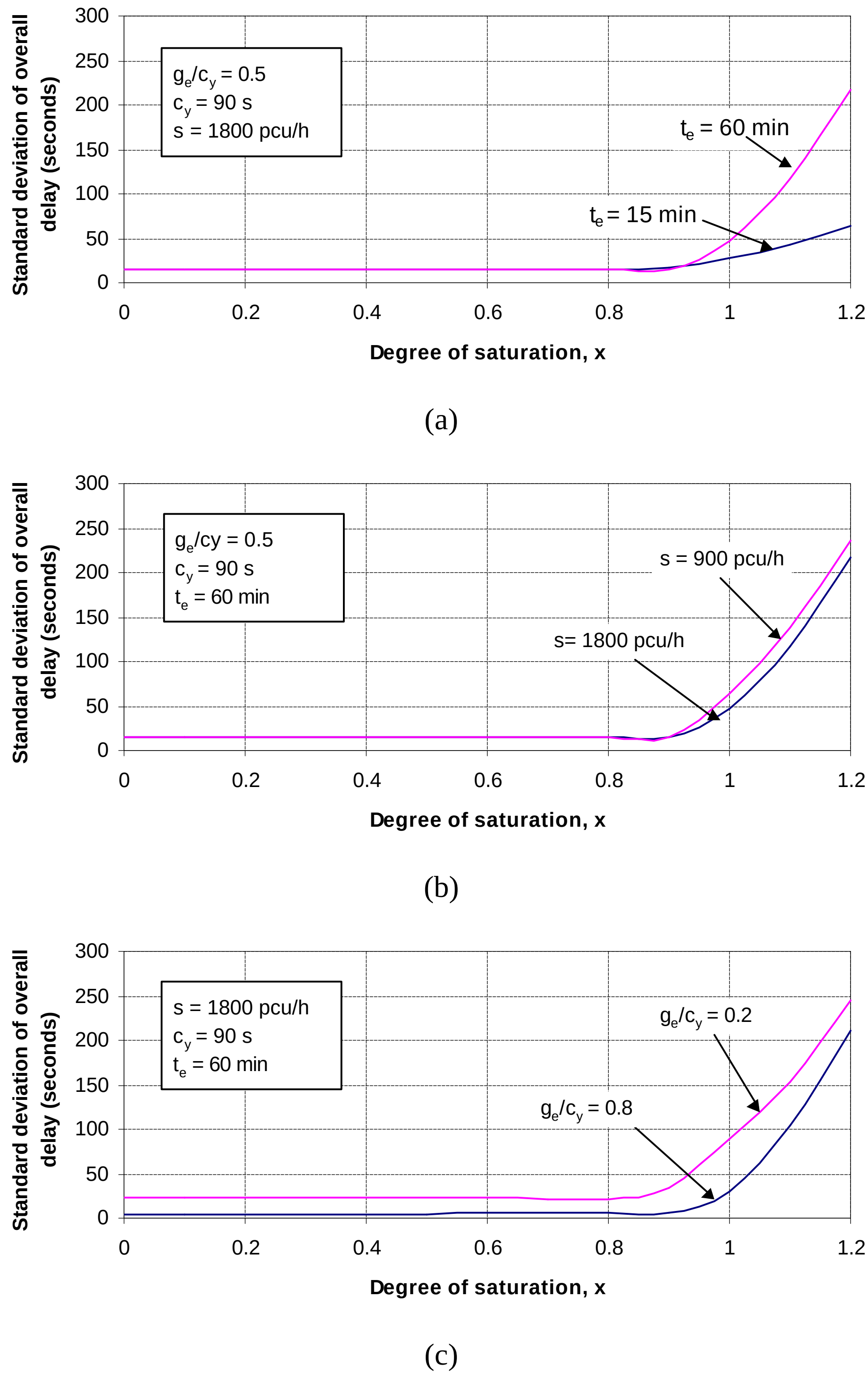

Figure 7. Influence on the standard deviation of overall delay: a) evaluation time, b) saturation flow, and c) ratio of effective green to cycle time

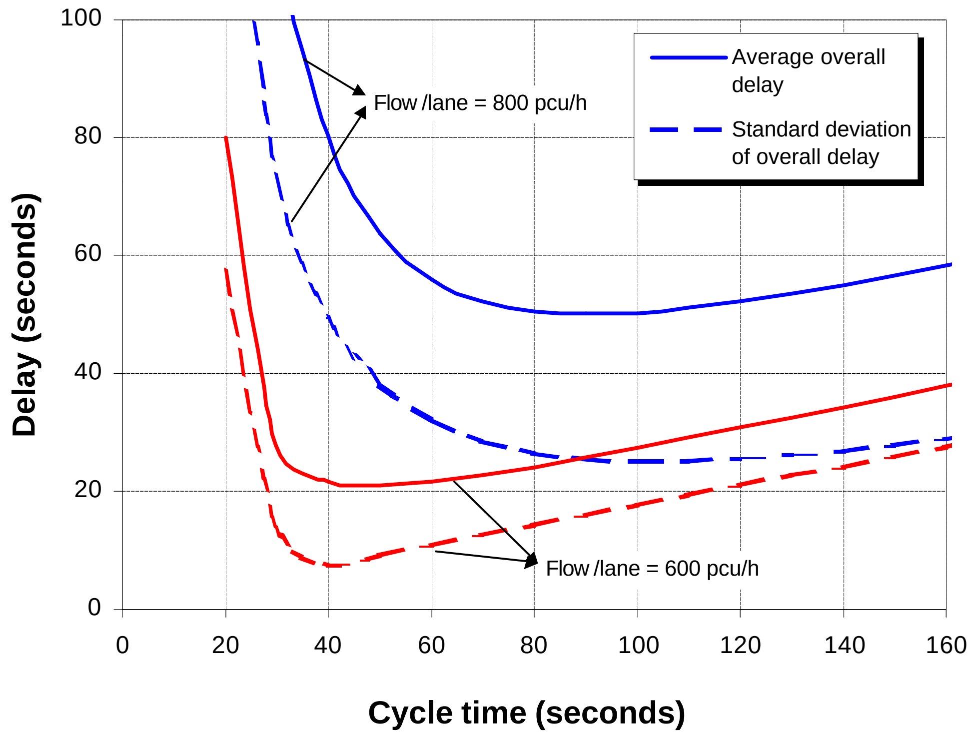

Figure 8. Relationship between the optimal cycle times with respect to mean and standard deviation of overall delay (Two phase, four approach intersection with equal flows on all approaches; saturation flow=1800 pcu/h; lost time = 4 seconds /phase)

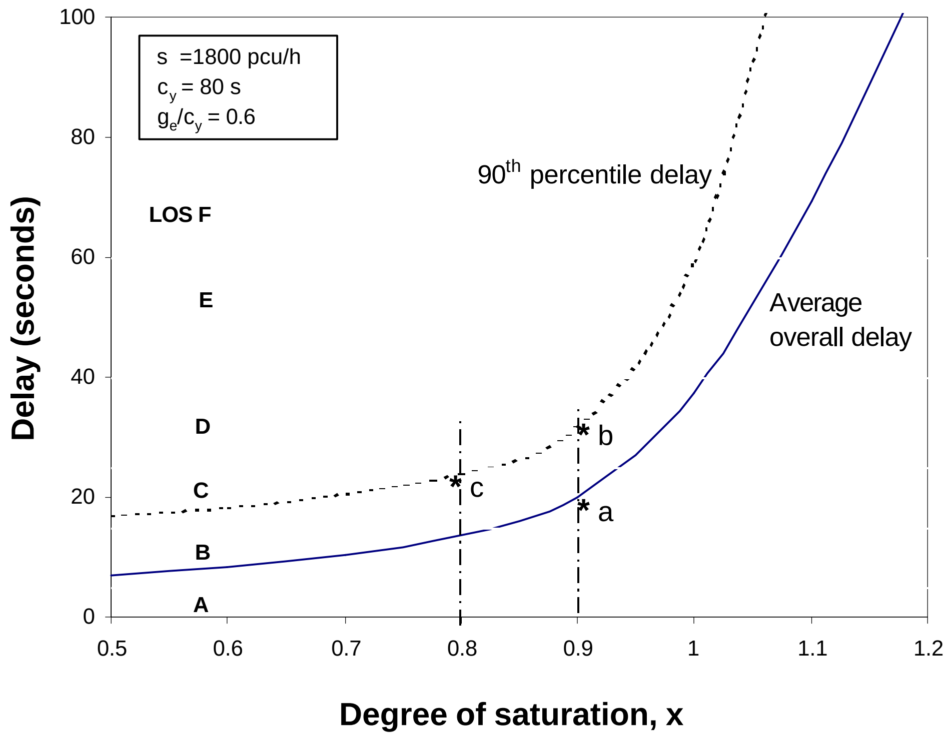

Figure 9. Level of service and delay variability

CONCLUSIONS AND FUTURE RESEARCH

This paper has described the development of an analytical model for estimating the variance of delay at signal controlled approaches. The model was constructed on the basis of the delay evolution patterns under two extreme traffic conditions: highly undersaturated and highly oversaturated conditions. A discrete cycle-by-cycle simulation model was developed and used to generate data for calibrating and validating the proposed models. The results of a correlation analysis indicate a remarkable agreement between the model estimates of the standard deviation of delay and simulation results (R^2 = 99.1%).

The developed model provides a valuable tool for the planning, design and analysis of signal controls. Practical applications have been demonstrated through its use in analyzing factors influencing delay variability, in determining optimal cycle times with respect to delay variability, and in assessing level of service according to the percentiles of overall delay.

The proposed analytical models were calibrated and validated with simulation results that are based on several important assumptions including random traffic arrivals with constant flow rate and unlimited queuing space. These assumptions may be overly restrictive and are likely to be violated in practice. The impact of these assumptions on the validity of these models has not yet been determined. It is recommended that future research focus on the following aspects. The potential impacts of the assumptions applied in this paper should be quantified. Field data should be used in conjunction with simulation results to calibrate and verify the proposed models.

ACKNOWLEDGEMENTS

This research was supported by the Natural Sciences and Engineering Research Council of Canada.

REFERENCES:

Abours, S. (1986) Estimation of Travel Time from Occupancy on an Urban Network, an Experiment in Paris, France. Second International Conference on Road Traffic Control, Institution of Engineers, London, UK, 1986.

Akcelik, R. (1981) Traffic Signals: Capacity and Timing Analysis, Research Report No. 123, Australian Road and Research Board, Vermount South.

Akcelik, R. (1988) “The Highway Capacity Manual Delay Formula for Signalized Intersections” ITE Journal 58 (3), pp. 23-27.

Allsop, R. E. (1972) “Delay at a Fixed Time Traffic Signal – I: Theoretical Analysis. Transportation Science, Vol. 6, 1972, pp. 260-285.

Boyce, D., Rouphail, N., and A. Kirson, (1993), "Estimation and Measurement of Link Travel Times in the ADVANCE Project", IEEE-IEE Vehicle Navigation & Information Systems Conference, Ottawa - VINS'93.

Brilon, W. And Wu, N. (1990) Delays at Fixed Time Traffic Signals Under Time-Dependent Traffic Condition, Traffic Engineering and Control 31 (12), pp. 623-631.)

Cronje, W.B. (1983), “Derivation of Equations for Queue Length, Stops, and Delay for Fixed-Time Traffic Signals” *Transportation Research Record* 905, TRB, Washington D. C., 1983.

Fu, L. and S. Teply, “Improving the Reliability of Paratransit Service Using Better Travel Time Information”, Proceeding of the Joint Conference of the Canadian Institute of Transportation Engineers, Pacific Northwest ITE Quad Section and Western Canada Traffic Association, April 13~16, 1997, Vancouver, British Columbia

Fu, L. Real-Time Vehicle Routing and Scheduling in Dynamic and Stochastic Traffic Networks, Ph. D. Dissertation, University of Alberta, Edmonton, Alberta, 1996.

Hock, N. C., *Queueing Modelling Fundamentals*, John Wiley & Sons, Baffins Lane, Chichester, England,1996

Kimber, R. M. and Hollis, E. M. (1979) Traffic Queues and Delays at Road Junctions. LR 909, Transportation Road Research Laboratory, Crowthorne.

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

Medhi, J. Stochastic Models in Queueing Theory, Academic Press, Inc. Harcourt Brace Jovanovich Publishers, San Diego, 1991

Newell, G. F. 1982, *Applications of Queueing Theory*, Chapman Hall, London.

Olszewski, P. (1993) “Modeling of Queue Probability Distribution at Traffic Signals. In *Transportation and Traffic Theory*, ed. M. Koshi, Elsevier, New York.

Olszewski, P. (1994) “Modeling Probability Distribution of Delay at Signalized Intersections. *Journal of Advanced Transportation*, Vol. 28, No. 3, pp. 253-274.

Rouphail, M. N. and N. Dutt, (1995), “Estimating Travel Time Distribution for Signalized Links: Model Development and Potential IVHS Applications”, Proceeding of the annual meeting of ITS America, Vol. 1, Washington, D. C., March 15-17, 1995.

Rouphail, N. M. and Akcelik, R. (1990) “Oversaturation Delay Estimates with Consideration of Peaking, Transportation Research Record 1365, TRB, Washington D. C.. pp. 71-81.

Sen, A., P. Thakuriah, X. Zhu and A. Karr, (1997) “Frequency of Probe Reports and Variance of Travel Time Estimates”, *Journal of Transportation Engineering*, Vol. 123, No. 4, 1997.

Sen, A., S. Soot, J. Ligas and X. Tian, (1997) “Arterial Link Travel Time Estimation: Probes, Detectors and Assignment-type Models. Preprint No. 970943, Transportation Research Board, 76th Annual Meeting, Washington, D. C.

Takaba, S. “Estimation of Measurement of Travel Time by Vehicle Detectors and License Plate Readers, Second Vehicle Navigation and Information Systems Conference. Michigan, Society of Automotive Engineers, Warrendale, PA., 1991.

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

Teply, S., D. I. Allingham, D. B. Richardson and B. W. Stephenson, *Canadian Capacity Guide for Signalized Intersections*, Second Edition (S. Teply, ed.), Institute of Transportation Engineering, District 7, Canada, 1984.

Teply, S., D. I. Allingham, D. B. Richardson and B. W. Stephenson, *Canadian Capacity Guide for Signalized Intersections*, Second Edition (S. Teply, ed.), Institute of Transportation Engineering, District 7, Canada, 1995.

TRB, (1994), *Highway Capacity Manual*, 1994 Update, Special Report 209, Transportation Research Board, National Research Council, Washington, D. C., 1985.

Van Aerde, M., B. Hellinga, L. Yu, and H. Rakha (1993) Vehicle Probes as Real-Time ATMS Sources of Dynamic O-D Travel Time Data. Large Urban Systems - *Proceedings of the ATMS Conference,* in St. Petersburg, Florida. pp. 207- 230.