Overflow Delay at a Signalized Intersection Approach Influenced by an Upstream Signal: An Analytical Investigation

A. TARKO, N. ROUPHAIL and R. AKÇELIK

REFERENCE:

NOTE:
This paper is related to the intersection analysis methodology used in the SIDRA INTERSECTION software. Since the publication of this paper, many related aspects of the traffic model have been further developed in later versions of SIDRA INTERSECTION. Though some aspects of this paper may be outdated, this reprint is provided as a record of important aspects of the SIDRA INTERSECTION software, and in order to promote software assessment and further research.
Overflow Delay at a Signalized Intersection Approach Influenced by an Upstream Signal: An Analytical Investigation

Andrzej Tarko, Nagui Roupail, and Rahmi Akcelik

Some initial formulations of general overflow delay formulas that can be applied to isolated intersection approaches as well as those contained in a signalized arterial network are discussed. The main issues addressed are the deficiencies in models of arrival that have been developed to account for the filtering effect of upstream signals, which use the Poisson arrival process as their foundation. A cycle-by-cycle simulation model is the investigation tool. This level of modeling allows for the estimation of the uniform and overflow delay components separately. The results of the model are first tested for the isolated intersection scenario. The simulation model results are then extended to a two-intersection system, and a generalized delay model is calibrated to encompass both the isolated and the system cases. The initial work is limited, however, to fixed-time signal control with no platoon dispersion and no secondary or midblock flows between the two approaches. Only single-lane flow cases are considered. It was found from the simulation results that the inclusion of a parameter $X_c$ (a value of the degree of saturation below which the overflow delay is negligible) in the delay model is justified. The study also confirmed that progression quality has no effect on the overflow delay estimate. Two competing overflow delay model forms were investigated; one used the variance-to-mean ratio of upstream departures and the other used the capacity differential between intersections to reflect the filtering effect. Whereas the first form has been widely quoted and advocated in the literature, the second approach provided better predictions in this study.

Vehicular delay at signalized intersections is a critical component of travel time in an urban road network. Delay contributes to vehicle operating costs and is intimately affected by the traffic engineer's decisions. The Highway Capacity Manual (HCM) (7) uses delay as the sole criterion for determining level of service at signalized intersections and indirectly for level-of-service evaluation on urban arterials.

Of the earlier delay formulas, Webster (2) applied steady-state queueing theory augmented by simulation to evaluate the overflow delay component. Several general formulas have since been proposed that use an index of the dispersion of the arrival and departure processes [see e.g., Darroch (3) and Gazis (4)]. None of the models, however, is directly related to coordinated signalized intersections.

To overcome the difficulties encountered with steady-state queuing models (i.e., infinite delays at capacity), time-dependent delay models were originally conceived by Robertson (5) and were elaborated further by Kimber and Hollis (6) using the so-called coordinate transformation method. Although there is no rigorous theoretical basis for this approach (7), empirical evidence indicates that these models yield reasonable results. A number of time-dependent formulas have found their way into the capacity guides of several countries, including the United States (1), Canada (8), and Australia (9).

This brief overview points to the importance of providing rational delay estimates for internal signals in a network. The work reported on in this paper follows on from the initial work reported in Roupail and Akcelik (10). Initially, the assumptions are kept simple to be effective. Hence, the model will be investigated for a single traffic stream flowing between two closely spaced intersections (i.e., no platoon dispersion or minor flows). The two signals operate at a fixed-cycle length, but could vary in splits and offsets.

The paper is organized as follows. First, a brief discussion of the delay model form is presented and current attempts to model overflow delay at isolated and coordinated signals are discussed. Next, the simulation methodology is described, along with validation results from the single intersection case. The extension to the two-intersection system is discussed next. Two general model forms are then presented to describe the filtering effect of the upstream signal on downstream delay: the first uses the variance-to-mean ratio of upstream departures (equivalent to downstream arrivals), $I$; the second uses a function of the capacity differences at the two intersections. Both model forms are compared and evaluated. The paper concludes with possible extensions of the concept to multiphase, multistream cases, and further research needs in that area are briefly highlighted.

**Analytical Delay Expressions**

The average approach delay per vehicle at a signalized intersection approach can be expressed as the sum

\[ d = d_1 + d_2 \]  

(1)

where $d_1$ is the uniform delay component, which refers to the average vehicle delay experienced assuming that traffic demand is the same for all signal cycles, and $d_2$ is the overflow delay component, which consists of both the random and time-dependent oversaturation effects for a sustained time period, $T$. 

---

A. Tarko and N. Roupail, Urban Transportation Center, University of Illinois at Chicago, 1033 West Van Buren, Suite 700 South, Chicago, Ill, 60607. R. Akcelik, Australian Road Research Board, P.O. Box 156, Nunawading 3131, Australia.
The first delay component \( d_1 \) is known as Webster’s first term and is given by

\[
d_1 = \frac{C(1 - \lambda)^2}{2(1 - \lambda \cdot X)}
\]

(2)

where

\( \lambda = \) effective green to cycle ratio,
\( X = \) volume-to-capacity ratio, and
\( C = \) cycle length.

Several delay models have been derived with the assumption of steady-state conditions (2). These models estimate the queue length under stochastic equilibrium conditions. The queue equilibrium can be attained only if the prevailing traffic volume and capacity are stationary for an indefinite period of time (in practice, however, it need only be sufficiently long) and if the expected flow rate is below capacity.

The coordinate transformation technique converts steady-state queueing models to time-dependent models (9). This process allows the relaxation of the steady-state model restrictions and the extension of the delay model to oversaturated conditions. Acelik (11) proposes a generalized time-dependent expression in the form

\[
d_2 = 900TX^a \left[ X - 1 + \sqrt{(X - 1)^2 + \frac{m(X - X_0)}{QT}} \right]
\]

(3)

where

\( Q = \) capacity [vehicles (veh)/hr]
\( X_0 = \) volume-to-capacity ratio below which overflow delay is negligible.

and

\( X_0 = a + b \cdot s \)

where

\( s = \) saturation flow rate (veh/sec),
\( g = \) effective green, and
\( a, b = \) parameters.

The delay model parameters \( a, b, n \), and \( m \) depend on the distribution of arrivals and departures.

Model parameters as presented by Acelik (11) are given in Table 1.

### TABLE 1 Calibration Parameters for \( d_1 \) in Selected Capacity Guides

<table>
<thead>
<tr>
<th>Method</th>
<th>Model Parameters*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n )</td>
</tr>
<tr>
<td>1985 HCM</td>
<td>2</td>
</tr>
<tr>
<td>Australian</td>
<td>0</td>
</tr>
<tr>
<td>Canadian</td>
<td>0</td>
</tr>
<tr>
<td>HCM (Acelik)</td>
<td>0</td>
</tr>
</tbody>
</table>

*Source: Acelik (11).
*bValue 6 When Platooning Occurs.
estimates at isolated and internal signals. What is lacking, however, is a systematic analysis of the relationship between signal system design parameters (such as cycle, upstream capacity, or offset) and overflow delay at a downstream approach. The goal of this research was to investigate the feasibility of a generalized overflow delay model applicable for both internal and external links in a system.

METHODOLOGY

A cycle-by-cycle simulation model developed by Roushail and Akcelik (10) has been used to generate the data base to be used for model development. Subsequently, multiple regression techniques (linear and nonlinear) were applied for model calibration in the form given by Equation 3. The simulation model is described next, then some issues regarding analytical model calibration are discussed.

Simulation Model Description

A discrete, macroscopic cycle-by-cycle simulation approach was adopted for modeling a two-intersection system. Vehicles are represented as individual entities, but delays are computed for groups of vehicles that have the same headway properties (hence the macroscopic designation). Delay calculations extend the original work of Staniewicz and Levinson (19). Following is the list of assumptions and limitations of the model:

1. Assumptions regarding the arrival process at the upstream intersection:
   a. The number of arrivals per cycle is generated according to the Poisson distribution; and
   b. The arrival pattern within a cycle is strictly uniform, with the arrival headway equal to the cycle length divided by the number of generated arrivals (which varies each cycle).

2. Assumptions regarding the departure process at the upstream intersection:
   a. There are no departures during the effective red time;
   b. The vehicles in the queue (earliest departure time exceeds arrival time) depart at a constant saturation headway;
   c. The vehicles arriving after the queue has discharged cross the stopline at the arrival headway; and
   d. If a cycle is overloaded, overflow vehicles are released first, first-served in subsequent cycles.

The arrival process at the downstream intersection is equivalent to the departure profile at the upstream intersection shifted by the cruise travel time between stoplines. This is the consequence of the following assumptions:

- All vehicles travel between stoplines at the same speed, immediately accelerating and decelerating at the stopline when delayed and
- There are no midblock vehicle or pedestrian flows.

The departure process at the downstream intersection follows the same principles as those given for the upstream approach. Model parameters for signal control are cycle length, effective green times, and signal offset. Parameters for traffic flow are traffic volume, saturation flow rates, and cruise time.

Figure 1 is a schematic representation of the model. The structure of the model allows for the estimation of all delays for two arrival types—random arrivals upstream and platoon arrivals downstream. More important, the simulation provides an option to generate a fixed number of arrivals in each cycle, which is concomitant with the uniform delay component, \( d_1 \) (Equation 1). Note that for the upstream approach, \( d_1 \) is equivalent to Webster's first term (Equation 2). Downstream uniform delays on the other hand are very much tied to progression quality (offset). The ability to generate each delay component separately at each intersection is a key strength of the simulation model.

Delay Estimation Process

All runs were performed assuming steady-state queueing conditions (i.e., \( T = \infty \) in Equation 3). This required that (a) a sufficiently large number of cycles be simulated and (b) the v/c ratio not approach 1. The closer v/c is to 1, the longer the simulation time period required. The simulation model was run for 100 cycles (2 to 3 hr), which allowed the investigators to neglect the initial nonsteady conditions and to cut off the overflow delay at the end of the simulation period. The resulting bias is estimated at less than 0.4 percent for the assumed maximum value of v/c = 0.9. After a Poisson run was completed, a tandem run was initiated using the average demand from the first run. In fact among the 100 simulated runs in the second case, some had \( n \) arrivals, while others had \( n + 1 \) arrivals, to produce the desired average per cycle. The difference in the average vehicle delay produced by the two tandem runs is, by definition, the overflow delay component, as defined in Equation 1.

The coordinate transformation technique was subsequently used to enable model evaluation by comparison with other time-dependent formulas currently used in various countries.

Model Calibration

A generalized overflow delay model in the steady-state form, based on Akcelik (11) was calibrated from simulation. The model form is

\[
d_2 = \frac{k[X - (a + bsg)]I}{Q(1 - X)}
\]

where

- \( X \) = degree of saturation,
- \( s \) = saturation flow (veh/sec),
- \( g \) = effective green (sec),
- \( I \) = variance-to-mean ratio of an arrival process,
- \( Q \) = capacity (veh/sec), and
- \( k, a, b = \) model parameters (in Equation 3, \( m = 8k \)).

For the purpose of calibration using regression analysis, Equation 4 is rewritten as

\[
d_2 = k \frac{X1}{Q(1 - X)} + ka \frac{-I}{Q(1 - X)} + kb \frac{-sgI}{Q(1 - X)}
\]
in which the terms $k$, $ka$, and $kb$ are regression parameters. The calibration data set for the isolated case consisted of 480 simulation runs of the following combinations:

- Cycle length ($C$): 80, 100, and 120 sec;
- Effective green-to-cycle ratio ($g/C$): 0.1, 0.3, 0.5, and 0.7; and
- Degree of saturation ($X$): 0.5, 0.7, 0.8, and 0.9.

The simulation data set corresponds to the reasonably wide range of single lane capacities ($sg = 4$ to 42 veh/cycle and $s = 0.5$ veh/sec).

**Model Evaluation**

The two simulated delay components in Equation 1 were evaluated independently. The simulated uniform component is shown in Figure 2 against Webster's first term. For all practical purposes, the differences are minimal. A systematic bias of about 0.5 sec was found to result from the discrete nature of the simulation model.

The calibrated overflow delay component based on Equation 4 is expressed by the model ($R^2 = .912$)

$$d_2 = 0.456 \frac{[X - (sg/100)]}{Q(1 - X)}$$

(6)

Note that the $I$ value in the simulation varies about 1.0 from one run to the next. For practical purposes, $I = 1$ for isolated intersections. Hence

$$d_2 = 0.456 \frac{[X - (sg/100)]}{Q(1 - X)}$$

(7)

A comparison of simulation results with Equation 7 estimates is shown in Figure 3. The resultant overflow delay model has parameters that are different from those listed in Table 1. Because of multicollinearity effects among parameters, how-

**FIGURE 1** Schematic of simulation model concept.

**FIGURE 2** Comparison of uniform delays—simulated and Webster first term.
ever, a better test is to directly compare the time-dependent form of the delay model in Equation 7 (herein termed the M3 model) with those listed in Table 1 (except for Akcelik's HCM formula). This is shown in Figures 4 and 5 for cycle capacities (sg) of 10 and 40 vehicles, respectively. The graphs indicate that when both cycle capacity and degree of saturation are low, the results given by the M3 model are almost identical to the HCM formula. The differences between the same two methods become more noticeable as X increases. For high X's and low cycle capacity, the model tracks the Canadian formula. In Figure 5 (high cycle capacity) the model tracks the Australian formula. By virtue of the coordinate transformation method, the M3 model (along with the Australian and Canadian models) becomes asymptotic to the deterministic delay model, d = 450 (X - 1). Overall, the model calibrated from the simulation compares very favorably with well-established overflow delay models cited in the literature.

**EXTENSION TO TWO-INTERSECTION SYSTEM**

Two additional factors have been included in this extension—signal offset between the two approaches and variations in upstream and downstream capacity. The first is related to the cyclic arrival pattern, and the second is concerned with the magnitude of arrivals per cycle as well as cycle-to-cycle fluctuations (see Figure 1). The calibration data set is described as follows:

- Cycle length (C): 80, 100, and 120 sec;
- Downstream approach degree of saturation (X): 0.6, 0.8, and 0.9;
- Downstream effective green times (g): 24 and 36 sec for C = 80, 44 sec for C = 100, and 36 and 54 sec for C = 120 (sg = 12 to 27 veh/hr); and
- Upstream effective green g_u varied from g up to a maximum value that is cycle-dependent.

The data reflect the tendency of coordinated phases to dominate the green allocation between coordinated and uncoordinated movements. The obvious case of zero delay when the upstream capacity is lower than downstream is excluded from the data set (e.g., sg_u < sg).

**Offset Impact on Overflow Delay**

The literature is replete with models that relate progression quality to uniform delay [see Fambro (12)]. To confirm previous findings reported in the literature, which show no correlation between overflow delays and offsets (12), a separate simulation experiment was conducted. In it C = 100 sec, g = 44 sec, g_u = 52 sec, cruise time = 20 sec, and X = 0.90. The model was run for each offset between zero and cycle length in intervals of 10 sec. A sample of the results is shown in Figure 6. As suspected, the overflow delay was highly variable (note X = 0.90) and independent of offset. On the other hand, the uniform term was very sensitive to offset selection. The minimum delay was attained at and around the travel time offset of 20 sec.

**Steady-State Delay Model Calibration**

Using Departing Stream I Ratio

By applying the same concepts described in the section, Model Evaluation, for the downstream approach as suggested in the
literature (17,19,20), the steady-state form of the overflow delay model in this case (note $R^2 = 0.738$) is

$$d_2 = \frac{1.02(X - [0.55 + (sg/300)])I}{Q(1 - X)}$$  \hspace{1cm} (8)

The differences between the models (Equations 6 and 8) are partly explained by the different calibration data sets. The isolated intersection case included some data points with very low cycle capacity (4 to 6 veh/cycle) and consequently yielded much higher overflow delays than in other cases. Twelve vehicles per cycle is the lowest capacity value considered for coordinated intersections. Further, the upstream signal impact may not have been appropriately represented through the single index, $I$.

In general, the $R^2$ in the model (Equation 8) is smaller than for the isolated case, although the exact value of $I$ (simulated) was used. This has been one of the reasons for seeking another form of the overflow delay model. Comparison of the calculated delays in the model (Equation 7) with the simulated values revealed some bias for small values of delay.

Figure 7 illustrates the variation of overflow delay and the coefficient $f$ with the difference between upstream and downstream capacities ($sg = 20$ veh/cycle, $C = 120$ sec, and $X = 0.85$). As suspected, overflow delay vanishes when the upstream capacity is less than or equal to the capacity at the downstream intersection. Clearly, this is the result of imposing an upstream constraint, which prevents the downstream intersection from cycle failure. On the other hand, overflow delay is found to be virtually fixed when the capacity difference is large (as the upstream capacity increases). In this case the downstream intersection can be considered to be isolated (from an overflow delay standpoint). The “smooth” transition from the coordinated to the isolated case (at high capacity differentials) implies that distinguishing between the two cases is artificial, especially because the offset impact is nonexistent.

Recall that attempts to express the upstream constraint by means of the coefficient $I$ in the overflow delay formula have been undertaken earlier by Hutchinson (17) and Van As (16). As shown in Figure 7, the coefficient $I$ and overflow delay reach a zero value at different coefficients on the x-axis, although their trends are similar. For traffic arriving downstream, the coefficient $I$ approaches zero when the upstream approach is close to saturation. On the other hand, overflow delay vanishes when the upstream capacity is less than the downstream capacity. Figure 7 illustrates the situation when the downstream degree of saturation is less than unity and the coefficient $I$ approaches zero ($X_u = 1$) on the left side of the point at which overflow delay vanishes. Clearly this is not desirable for estimating a valid $d_2$ value using Equation 8.

Using Capacity Differential Method

Taking into account these facts and the practical difficulties with the estimation of $I$, another more general overflow delay formula is suggested. The form of the analytical model for overflow delay (Equation 4) is retained, but without the coefficient $I$, as follows:

$$d_2 = \frac{k(X - X_u)}{Q(1 - X)}$$  \hspace{1cm} (9)

where

$$k = k_f l_c$$

with $k_c$ is slope parameter for the isolated case ($k$ in Equation 4, and $f = \text{adjustment factor for upstream conditions that is expressed as a function of the difference between the upstream and downstream capacities, and}$ $X_u = b(sg)$ according to results obtained for the isolated case (Equation 6).

The function $f = f[(sg)_u - (sg)]$ shown in Figure 7 should meet the requirements

$$f = 1 \text{ when (sg)}_u >> (sg),$$

$$0 < f < 1 \text{ when (sg)}_u > (sg),$$

$$f = 0 \text{ when (sg)}_u < (sg)$$

where $(sg)_u$ and $(sg)$ are the upstream and downstream capacities in vehicles per cycle, respectively.

The following functional form, which is consistent with Figure 7, has been selected for calibration:

$$f = \begin{cases} 
1 - e^{-\alpha [(sg)_u - (sg)]} & (sg)_u > (sg) \\
0 & \text{elsewhere}
\end{cases}$$  \hspace{1cm} (10)
where \( \alpha \) is a calibration parameter. To achieve the generalized model form, the data base for model calibration included simulated data from both isolated and coordinated cases. Because of limited computer memory, one-half the number of data points were randomly drawn from the original data sets before combining. The calibration data set thus contained 738 observations.

Multiple linear regression was used to determine the best estimates of parameters \( k_s \) and \( b \). Because of the nonlinear form of the function given by Equation 10, the parameter \( \alpha \) cannot be calibrated directly. The optimal value \( \alpha = 0.5 \) was determined from a series of tests which maximizes the \( R^2 \) value (\( R^2 = 0.819 \)). The resulting model has the form for \((sg)_n > (sg)\) and \(X > (sg)/100\), and 0 elsewhere.

Equation 11 provides calibrated values of all delay parameters as follows:

\[
d_d = \frac{0.408[1 - e^{-0.5\{(sg)_n - (sg)\}}][X - (sg)/100]}{Q(1 - X)}
\]

(11)

where

\[
k_s = 0.408,
\]

\[
f = 1 - \exp[-0.5((sg)_n - (sg))],
\]

\[
X_n = (sg)/100.
\]

A comparison of the analytical delay estimates with simulated values is shown in Figure 8 in which \( f \) values calculated from Equation 10 and \( \alpha = 0.5 \) are compared with the simulated values of \( f \). It is clear that the factor \( f \) provides a more appropriate means for estimating the overflow delay when compared with the coefficient \( f \).

**Time-Dependent Delay Model**

Transformation of the steady-state model (Equation 11) into a time-dependent formula is carried out by setting \( n = 0 \) and using the relationship \( m = 8k \) in Equation 3 (11) with \( k = k_f \). The parameters \( k_s \), \( f \), and \( X_n \) are given by Equation 11. Furthermore, \( m_c \) was substituted for \((sg)_n\).

\[
d_d = 900T \left[ \frac{X - 1}{Q(1 - X)} + \frac{3.3[1 - e^{-0.5(m_c - sg)}]X - (sg)/100]}{QT} \right]
\]

(12)

where

\[
T = \text{time period (hr),}
\]

\[
X = \text{degree of saturation,}
\]

\[
Q, sg = \text{lane capacity (veh/hr, veh/cycle, respectively), and}
\]

\[
m_c = \text{maximum number of arrivals per cycle.}
\]

Because of the imposed constraint on demand represented in Equation 12 by \( m_c \), the observed degree of saturation at the analyzed approach cannot exceed value \( m_c/sg \). Examples of values given by the time-dependent formula are shown in Figure 9 (\( Q = 900 \text{ veh/hr}, C = 80 \text{ sec}, \) and \( T = 0.25 \text{ hr} \).

The impact of the constrained demand disappears rapidly with an increase in the difference between upstream and downstream capacity. When the difference approaches 6 veh/cycle, the resulting overflow delays approach those occurring at an isolated intersection.

**IMPLEMENTATION CONSIDERATIONS**

The analytical investigation of overflow delay models presented was aimed at finding a relatively simple model that could incorporate the upstream signal impact. The results obtained should be considered preliminary results. Simplifying assumptions in the simulation model limit the direct implementation of the analytical model (Equation 12) to cases for which an unopposed single vehicle stream is controlled by
a downstream pretimed signal and filtered upstream by a bottleneck. The bottleneck is located close to the signal and is characterized by a fixed capacity (expressed in vehicles per downstream signal cycle). This may describe a pedestrian signalized crossing or an unsignalized bottleneck (e.g., traffic lane closure or parking vehicles) located upstream to the subject approach at a signalized intersection.

Model expansion to more general cases requires the investigation of the following issues:

- How turning movements at the upstream signals and unsignalized midblocks contribute to the total constraint on arrivals at the downstream signal;
- How variations in capacity of minor streams (e.g., filter left turns) should be incorporated in the model (Equation 11) (a substantial short-term variability in capacity of opposed streams precludes the use of its mean value in the strongly nonlinear function, Equation 11); and
- To what extent does cruise time variance influence the distribution of vehicle arrivals at the downstream intersection, particularly with regard to the maximum number of arrivals.

These factors should be incorporated into the model through the adjustment of the maximum number of arrivals, $m_0$, or the factor, $f$, itself.

CONCLUSIONS

The analysis has shown that $X_0$, proposed by Akcelik for the overflow delay model, is justified. In particular, its use is justified for the case of low capacity, as in a lane-by-lane analysis in SIDRA (27).

Overflow delay is independent of progression quality, even in the case of no platoon dispersion when the effect is expected to be the strongest. Thus, there is justification for adjusting only the first component of the HCM delay model with a PF factor, as suggested by Fambro (13). The delay overestimation with the HCM formula at degrees of saturation higher than 1.2 has been confirmed.

When the upstream signal meters traffic flows, the overflow delay at the downstream intersection can decrease dramatically. Although this metering effect is reflected in the variance-to-mean ratio of arrivals at the downstream signal, its use can cause significant delay overestimation. The model proposed in this paper, which uses differential capacity measure, appears to be more appropriate for overflow delay estimation.

Traffic flow metering implies the existence of a maximum degree of saturation that can be observed at a downstream intersection. In the control optimization methods for signalized networks (e.g., TRANSYT), this fact is unfortunately neglected.

The model developed should be considered a preliminary model. It needs to be verified in field studies. Moreover, the inclusion of additional factors into the delay model (e.g., turning movements, midblock flow sources, or cruise time variations) would enhance its usefulness for practical implementation.

REFERENCES


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