DELAY AT UNSIGNALIZED INTERSECTIONS

by

Prof. Dr. Werner Brilon
Ruhr-University Bochum
D 44 780 Bochum
Germany

Tel: +49 234 322 5936
Fax: +49 234 32 14 151
e-mail: werner.brilon@rub.de

Citation:

The estimation of average delay for vehicles at an intersection by analytic methods is based on queuing theory. Due to the variability of traffic demand over time the estimation of delays for time-dependent flow and capacity - also with a temporary overload - is of primary interest. For the solution different approximate approaches are currently used in practice. They are based on the approximation of the priority system by an M/M/1-queue and on the so-called coordinate transformation technique developed by Kimber and Hollis. The original approach can be split into nine different cases. The paper investigates the background of these possible solutions and their quality of approximation. Especially the kind of definition for the sum of all delays plays an important role. For each of the nine cases, average delay formulas are derived. The complete set of results offers new solutions - especially for the more realistic cases. Thus, initial queues at the beginning of the observed peak period as well as different conditions in the post-peak period can be described. As methods for validating these formulas numerically exact results could be calculated by Markov-chains. Also stochastic simulations and empirical data are used to check the approximate solutions against reality. As a result, a set of equations is recommended which can be applied to estimate average delays at unsignalized intersections for well-defined traffic conditions. These seem to be better fitted to the modeled priority type intersections than current approaches.
INTRODUCTION

Guidelines for unsignalized intersections use different methods for the estimation of delays; e.g. the HCM (1) uses the equation by Akcelik and Troutbeck (2). This method is based on an approximate solution which in its basic ideas goes back to Kimber and Hollis (3). That report also contains other approaches leading to rather complex formulas (e.g. (3), eq. 22). One other solution has been proposed by the author (4). All these equations lead to slightly different results. There is no clear insight, which of the existing approaches is more correct or if alternative solutions might be possible.

For average delay estimation at intersections it is important to be able to treat time dependent cases where the traffic volumes at the intersection vary over time. Special problems occur, if the traffic demand exceeds the capacity for a specific period. Here delay estimation turns out to be an application of a more general problem of mathematical queuing theory.

This paper is an extension of another recent paper by the author (5). For several more details the reader is referred to that source. The intension of the present paper is to demonstrate the practical implications of different solutions for the average delay. This seems to be important, since without an understanding of the basics for oversaturated delay estimation significant misinterpretations might become possible. The paper shows that all solutions for average delays at unsignalized intersections are members of a larger family of approximate solutions. A classification of different possibilities for average delay definition is offered. Among all potential conditions the traditional formulas only constitute a solution for rather specific and even unrealistic cases. These considerations underline, that a sound understanding of the formulas’ background is crucial for a correct application. The key to the assessment of a useful delay estimation technique is, however, the comparison with exact and realistic results. Therefore, Markov chain techniques, simulations, and empirical data have been used.

THE "PRIORITY SYSTEM"

As a "priority system" we define the conflict of two movements (FIGURE 1). Here one minor movement is crossing a major, i.e. prioritized, movement. This priority system can be understood as a queuing system where the first space for a vehicle next to the stop line is regarded as the service counter. The further waiting positions constitute the space for the queue. We define:

\[
\begin{align*}
    s &= \text{service time} = \text{time spent by vehicles in the first position} (= 1 / c) \\
    c &= \text{capacity} \\
    d &= \text{delay} = \text{time spent by vehicles in the queue (without the first position)} \\
    w &= \text{waiting time} = \text{time spent by minor street vehicles in the priority system} \\
    &= d + s \text{ (definition according to Heidemann (6))}
\end{align*}
\]

Averages are denoted by small letters, whereas capitals stand for the sum of all delays (D) and the sum of all waiting times (W) respectively. Neither delay nor waiting times are including any geometric delay, like the HCM (1) does when it states to calculate "control delay".

For an overview about the mathematical treatment of the priority system reference is made to Troutbeck and Brilon (7) or Luttinen (8). For our considerations it is important to mention that the model for capacity calculation and traffic performance measures (like delay) are kept independent from each other. Thus, it is possible to calculate capacity as the basic value for
succeeding calculations by any available method. These could be the empirical regression method (cf. Kimber and Coombe (9)), the gap acceptance theory, or the conflict method (10); (11). These methods will not be discussed here in detail.

For undersaturated conditions, i.e. demand q is less than capacity c, the approximation of the priority system by an M/M/1-queue is essential (cf: (3), (12), (6)). Other steady-state delay equations based on gap acceptance theory (like Tanner (13); Yeo and Weesakul (14); cf. also (7)) are too complicated to be used in the following mathematical derivations. The M/M/1-queue is that standardized queuing system with the largest degree of randomness. The two 'M' stand for exponential distributions of arrival gaps and service times. '1' means: one service channel. Unfortunately, the M/M/1-delay is not necessarily identical to the priority system delay (cf. (5)). It should be noted that the representation of the priority system by the M/M/1-queue can cause a bias, especially for large major traffic volumes, i.e. for small capacities. This drawback, however, is inevitable for further treatment of delay for temporary oversaturation. Fortunately the deviations of the priority system from the M/M/1-queue are small for usual parameter combinations.

For the M/M/1-queue the average time of customers in the system (i.e. the waiting time) is

\[ w = \frac{1}{R} = \frac{1}{c - q} = \frac{1}{c \cdot (1 - x)} \]  

The average delay for the M/M/1-queue is

\[ d = w - \frac{1}{c} = \frac{x}{c \cdot (1 - x)} \]  

where

- \( w \) = average waiting time (s)
- \( d \) = average delay (s)
- \( R \) = reserve capacity = c - q (veh/s)
- \( c \) = capacity (veh/s)
- \( q \) = demand volume (veh/s)
- \( x \) = degree of saturation = q/c (-)

**COORDINATE TRANSFORMATION**

To estimate delays for situations where the demand volume q is variable over time and where it could even exceed the capacity c during a specific peak period an approximation is used. This approximation goes back to a concept developed by Kimber and Hollis (3). The method is characterized as coordinate transformation technique by several authors. It is illustrated in FIGURE 2a. There we see the M/M/1-waiting time (eq. 1) as a function of the degree of saturation (= x = q/c ) and the deterministic delay \( d_d \). The deterministic delay is valid for a D/D/1-queueing system. The idea for the approximate solution is:

- For very low saturation the system adapts to a changing demand very quickly. Thus, the time-dependent solution will be very close to the stationary solution represented by the M/M/1-queue.
- For extreme oversaturation (i.e. large x and long period T of oversaturation) a long queue has to be expected. Thus, the randomness of relatively small variances in the arrival pattern at the end of the queue or in the departure process become rather unimportant. Thus, the average delay approaches the deterministic delay. Therefore, the solution for the average delay in the time-dependent system should be a transition between the steady state delay (M/M/1) and the deterministic delay. This transition
curve is what we are looking for as the estimation of average delay for a time dependent queuing system. The same kind of transition can also be constructed using the reserve capacity $R$ (FIGURE 2b).

Three kinds for the sophistication of the approximation can be defined:

A1. additive; x-axis (see FIGURE 2a)\[ \alpha = \beta \]

A2. multiplicative; x-axis (see FIGURE 2a)\[ \frac{\alpha}{1 - \frac{\beta}{x_d}} \]

A3. additive; R-axis (see FIGURE 2b)\[ \alpha^* = \beta^* \]

where

\[ x = \text{degree of saturation} = \frac{q}{c} \]
\[ R = \text{reserve capacity} = c - q \text{ (veh/h)} \]

\[ \alpha, \beta, \alpha^*, \beta^* : \text{parameters, see FIGURE 2} \]

A potential fourth case, a multiplicative R-based approach gives no sense.

Each of these assumptions leads to a specific solution. There is no aspect to prefer one of these approximations. The only reason will be to identify that approach, which is leading to the most realistic solution.

**DETERMINISTIC DELAY**

For the solution we first have to identify the deterministic delay in FIGURE 2. As the D/D/1 system we understand a queuing system where all customers arrive with a headway of $1/q$ ($q = \text{demand volume}$) and where they are served with a constant service time of $s = 1/c$ ($c = \text{capacity}$). For $x = q/c < 1$ (i.e. $R > 0$) there are no delays for customers in such a system. Then the only time which they spend in the system is the service time $s$.

For the analysis of deterministic delay the traffic demand pattern over time has to be defined. Our considerations are illustrated in FIGURE 3. Here we assume that the traffic demand $q$ is constant during the observed peak period. Also the capacity $c$ is assumed to remain constant during this time. However, after the peak both the demand $q_1$ and the capacity $c_1$ may be different compared to the peak period. For these more general circumstances also an initial queue of length $N_0$ is assumed to exist at the beginning of the peak period. Since we concentrate on peak intervals with a potential temporary oversaturation the condition

\[ 0 \leq q_1 < c_1 \quad \text{where} \quad q_1 < q \quad (3) \]

should be valid.

Two different clearance times for the queue must be defined:

- Period $a$ is the time after which the last vehicle arriving during the peak departs.
- $a_1$ is the time after which the expected length of the deterministic queue becomes zero.

FIGURE 3b shows the cumulative number of arrivals (upper line A-B-C-D) and of departures (lower line A-F-D). The total delay then is the area between the two cumulative curves. There are, however, several possibilities which part of the area should be regarded as the relevant sum of delays. Here, different cases for the deterministic delay $D_D$ can be distinguished (TABLE 1). As a general formula for the sum $D_D$ of all delays we can use:

\[
D_D = \begin{cases} 
\frac{1}{2} \cdot \left[ N_0 \cdot T + N_T \cdot (T + a) \right] & \text{for} \quad N_0 > T \cdot c \cdot (1 - x) \\
\frac{N_0^2}{2 \cdot c \cdot (1 - x)} & \text{elsewhere}
\end{cases}
\]

(4)
\[ N_T = \max \left\{ N_0 + c \cdot (x - 1) \cdot T, 0 \right\} \]  

where

- \( N_T \) = maximum deterministic queue length (veh)
- \( N_0 \) = initial queue length (veh)
- \( c \) = time to dissolve the queue due to oversaturation after the peak (according to TABLE 1) (s)

Each case has advantages for specific applications:

**D1**  This definition restricts the consideration on delays which do occur exactly during the relevant peak interval. This definition is the one to be applied when delays from successive peak intervals are added, e.g. over all hours of a whole day.

**D2**  This is the delay, which traffic engineers usually define, when they estimate delays by empirical methods. The D2-definition contains delays experienced after the end of the peak period. But it avoids the integration of delays experienced by vehicles arriving after the peak.

**D3**  This is the total delay which is induced into the system by the temporary overload. But it contains delays experienced after the end of the peak. Even delays for vehicles arriving after the considered peak period are involved into the total delay. For an economic assessment of delays, caused by specific peak periods, this definition is the preferential one, since it represents the total consequences of the overload happening during the peak period.

Each of these 3 D-cases can be combined with cases A1 - A3, such that we receive 9 possible cases, each of which leads to a specific delay equation. It should be noted that D_D covers delays – not waiting times.

From the sum \( D_D \) of total deterministic delay the average delay \( d_D \) is derived by relating \( D_D \) to those vehicles which are exposed to become involved into the queue of waiting vehicles. This means

\[ d_D = \frac{D_D}{N} \]  

where

- \( d_D \) = average deterministic delay (s)
- \( D_D \) = sum of all deterministic delays (s)
- \( N \) = number of vehicles exposed to contribute to \( D_D \) (-)

\( N \) contains all vehicles which arrive during the relevant time period T with the consequence that  
\[ N = q \cdot T = x \cdot c \cdot T \] . Other assumptions for \( N \) might be possible (see (5)) but they give no sense in practice.

One example for the deterministic delay depending on the degree of saturation \( x \) or of the reserve capacity \( R \) is shown in FIGURE 4. There is one limiting case:

\[ x = x_g = 1 - \frac{N_0}{T \cdot c} \quad \rightarrow \quad R = R_g = \frac{N_0}{T} \quad \text{with} \quad d_g = \frac{T \cdot N_0}{2 \cdot (T \cdot c - N_0)} \]  

This is the point (marked in FIGURE 4) which - as a maximum - enables a dissipation of the initial queue (length \( N_0 \)) within the peak period of duration T. Beyond this point (i.e. for \( x > x_g \) or for \( R < R_g \)) the queue at the end of the peak period will be \( > 0 \) for the deterministic system. We see that \( N_0 \) has an influence on the shape of the curves for small \( x \) (i.e. \( x < 1 \)).
For $N_0 > 0$ the value of $d_D$ increases to infinity for $x \to 0$, since due to $N_0$ there is always some delay experienced in interval $(t, t+T)$. With eq. 6, for small $x$ the number of $N$ is also small. This will need some special treatment later in this paper. Only for $N_0 = 0$ the deterministic delay starts from the point $(x = 1, d_D = 0)$. For case D2 the relation $D_D = \text{Function}(x)$ is always linear or nearly linear. D3 is identical with D2 for $q_1 = 0$. With increasing $q_1$, the D3-curve becomes increasingly concave. In the limiting case of $q_1 = c_1$ the curve for case D3 grows to infinity at $x = 1$. The D1-curve is always convex.

At first we concentrate on the simple case where $N_0 = 0$ (i.e. no initial queue) and $c_1 = c$ (i.e. constant capacity). This simplified case is not relevant for case D1 and D3. Thus, first of all we can derive equations representing the average delay for case D2, which then will be used as a reference.

- Case D2 + A1:
  \[
  d = \frac{1}{4 \cdot c} \left[ 2 + c \cdot T \cdot (x - 1) + \sqrt{(2 + c \cdot T \cdot (x - 1))^2 + 8 \cdot x \cdot c \cdot T} \right] \tag{8}
  \]

- Case D2 + A2:
  \[
  d = \frac{T}{4} \cdot \left[ x - 1 + \sqrt{(x - 1)^2 + \frac{8 \cdot x}{c \cdot T}} \right] \tag{9}
  \]

This equation - resulting from the multiplicative approximation A2 - is identical with the Akcelik-Troutbeck-equation (2) which is also used in the HCM (1).

- Case D2 + A3:
  \[
  d = \frac{1}{4 \cdot c} \left[ R \cdot T + 2 - \sqrt{(R \cdot T - 2)^2 + 8 \cdot c \cdot T} \right] \tag{10}
  \]

Even if these formulas look quite different from each other, the numerical results are rather similar. This is pointed out in FIGURE 5 for one example. Nevertheless, it is of interest which of the three approximations A1 - A3 is closer to those average delays, which are the correct values from a theoretical point of view.

**MARKOV-CHAINS**

The only possibility to calculate exact results for the average delay within a time dependent queuing system is given by numerical methods in the form of so-called Markov-chains. This has been described by the author (5). The method produces an exact solution for the average queue length of an M/M/1-system, in which the capacity and the demand can be modified over time in each desired manner, e.g. in 1-minute intervals. The pattern of capacities and demand volume, e.g. like FIGURE 3a can quite easily be implemented. Then the area under the curve for the expected queue length over time is equal to the sum of all delays (like $D_D$ in eq. 4; here, however, for the stochastic system), which together with eq. 6 gives the average delay $d$ for cars arriving during the peak.

Markov-chain calculations have been performed for a variety of parameters. Especially results for average delay according to case D2 were evaluated for several combinations of $c$ and $T$ with emphasis on longer peak periods in the range of $T = 0.25$ h through 1 h. It turned out that on the scale of FIGURE 5 the relation for $d = F(x)$ or $d = F(R)$ estimated by eq. 8 – 10 matched quite well with Markov-chain results. The differences, however, are so small that on the scale of FIGURE 5 they are not well visible. The differences are shown in a finer resolution in (5) for $T = 1$ h. We see that the differences are quite small in absolute terms. The degree of approximation seems to be quite acceptable for practical application. The tendency of the 3 curves remains also similar for other combinations of parameters. Differences become rather significant, if the capacity has very
low values. The results have also been compared by their residual standard deviations between analytic results and Markov-chain calculations (TABLE 2). We see again that the differences are more severe for lower capacities. Here the A2-approximation seems to be the closest. For larger average or capacities, however, approximation A3 is the best, closely followed by A2. In any case the A1-approximation (i.e. the additive approach over x) turns out to be of lowest quality.

SIMULATION

Another method for testing the approximate results with correct values is to use stochastic simulation. A computer program for simulating the priority system (FIGURE 1) has been written. The statistical variation of inputs is only concerning the arrival times of vehicles for both traffic streams. As parameters the following values have been applied: constant critical gaps and follow-up times $t_c = 6 \text{s}$ and $t_f = 3 \text{s}$; drivers behaving consistently and homogeneously; peak period of duration $T$ (here $T = 15 \text{ minutes}$); average delay according to definition D2. FIGURE 6 shows that the calculated curves do - on average - follow the simulated points. The results for the 15-minute average of delay do, however, vary over a quite remarkably wide area. The standard deviation of delays (standard deviation between average delays over 15-minute intervals) is always in the range of 0.7 of the mean of the average delays. For $x < 0.8$ the standard deviation of the 15-minute average delay is even larger than the average.

This illustrates: Even if the statistically true value of average delay can be well estimated by the analytical functions, the average values for delay observed in practice could deviate considerably from these true values. This makes it rather difficult to verify delay equations by empirical methods.

The three curves in FIGURE 6 represent the approximations for case A1 (upper curve, eq. 8), A2 (eq. 9), and A3 (lower curve, eq. 10). The residual standard deviations for the simulated points (simulated average delay relative to eq. 8 – 10) are also shown in TABLE 2. If we try to interpret the small differences, the results support the A3-solution. The difference to the A2-solution is, however, quite small.

All equations compared to simulation results show a tendency to slightly overestimate average delays in the range of $x = 1$. Further analysis shows that this effect is a consequence from using the M/M/1-queue as an approximation for the priority system.

On this background, it can be stated, that the approximation method A3 gives the best correlation to simulation and Markov-chain results. The method A2 is, however, very close up. Based on this experience and in spite of knowing that the same result must not necessarily be obtained also for cases D1 and D3, method A1 is not further treated here.

EFFECT OF $N_0$

Up to now we have studied the unrealistic simplified case of $N_0 = 0$, $c_1 = c$, and $x_1 = q_1/c_1 = 0$. To adjust the solution to more realistic circumstances $N_0$, the initial queue length, should be allowed to have any positive integer value. Then the sum of deterministic delays assumes a function over $x$ or over $R$ like it is illustrated in FIGURE 7.

The direct application of the principle of approximations A1 - A3 does not lead to useful results. Besides the fact that the equations assume unreasonable complicated functions, there is also the problem that the average delay does not only increase for large $x$ (or small $R$) but also in the vicinity of $x \to 0$ (i.e. $R \to c$) due to eq. 6. If $x = 0$, then the term $N = q \cdot T = x \cdot c \cdot T$ in the numerator of eq. 6 becomes 0, such that with
\[ D_D = D_0 = \frac{N_0^2}{2 \cdot c} \]  \hspace{1cm} (11)

where
\[ D_0 = \text{sum of delays for the case of no traffic arriving} \]
\[ \text{during the interval } (t, t+T) \]  \hspace{1cm} (8)

there will be a minimum sum of delay in the denominator. Such a function is not accessible to approximation A1 - A3.

One solution might be to perform the same type of approximation (cases A1, A2, A3) on the scale of D, i.e. with the sum D of delays instead of the average individual delay d. Trying this, the result is quite discouraging because the equations become unacceptably complicated. Nevertheless, a useful solution was found via the treatment of the sum of delays calculated from eq. 9 and 10 by multiplication with \( N = q \cdot T = x \cdot c \cdot T \), which is the number of vehicles arriving during the peak interval \( (t, t+T) \). In addition to this sum of delays, the minimum amount of delays \( D_0 \), which goes back to the initial queue, has to be added. Also the asymptote for the deterministic delay has to be transformed from \( d_D = T/2 \cdot (x-1) \) (in case \( D_2; N_0 = 0 \)) to \( d_D = T/2 \cdot (x-\hat{x}) \) where
\[ \hat{x} = 1 - \left( \frac{1}{T \cdot c} + \frac{1}{T \cdot c - N_0} \right) \cdot N_0 \]  \hspace{1cm} (12)

Thus, for case D2 the sum of delays can be described with a rather good approximation by the equation
\[ D = D_0 + \frac{T^2}{4} \cdot c \cdot x \cdot \left( \Delta x + \sqrt{\Delta x^2 + \frac{8}{c \cdot T}} \cdot x \right) \]  \hspace{1cm} (13)

where
\[ \Delta x = x - \hat{x} \]

The similar result can be obtained by approximation A3 using the reserve capacity \( R \).
\[ D = D_0 - \frac{(c-R) \cdot T}{4 \cdot c} \cdot \left( \Delta R \cdot T + 2 - \sqrt{(\Delta R \cdot T - 2)^2 + 8 \cdot c \cdot T} \right) \]  \hspace{1cm} (14)

where
\[ \Delta R = R - \hat{R} \]
\[ \hat{R} = \frac{N_0 \cdot (2 \cdot T \cdot c - N_0)}{T \cdot (T \cdot c - N_0)} \]

Then the average delay d is calculated from eq. 6 with D from eq. 13 or 14 and \( N = q \cdot T = x \cdot c \cdot T \). Numerical comparisons for several examples show (5): There is a rather good correspondence between the approximate formulas and the exact values represented by Markov-chain results. Only in the range of \( x = 0.3 \) to 0.7 there are smaller differences which increase if \( N_0 \) grows into unrealistic large values (e.g. above 30 vehicles). In any case the preciseness is quite sufficient for practice. For larger \( x \) the degree of approximation is quite good for any \( N_0 \)-value.

**EFFECTS OF THE POST-PEAK PERIOD**

The influence of the capacity and demand, which is prevailing after the peak period and which has an influence on the delay for the vehicles arriving during the peak interval \( (t, t+T) \) in cases D2 and D3 still has to be solved.
For case D2 the effect of the post-peak capacity $c_1$ (which may be different from the peak period capacity $c$) can be taken into account by adding the difference in deterministic delay to the sum of all delays. As a consequence, the sum of delays (eq. 13 and 14) has to be modified as

$$D_{c1} = D - \frac{N_T^2}{2} \left( \frac{1}{e} - \frac{1}{c_1} \right)$$  \hspace{1cm} (15)

where

- $D_{c1}$ = sum of delays for $c_1 \neq c$ (s)
- $D$ = sum of delays for $c_1 = c$ (eq. 13 or 14) (s)
- $q_1$ = post peak traffic demand (veh/s)
- $c_1$ = post peak capacity (veh/s)
- $N_T$ = max. deterministic queue length at the end of the peak period (cf. eq. 5) (-)

In the similar way the additional sum of delay which is due to definition D3 is given by

$$D_{D3} = D + \frac{N_T^2}{2 \cdot c_1} \frac{q_1}{c_1 - q_1} = D + \frac{N_T^2}{2} \frac{x_1}{c_1 \cdot (1 - x_1)}$$  \hspace{1cm} (16)

where

- $D$ = sum of delays for $c_1 = c$ (eq. 13 or 14) (s)
- $x_1$ = degree of saturation after the peak period $q_1/c_1$ (-)
- $R_1$ = reserve capacity after the peak period $c_1 - q_1$ (veh/s)

Finally, the whole set of equations, based on approximations A2 and A3, is given in TABLE 3 as an overview. The formulas in both columns of the table are alternative to each other. Here, the average waiting time $w$ is calculated from the average delay $d$ by adding the weighted average of the service times ($s = 1/c = \text{peak period service time}$, weighted by $c \cdot T$, and $s_1 = 1/c_1 = \text{post-peak service time}$, weighted by $c_1 \cdot a$). Here the solution for D2-A3 is recommended for practical application.

FIGURE 8a shows the average delay using this approach depending on $R$ for the case of an initial queue $N_0 = 0$. For the post peak period both the minor as well as the major traffic volume have been assumed as 0.8 times the peak demand flows. Here we see that down to a reserve capacity of 100 veh/h the delay is always of similar size. Only below $R = 100$ veh/h the delay curves are spreading out with the largest delays happening for small capacities. In any case the required queuing delay ($d_{\text{LOS} = D < 45}$ s) as it is demanded for LOS = D or better by the HCM (11, chapter 17) can be guaranteed with a reserve capacity $R > 80$ veh/h for $N_0 = 0$. FIGURE 8b shows the average delay in relation to the reserve capacity how it is influenced by the initial queue $N_0$. The graph is valid for $q_1 = 0.9 \cdot q$ and $c_1 = c/0.9$.

**CASE D1**

To describe case D1 the following equation for the average delay can be given for case D1 – A2.

$$D_1 = D_0 + \frac{c \cdot T^2}{4} \left( x - 1 - \frac{2 \cdot x}{c \cdot T} + \sqrt{(x - 1)^2 + \frac{4 \cdot x}{c \cdot T} \left( 1 + x + \frac{x}{c \cdot T} \right)} \right)$$  \hspace{1cm} (17)
A solution for the case D1 – A3 is not possible. For the generalized case it leads to an undefined area in the range of $R = \hat{R}$.

EMPIRICAL EVIDENCE

The result have also been compared to empirical data which were obtained from Brilon and Weinert (15) for rural T-junctions in Germany, here for the left turns from the minor road. The comparison could only be made for 5-minute intervals, since longer periods with stationary traffic volumes can hardly be found in reality.

FIGURE 9 shows two examples of rather good correspondence between empirical and theoretical delays for the D2 case. There were other intersections where the correspondence was worse – partly also due to priority reversal behavior of drivers.

CONCLUSION

Average delay estimation at unsignalized intersections is based on approximate equations. These approximations, in principle, match well with more precise estimates for average delay. The set of possible solutions is, however, wider than usually assumed. The most commonly applied formulas are just one special case out of a larger family of potential solutions. It must also be noticed that the most popular solutions, like the Akcelik-Troutbeck-equation (2), are only valid for rather simplified conditions which are not too representative for real world conditions.

Among the alternative sophistications for the approximation the additive x-axis (A1) approach should not be used. The best correspondence between exact and estimated values is obtained by the A3 approach using the reserve capacity.

The classification of definitions for the average delay (case D1 - D3) reveals systematically different results for the average delay. These must be carefully interpreted when using the estimated values. Each of these definitions has its own justification. Usually for practice, however, the D2-case is recommended, since for empirical evaluations of average delays this definition is the usual one.

These results are also confirmed by stochastic simulation. Here, in addition, we also get information about the high values of standard deviations of delays with a coefficient of variation for the 1-hour average delay in a range of 0.7, which reveals another indication that a careful interpretation of average delays at intersections is essential. It can also be shown that empirical observations are in coincidence with the theoretical results.

As a result of the investigation a set of delay equations is recommended, which can be applied to estimate average delays at unsignalized intersections for well-defined traffic conditions. TABLE 3 (right column) contains the recommended set of equations, where case D2 represents the kind of average delay definition as it is used in practice. The paper makes clear that instead of an uncritical use of delay formulas a well sophisticated selection of the adequate equation is required also in practice.
References

List of table titles and figure captions

TABLE 1 Cases for the Classification of Total Delay

TABLE 2 Comparison of results for average delay $d$: Residual standard deviation for Case D2, Cases A1 - A3 ; $T = 15$ minutes

TABLE 3: Formulas for Application

FIGURE 1 The “priority system”

FIGURE 2 Illustration of the method of approximation
a) on the x- axis b) on the R-axis

FIGURE 3 Deterministic delay for the general case.

FIGURE 4 Deterministic delay as a function of the degree of saturation (upper row) and of the reserve capacity $R$ (row on the bottom). For this example $N_0$ has been chosen as 10 (left side) and 0 (right side; both figures: $c = 300$ veh/h; $c_1 = 400$ veh/h; $q_1 = 0.9 \cdot q$; $T = 15$ minutes).

FIGURE 5 Average delay $d$ as a function of the degree of saturation $x$ (left side) and in relation to the reserve capacity $R$ (right side). The figures compare the three kinds of approximation A1 (upper curve), A2 (middle), A3 (lower curve). Also the Markov chain results are indicated by the quadratic points. Parameters for this example: $N_0 = 0$, $c = c_1 = 300$ veh/h, $T = 15$ minutes.

FIGURE 6 Simulation results for the priority system. Each point represents the average delay over a 1-hour-peak period according to case D2 in relation to the degree of saturation $x$. The larger points represent the standard deviation between the results for average delay in the corresponding range of $x$. Parameters for this example:
 a) $T = 15$ minutes; $N_0 = 0$, $c = c_1 = 600$ veh/h and b) $c = c_1 = 300$ veh/h.

FIGURE 7: Average delay according to eq. 13 for an initial queue of $N_0 = 20$ depending on $x$ and according to eq. 14 depending on $R$. Parameters for this example: $T = 1$ h; $N_0 = 20$, $c = c_1 = 600$ veh/h. The figures on the left show results for the full scale of $x$ and $R$ whereas the right side is focusing on the more interesting area close to $x = 1$ ($R = 0$).

FIGURE 8: Average waiting time $w$ as a function of the reserve capacity $R$ (case D2-A3) for $T = 15$ minutes.
FIGURE 9 Comparison of measured waiting times with calculated values. The dark dots represent the average waiting time measured according to definition D2. The small circles represent the calculated waiting time for case D2-A2 calculated for the same traffic volumes as during the measurement interval (T = 300 s).
### TABLE 1 Cases for the Classification of Total Delay

<table>
<thead>
<tr>
<th>case</th>
<th>area in FIGURE 4b</th>
<th>equation for the term a in eq. 4 (for the calculation of $D_D$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>A B C F</td>
<td>$a = 0$</td>
</tr>
</tbody>
</table>
| D2   | A B C E F         | $a = \frac{N_f}{c_1}$  

  a is the time after which the last vehicle arriving during the peak departs. |
| D3   | A B C D E F       | $a = a_1 = \frac{N_f}{(c_1 - q_1)}$  

  $a_1$ is the time after which the expected length of the deterministic queue becomes zero. |
TABLE 2 Comparison of results for average delay d: Residual standard deviation for Case D2, Cases A1 - A3; T = 15 minutes

<table>
<thead>
<tr>
<th>capacity</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
</tr>
</thead>
<tbody>
<tr>
<td>c = 100</td>
<td>12.67</td>
<td>9.45</td>
<td>21.35</td>
</tr>
<tr>
<td>c = 300</td>
<td>9.24</td>
<td>4.09</td>
<td>1.55</td>
</tr>
<tr>
<td>c = 600</td>
<td>7.47</td>
<td>5.05</td>
<td>3.12</td>
</tr>
</tbody>
</table>

compared to Markov-chain calculation ¹)

<table>
<thead>
<tr>
<th>capacity</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
</tr>
</thead>
<tbody>
<tr>
<td>c = 100</td>
<td>65.40</td>
<td>54.00</td>
<td>47.80</td>
</tr>
<tr>
<td>c = 300</td>
<td>30.20</td>
<td>27.10</td>
<td>25.30</td>
</tr>
<tr>
<td>c = 600</td>
<td>19.30</td>
<td>17.90</td>
<td>17.00</td>
</tr>
</tbody>
</table>

compared to simulation ²)

<table>
<thead>
<tr>
<th>veh/h</th>
<th>s</th>
</tr>
</thead>
</table>

¹) values = standard deviations for the difference between the result from eq. 8 - 10 and the Markov-chain result; compared over each x ∈ [0.5 (0.05) 1.4]

²) values = standard deviations for the difference between results from eq. 8 - 10 and the simulated average delay; compared for 1000 1-hour simulation runs with various combinations of q and c within the interval (x= 0.5, 1.4)
### TABLE 3: Formulas for Application

<table>
<thead>
<tr>
<th></th>
<th>A2 (using $x = \frac{q}{c}$)</th>
<th>A3 (using $R = q - c$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>D1</strong></td>
<td>$D_0 = D_0 + \frac{c\cdot T^2}{4} \left( x - 1 - \frac{2\cdot x}{c\cdot T} \right) \left( x - 1 \right)^2 + \frac{4\cdot x}{c\cdot T} \left( 1 + x + \frac{x}{c\cdot T} \right)$</td>
<td>-</td>
</tr>
<tr>
<td><strong>D2</strong></td>
<td>$\Delta x = x - \frac{1}{c\cdot T} (\frac{1}{c\cdot T} + \frac{1}{c\cdot T - N_0}) \cdot N_0$</td>
<td>$\Delta x = D_0 - \frac{(c - R)\cdot T}{4\cdot c} \left( \frac{\Delta R \cdot T + 2 - \sqrt{(\Delta R \cdot T - 2)^2 + 8\cdot c\cdot T}}{2} \right)$</td>
</tr>
<tr>
<td></td>
<td>$a = \frac{N_T}{c_i}$</td>
<td>$D_0 = \frac{N_0^2}{2 \cdot c}$</td>
</tr>
<tr>
<td><strong>D3</strong></td>
<td>$D_1 = D_0 + \frac{N_T}{2} \cdot \frac{x_i}{c_i \cdot (1 - x_i)}$</td>
<td>$D_1 = D_1 + \frac{N_T^2}{2} \cdot \frac{c_i - R_0}{R_i \cdot c_i}$</td>
</tr>
<tr>
<td></td>
<td>$a = \frac{N_T}{c_i \cdot (1 - x_i)}$</td>
<td>$a = \frac{N_T}{R_i}$</td>
</tr>
<tr>
<td></td>
<td>$N_T = N_0 + c \cdot (x - 1) \cdot T$</td>
<td>$N_T = \text{Max}{N_0 - R \cdot T}$</td>
</tr>
<tr>
<td><strong>d</strong></td>
<td>$d = \frac{D}{x \cdot c \cdot T}$</td>
<td>$d = \frac{D}{(c - R) \cdot T}$</td>
</tr>
<tr>
<td><strong>w</strong></td>
<td>$w = d + \frac{T + a}{c \cdot T + c_i \cdot c}$</td>
<td></td>
</tr>
</tbody>
</table>
FIGURE 1 The “priority system”
FIGURE 2 Illustration of the method of approximation
a) on the x-axis  b) on the R-axis
FIGURE 3 Deterministic delay for the general case.
FIGURE 4 Deterministic delay as a function of the degree of saturation (upper row) and of the reserve capacity R (row on the bottom). For this example $N_0$ has been chosen as 10 (left side) and 0 (right side; both figures: $c = 300$ veh/h; $c_1 = 400$ veh/h; $q_1 = 0.9 \bullet q$; $T = 15$ minutes).
FIGURE 5 Average delay $d$ as a function of the degree of saturation $x$ (left side) and in relation to the reserve capacity $R$ (right side). The figures compare the three kinds of approximation $A_1$ (upper curve), $A_2$ (middle), $A_3$ (lower curve). Also the Markov chain results are indicated by the quadratic points. Parameters for this example: $N_0 = 0$, $c = c_1 = 300$ veh/h, $T = 15$ minutes.
FIGURE 6 Simulation results for the priority system. Each point represents the average delay over a 1-hour-peak period according to case D2 in relation to the degree of saturation $x$. The larger points represent the standard deviation between the results for average delay in the corresponding range of $x$. Parameters for this example:

a) $T = 15$ minutes; $N_0 = 0$, $c = c_1 = 600$ veh/h and

b) $c = c_1 = 300$ veh/h.
FIGURE 7: Average delay according to eq. 13 for an initial queue of $N_0 = 20$ depending on $x$ and according to eq. 14 depending on $R$. Parameters for this example: $T = 1$ h; $N_0 = 20$, $c = c_1 = 600$ veh/h. The figures on the left show results for the full scale of $x$ and $R$ whereas the right side is focusing on the more interesting area close to $x = 1$ ($R = 0$).
FIGURE 8: Average waiting time $w$ as a function of the reserve capacity $R$ (case D2-A3) for $T = 15$ minutes.

a) influence of capacity $c$ for an initial queue length $N_0 = 0$

b) influence of initial queue length $N_0$ at a capacity of 300 veh/h
FIGURE 9 Comparison of measured waiting times with calculated values. The dark dots represent the average waiting time measured according to definition D2. The small circles represent the calculated waiting time for case D2-A2 calculated for the same traffic volumes as during the measurement interval (T = 300 s).