DRAŽEN CVITANIĆ, D. Sc.
e-mail: drazen.cvitanic@gradst.hr
DEANA BREŠKI, M. Sc.
e-mail: deana.breski@gradst.hr
University of Split,
Faculty of Civil Engineering and Architecture
Matice hrvatske 15, HR-21000 Split, Republic of Croatia
BORIS VIĐAK, M. Sc.
e-mail: boris.vidjak@igh.hr
Civil Engineering Institute of Croatia, Business Centre Split
Matice hrvatske 15, HR-21000 Split, Republic of Croatia

Science in Traffic and Transport
Original Scientific Paper
Accepted: May 16, 2006
Approved: Mar. 1, 2007

REVIEW, TESTING AND VALIDATION OF CAPACITY AND DELAY MODELS AT UNSIGNALIZED INTERSECTIONS

ABSTRACT

This paper deals with problems related to the capacity and delay models at unsignalized intersections. The parameters of various models are calibrated based on the on-site data obtained in the countries where the models were developed. This paper deals with two problems: The first is the reliability of applied methodology in general, and the second is the acceptability of model parameters for the use in prevailing road and traffic conditions in Croatia.

The first section presents the background review of the gap acceptance theory which is the basis for the development of the capacity and delay models. The mathematical derivation of the basic capacity models is presented as well as the resulting capacity models according to the assumed theoretical distribution of headways in the priority stream. This paper also presents the methods for the calculation of the capacity considering the existence of traffic streams of different hierarchies. The results of various theoretical capacity models testing are presented. The second section presents a review of queuing theory and the resulting stationary and time-dependent delay models. The underlying hypothesis as well as reliability and limitations of the delay models are presented. Also, the Highway capacity manual delay model was tested against field data.

KEY WORDS

unsignalized intersection, capacity, delay, level of service

1. INTRODUCTION

Highway alignment design as well as the selection of cross-section elements is based on the capacity analysis of every segment of road network (highway sections, intersections). The purpose of capacity analysis is to insure that planned highway network could deal with the present and future traffic flows with satisfying quality; i. e. Level of service (LOS). Since unsignalized intersections are the most common intersection type, their functionality has a great impact on the quality of traffic flows on the urban street network and especially on the suburban and rural highway network (at connections of state and county roads).

In the following chapters of this paper, first the development of available capacity methodologies is presented followed by the summary results of conducted capacity model testing (based on the field investigation of an unsignalized intersection in Croatia). The detailed procedure of field data collection, estimation of the critical gap and follow-up time, as well as the results of capacity models testing have already been presented in paper [16]. Next, the development of the modern delay models based on the queuing theory is presented. At the end of the paper the results of delay models testing are presented. Also, the reliability of the model and the applicability of model parameter values for the prevailing roads and the traffic conditions in Croatia are discussed.

1.1 Available methodologies

In the Republic of Croatia, according to *Guidelines about the base conditions that rural public road and their elements need to consider from the safety aspect* [1] the Highway capacity manual HCM methodology [2-4] is recommended for cross-section elements selection.

Today, a great part of the world (USA, German, Canada, Australia…) use theoretical models for the capacity and level of service analysis of unsignalized intersections. These models are based on the *gap acceptance theory* i. e. on the concept of defining the extent to which the drivers are able to utilize gaps of a particular size in the major traffic stream to leave the stop-line and cross (or merge) the intersection. The gap acceptance is defined by two main parameters: *the critical gap* (t_c) and the *follow-up time* (t_f). The critical gap is the minimum length time interval that allows intersection entry to one minor street vehicle. *The follow-up* time is the time span between the departure of one vehicle from the minor stream and the departure of the next, under condition of continuous queuing. Thus, the follow-up time defines the saturation flow rate for the minor street if there were no conflicting vehicles on the major street. The basic concept of the gap acceptance theory is illustrated in Figure 1.1.

In reality, the critical gap is not a constant value; it is a stochastic variable with different values for different drivers and for each individual driver in different conditions. According to the gap acceptance theory the driver is considered *consistent* if they always accept the same gap size. The drivers are considered *homogenous* if they accept the same gap size under the same conditions.

Lack of homogeneity and inconsistency lead to various complicated models. Many investigations [5], [6], [7], [8], showed that models assuming the homogeneity and consistency resulted in minor differences from the complicated models. For the simplicity, all the existing methodologies assume drivers homogeneity and consistency.

For the critical gap value usually the mean value of the distribution of accepted gaps is used and for the follow-up time the average value.

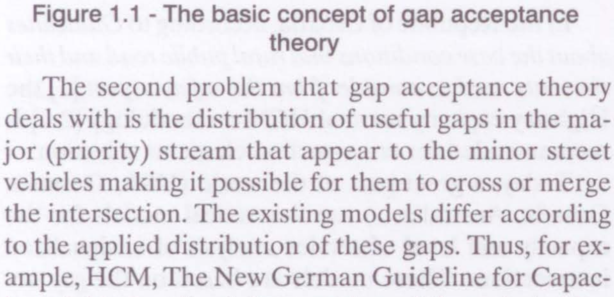

Figure 1.1 - The basic concept of gap acceptance theory

The second problem that gap acceptance theory deals with is the distribution of useful gaps in the major (priority) stream that appear to the minor street vehicles making it possible for them to cross or merge the intersection. The existing models differ according to the applied distribution of these gaps. Thus, for example, HCM, The New German Guideline for Capacity of Unsignalized Intersections [9], and A New Swedish Capacity Manual/CAPCAL 2 [10], use negative exponential distribution, while the methodology developed in Australia (AASIDRA) uses Cowan M3 distribution [11]. The probability density function (f) and cumulative distribution function (F) of the negative exponential distribution are given by (1.1).

$$F(h) = P(h \le t) = 1 - e^{-qt} \text{ for } t \ge 0;$$
$$F(h) = 0, \text{ for } t < 0, \text{ and}$$
$$f(h) = \frac{\mathrm{d}[P(h \le t)]}{\mathrm{d}t} = qe^{-qt} \tag{1.1}$$

Parameter q is the mean traffic volume of major stream (veh/hour). This distribution adjusts well to the real data when the traffic is less than the capacity but this distribution cannot describe the queue formation in the major stream. Many authors tried to define more realistic gaps distribution assuming that the number of vehicles drive in a queue and other drive free without the interaction with other vehicles.

From this type of distributions the most commonly used is the Cowan M3 distribution which models the proportion of free vehicles α with a shift exponential distribution, and the rest of 1-α vehicles have the same time gap t_m. The probability distribution function for this type of distribution is given by (1.2).

$$F(h) = P(h \le t) = 1 - \alpha e^{-\lambda(t - t_m)} \text{ for } t > t_m \text{ and}$$
$$F(h) = P(h \le t) = 0 \text{ for other } t \tag{1.2}$$

where λ is given by (1.3).

$$\lambda = \frac{\alpha q}{(1 - t_m q)} \tag{1.3}$$

This headway model is rather general. For $\alpha = 1$ the displaced exponential distribution is obtained; for $\alpha = 1$ and $t_m = 0$ the negative exponential distribution is obtained. This distribution adjusts well to the intensity and characteristics of traffic flow by parameters α and t_m.

1.1.1 Basic capacity models

Although on almost all intersections more than two traffic streams exist, the basic capacity models rely on the description of only two conflicting streams; the major (priority) and the minor one. The complicated, more realistic models are then evaluated from these basic models. The basic capacity equations are derived from the number of minor stream vehicles $g(t)$ which can enter into a major stream gap duration of t seconds. The expected number of these t-gaps per hour is equal to $q_p f(t)$, and the capacity (veh/hour) of the minor stream provided with these t-gaps in one hour is $3600\, q_p f(t) g(t,)$ where $f(t)$ is a probability density function for the distribution of gaps in the major stream, q_p is major (priority) street volume (veh/sec). To get the total capacity of minor street Q_m we have to integrate over the whole range of major stream gaps:

$$Q_m = q_p \int_0^\infty f(t) g(t) \mathrm{d}t \tag{1.4}$$

Now, the capacity for this simple model can be evaluated by elementary probability theory methods if we assume:

- t_c and t_f are constant values (driver population is homogenous and consistent),
- probability distribution for major stream headways,
- constant traffic volumes for each traffic stream.

Depending on the selection of two different formulations for function *g(t)* we get two different families of capacity equations. The first family assumes a stepwise constant function for *g(t)* shown in Figure 1.2. This approach results in:

$$g(t)=\sum_{n=0}^{\infty} n p_n(t) \tag{1.5}$$

where $p_n(t)$ is the probability that *n* minor stream vehicles enter a gap in the major stream of duration *t, i. e.*

$$p_n(t)=\{1 \text{ for } t_c+(n-1)t_f \le t \le t_c+nt_f \;/\; 0 \text{ elsewhere}\} \tag{1.6}$$

The second family of capacity equations assumes a continuous linear function as shown in Figure 1.2 and results in the following equation:

$$g(t)=0 \text{ for } t<t_0, \text{ and}$$

$$\frac{t-t_0}{t_f} \text{ for } t \ge t_0, \text{ where } t_0=t_c-\frac{t_f}{2} \tag{1.7}$$

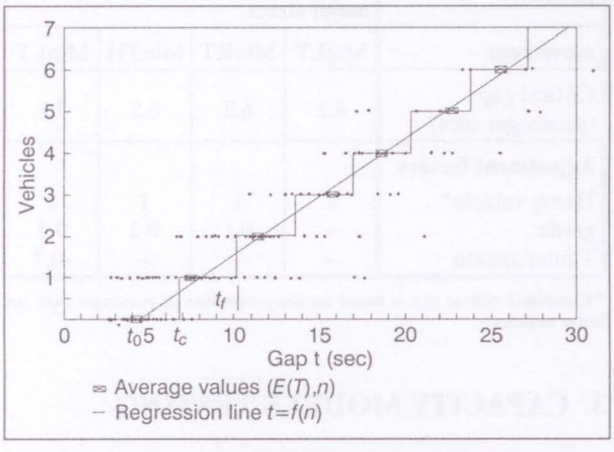

Figure 1.2 - Linear and stepwise function *g(t)* of accepted gaps

Combining equations (1.4) and (1.5), using the exponential distribution of headways in the major stream one can get the capacity equation (1.8) derived independently (in a different manner) by Drew [12], Buckley [13] and Harders [14]:

$$Q_m=q_p\frac{e^{-(q_p/3600)(t_c-t_f)}}{e^{-q_p \cdot t_f/3600}-1} \tag{1.8}$$

This model was used in HCM 1985, and again today in HCM 2000 i. e. HCS+. This model is known as Harders model. Siegloch [15] derived the following capacity equation of minor stream assuming linear function for *g(t)*:

$$Q_m=\frac{3600}{t_f}e^{-q_p t_0/3600} \quad \text{where } t_0=t_c-\frac{t_f}{2} \tag{1.9}$$

This model is used by HCM 1994. Both approaches for *g(t)* produce useful capacity formulae where resulting differences are small. More general solutions for the capacity estimation can be obtained by using the more realistic Cowan distribution and applying the stepwise function for *g(t)*. Then, the associated more general capacity equation is:

$$Q_m=\frac{\alpha q_p e^{-\lambda(t_c-t_f-t_m)}}{e^{-\lambda t_f}-1} \tag{1.10}$$

where $\lambda=\dfrac{\alpha q_p}{(3600-t_m q_p)}$, α = proportion of free vehicles, t_m = minimum inter vehicle tracking headway. This is the so called Troutbeck modification of Harders model shown by equation (1.8).

1.1.2 More realistic capacity models

These models describe the interaction among more streams of different rank of hierarchy. They are derived from the basic models using impedance factors from the queuing theory and also using empirical weighting factors for conflicting volumes of various traffic streams. Figure 1.3 shows the hierarchy of traffic streams for three-leg intersection.

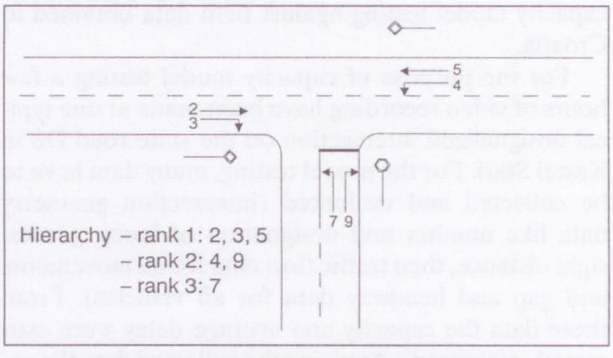

Figure 1.3 - Hierarchy of traffic streams on T intersection

The impedance factor p_0 presents the probability that no vehicle is queuing in the higher priority stream and it is given by:

$$p_0=1-\rho \tag{1.11}$$

where $\rho=q_p/Q_p$ (saturation flow rate), q_p = traffic volume on major (priority) stream (veh/hour), Q_p = major stream capacity (veh/hour).

According to the theory vehicle rank 3 can enter the intersection only when there are no queuing vehicles rank 2, i. e. only during the part $p_{0,\,rank2}$ of the total time. Therefore, the basic value of the capacity for vehicles rank 3, Q_m, must be reduced to $p_0 Q_m$ to get the real potential capacity Q_e as follows:

$$Q_{e,rang3} = p_{0,rang2} \cdot Q_{m,rang3} \tag{1.12}$$

For T intersection this means

$$Q_{e,7} = p_{0,4} \cdot Q_{m,7} \tag{1.13}$$

For each traffic stream the maximum potential capacity should be calculated by this method using the sum of all the conflicting traffic volumes with higher rank than the rank of traffic stream in question. In distinct methodologies for determining the conflicting volume the so-called weighting factors are used. These weighting factors are the result of comparison of real and modelled capacity. In Table 1.1 some of the weighting factors used in HCM 1994, HCM 2000 and German methodology are presented.

Table 1.1 - Weighting factors for conflicting traffic volumes

Subject movement	Movement No.	Conflicting traffic volume q_p
Major street left turn	4	$q_2 + q_3$
Minor street right turn	9	$q_2 + 0.5q_3$
Minor street left turn	7	$q_2+0.5q_3+q_5+q_4$

2. FIELD DATA SURVEY

The following two chapters present the summary results from the research published in [16] about the capacity model testing against field data obtained in Croatia.

For the purpose of capacity model testing a few hours of video recording have been made at one typical unsignalized intersection on the state road D8 in Kastel Stari. For the model testing, many data have to be collected and evaluated (intersection geometry data like number and designation of lanes, grades, sight distance, then traffic flow data for all movements and gap and headway data for all vehicles). From these data the capacity and average delay were estimated (measured). Applying the collected data the capacity and delay models were tested making comparison with the true (field measured) capacity and delay. Also, the critical gap was estimated. Field capacity estimation was conducted using the Kyte method [17]:

$$Q_m = \frac{3600}{t_u + t_{mvup}} \tag{2.1}$$

where

Q_m = true capacity of minor stream; t_u = average service delay of vehicles at the stop lane; t_{mvup} = average move-up time from second position in queue to the stop line.

This method is based on the assumption that the service delay is a randomly distributed variable affected by composition and volume of major conflicting streams as well as by the gap acceptance process. The sum of service and move-up time is a variable that presents the average time that each minor stream vehicle is on the stop line. Based on the queuing theory the capacity is then the inverse of the sum of service and move-up time.

The values of critical gap and follow-up times used in the HCM are estimated on the field data from intersections in the USA. Thus, the applicability and reliability of using these data for the prevailing conditions in Croatia could be questioned. The Troutbeck maximum likelihood method [18] for critical gap estimation was used in [16] because it gives the consistent estimation with respect to traffic volumes changes. It rests on the hypothesis that a critical gap is greater than the driver's maximum rejected gap and smaller than the driver's accepted gap. It assumes log-normal distribution of accepted and rejected gaps.

The investigations conducted in [16] indicated that there is no significant difference between values on the analyzed intersection and the values presented in Table 2.1 which are used in the HCM 2000.

Table 2.1 - Recommended values of critical gaps according to HCM

geometry	Single line per direction on the major street			
movement	MajLT	MinRT	MinTH	MinLT
Critical gap (passenger cars)	4.1	6.2	6.5	7.1
Adjustment factors				
Heavy vehicle*	1	1	1	1
grade	–	0.1	0.2	0.2
T intersection	–	–	–	-0.7

* Combined critical gap is based on the proportion of passenger cars and heavy vehicles

3. CAPACITY MODELS TESTING

For the capacity model testing, two measures of effectiveness were used in [16]: mean absolute error (MAE) and mean absolute percent error (MAPE). Both measures were used to evaluate the quality of models by comparison of the modelled and field capacity.

MAE and MAPE are defined by equations (3.1) and (3.2)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Q_m^j - Q_f^i\right| \tag{3.1}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Q_m^j - Q_f^i}{Q_f^i}\right| \tag{3.2}$$

where n is the number of data, Q_m is the model capacity and Q_f is the field capacity.

Supplemental parameters considered were the regression analysis parameters such as coefficient of determination R^2. First, the testing was performed using the parameters values (critical gap and follow-up time) estimated in the field, and then with the parameters from HCM 2000. This way, first the quality of the model was tested, and then the conclusions about the possibility of using the parameters from HCM for the prevailing road and traffic conditions in Croatia was evaluated. The basic capacity models testing indicated good correlations between the model results and the field capacity for both tested models (Harders model and model using Cowan distribution). The Harders model used in HCM gave better results than the more realistic model which uses parameter a for description of free flowing vehicles. Since in Croatia there is no traffic database needed for evaluating the dependency of this parameter and the traffic intensity, the simplest way is to use the Harders model where the unique model parameter is traffic flow q, which is the parameter commonly measured. Therefore, in [16] only the Harders model using exponential distribution was tested for the more complicated capacity models. The impedance effect, i. e. the impact that the queuing vehicles on the major street have on the minor street capacity, is tested using the specific field values of model parameters and the weighting factor of 1 for right turn manoeuvres from the major stream.

Such model testing as well as field investigations have shown that a major street left turn has a disproportionate effect on the minor street capacity compared with other movements because of queuing on the major street. The model testing presented in [16] showed significant improvement of the model estimation when factor of 2 for major street left turn movement as proposed in [8] was used. The results were symmetrically distributed around the 45 degree line as shown in Figure 3.1. Models testing showed that better

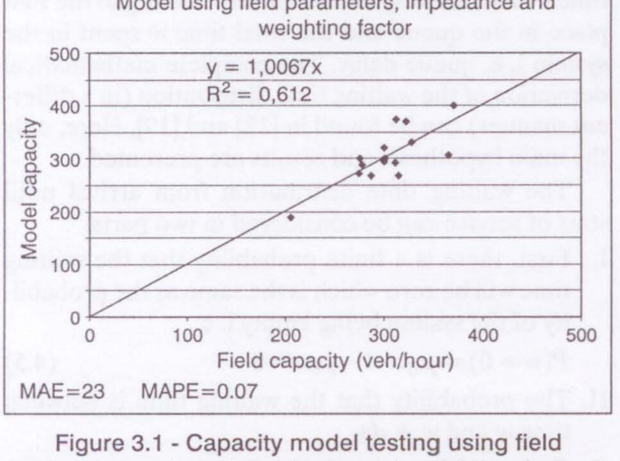

Figure 3.1 - Capacity model testing using field parameters values, impedance and weighting factor 2

results were achieved by using the site specific parameters, and the parameters values used in the HCM 2000 gave better results than values used in the HCM 1994. This investigation showed that the HCM capacity model resulted in fairly good estimation of the measured field capacity at one intersection in Croatia.

4. QUALITY OF TRAFFIC FLOW (LEVEL OF SERVICE)

4.1 INTRODUCTION

The usual way of solving operational problems involving traffic flow at the intersections is to ensure that the average capacity can handle the average flow, so that persistent traffic jams do not occur. However, because flow fluctuates, guaranteeing that highway capacity can handle traffic demand on the average does not preclude the formation of bottlenecks.

The quality of traffic flow i. e. Level of service (LOS) at an intersection is usually represented by control delay. It is average delay of a vehicle caused by the type of intersection control type (stop sign, traffic signal) consisting of queue delay and geometric delay. Geometric delay is caused by decelerating, accelerating and merging the intersection. Some parts of the geometric delay are already included in the queue delay. The average queue delay is a function of flow intensity in major and minor stream q_p and q_m, proportion of free flow vehicles and the distribution of platoon size length in the minor and major streams. The queuing theory deals with these problems so that the delay estimations are based on the distributions of arrivals and service times. All models which use gap acceptance theory are based on the queuing theory.

Figure 4.1 presents the components of delay.

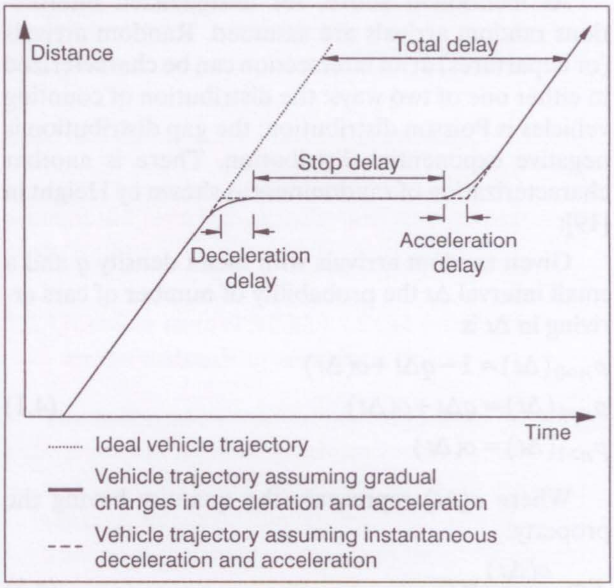

Figure 4.1 - Components of delay

4.2 MAIN HYPOTHESIS OF THE QUEUING THEORY

The queuing theory developed in order to describe the behaviour of a system providing services (capacity Q) for randomly arising demands (flow q). The fundamental idea of the theory is that congestion is manifested through delay caused by an interruption in the flow pattern. To specify the queuing system completely the following must be known: the distribution of arrivals, the queue discipline (first-in-first-out or random service), number and configuration of channels and distribution of service times of the channels. In general, the question of interest of the queuing theory are the distribution of the length of queue, the distribution of the waiting time in the queue and the percent of time during which the system is idle. The answers to these questions depend directly on the nature of the input and service time distributions. Usually, for describing the operation at unsignalized intersection the so-called M/M/1 system is used, where M denotes random distribution. So M/M/1 denotes a single channel queue with a random (Poisson arrivals) and random (negative exponential) service. For the intersections with traffic signals the more realistic assumption is that a vehicle will leave the intersection in uniform increments i. e. service time is uniformly distributed.

Figure 4.2 presents a scheme of a typical queuing system.

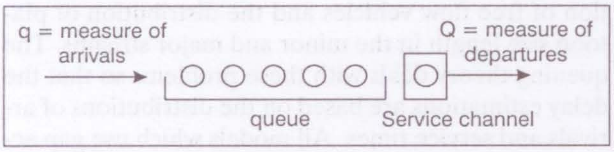

Figure 4.2 - Presentation of a queueing system

As mentioned above, for unsignalized intersections random arrivals are assumed. Random arrivals (or departures) at an intersection can be characterized in either one of two ways: the distribution of counting vehicles is Poisson distribution; the gap distribution is negative exponential distribution. There is another characterization of randomness as shown by Height in [19]:

Given random arrivals with mean density q and a small interval Δt the probability of number of cars arriving in Δt is

$$p_{n=0}(\Delta t) = 1 - q\Delta t + o(\Delta t)$$

$$p_{n=1}(\Delta t) = q\Delta t + o(\Delta t) \tag{4.1}$$

$$p_{n>1}(\Delta t) = o(\Delta t)$$

Where $o(\Delta t)$ represents the quantity having the property:

$$\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0 \tag{4.2}$$

These equations are results of using the Poisson distribution for the arriving vehicles. These three conditions are equivalent characterization of randomness of the car in the lane.

Based on these conditions it is now possible to define the following statements:

- the probability of arrival in a short time interval is proportional to the length of the interval (with a proportionally constant q)
- the service time has negative exponential distribution with mean $1/Q$.

Now, the key characteristics of the queuing theory for single channel random arrangement could be derived from the probability that the queue is in the some state n at time t considering the possible situations that a queue would be in state n at time $t+\Delta t$ as shown in [19]. Using the moment generating function we can see that a queue length is geometric distributed so that the asymptotic probability of the queue length probabilities is given by:

$$p_n = (1-\rho)\rho^n \tag{4.3}$$

Thus, the expectation and variance are:

$$E(n) = \frac{\rho}{(1-\rho)}; V(n) = \frac{\rho}{(1-\rho)^2} \tag{4.4}$$

For the proper understanding of the behaviour of the queue it is important to notice that the variance of the geometric distribution is always greater than its mean. When the queue approaches saturation, so that ρ is nearly unity, the variance becomes very large. It means that the average queue length is subject to enormous fluctuations which may lead to substantial waiting periods i. e. delays. In other words for small changes in capacity one can get large changes in delay.

4.2.1 Waiting times

The most important parameters describing traffic operations (merging, crossing) at an intersection is the distribution of such random variables as the waiting time v of an arriving vehicle before coming to the first place in the queue and the total time w spent in the system i. e. queue delay. The complete mathematical derivation of the waiting time distribution (in a different manner) can be found in [12] and [19]. Here, only the main hypothesis and results are presented.

The waiting time distribution from arrival until start of service can be considered in two parts:

I. First, there is a finite probability that the waiting time will be zero which is the same as the probability of the system being empty i. e.

$$P(w=0) = p_0 = 1-\rho, n=0 \tag{4.5}$$

II. The probability that the waiting time is between time w and $w + dw$

$$P(w < \text{waiting} < w+dw) = f(w)dw, n>0, \tag{4.6}$$

where $f(w)$ is the waiting time density function described by (4.5) and (4.6) and illustrated in Figure 4.3.

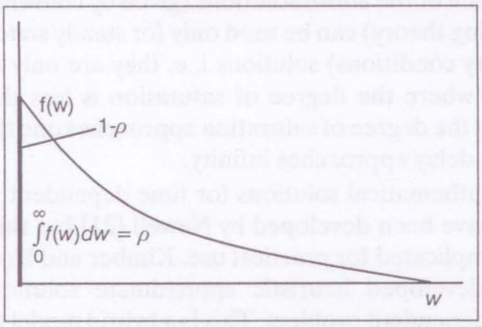

Figure 4.3 - Waiting time density function

The area below the function represents all possible combinations that will yield the waiting time between w and $w + dw$. Theoretically, such a delay is possible as long as there is at least one vehicle in queue. This can be expressed as

$$P(n\geq 1)=\sum_{n=1}^{\infty} p_n \tag{4.7}$$

In order for the waiting time of a vehicle to be exactly between w and dw all the vehicles in the system ahead of the subject vehicle except the one immediately ahead must be served in dw. This is the product of two probabilities:

$$p_{n-1}(w)=\frac{(Qw)^{n-1}e^{-Qw}}{(n-1)!} \text{ and} \tag{4.8}$$

$$p_1(dw)=Qdw \tag{4.9}$$

Summed over (4.8) and substituted in (4.9) results in

$$f(w)dw=\sum_{n=1}^{\infty} p_n p_{n-1}(w) p_1 dw=$$

$$=Q\rho(1-\rho)e^{-Qw}dw\sum_{n=1}^{\infty}\frac{(\rho Qw)^{n-1}}{(n-1)!} \tag{4.10}$$

Thus, the probability density function for the waiting time is

$$f(w)=\rho(Q-q)e^{-w(Q-q)},\ w>0 \tag{4.11}$$

And the cumulative form is

$$p(w>w)=\int_{w}^{\infty} f(w)dw=\rho e^{-w(Q-q)} \tag{4.12}$$

So the probability of waiting time being less than w is

$$P(0<w<w)=\int_{0}^{w} f(w)dw=\rho-\rho e^{-w(Q-q)}, \tag{4.13}$$

Using the moment generating function the average waiting time of one vehicle is

$$E(w)=M_w'(0)=\frac{q}{Q(Q-q)}=\frac{1}{Q}\frac{\rho}{1-\rho} \tag{4.14}$$

The density function for the total time of an arrival v in the system may be derived in the same manner as for w. The result is the following:

$$E(v)=\frac{1}{Q-q}=E(w)+\frac{1}{Q} \tag{4.15}$$

The relationship between the expected delay $E(v)$ and the expected number in the system $E(n)$ for the M/M/1 case is given by

$$E(v)=\frac{E(n)}{q} \text{ and } q=\frac{E(n)}{E(v)} \tag{4.16}$$

The expected values of state of the system with random arrivals and for some general distribution of service times is given by the so-called Pollaczek-Khintchine equation:

$$E(n)=\rho+\frac{q^2 Var(n)+\rho^2}{2(1-\rho)} \tag{4.17}$$

where $Var(u)$ is the variance of the service time distribution.

This equation is easy to use since $Var(v)$ is 0 for regular service and $Var(u)=1/Q^2$ for random service (negative exponential). The expected waiting times for random and uniform departures are given in Table 4.1.

Table 4.1 Mean values

	Mean queue length E(n)	Mean waiting time in the queue E(w)	Mean total delay (waiting and service time) E(u)
Random service	$\frac{\rho}{1-\rho}$	$\frac{1}{Q}\frac{\rho}{1-\rho}$	$\frac{1}{q}\frac{\rho}{1-\rho}$
Regular service	$\frac{2\rho-\rho^2}{2(1-\rho)}$	$\frac{\rho}{2Q(1-\rho)}$	$\frac{2-\rho}{2Q(1-\rho)}$

4.3 DELAY MODELS

4.3.1 Stationary delay models

Equations presented in Table 4.1 were the basis for developing of the modern intersection delay models. Fairly realistic equations for the delay estimates can be derived from the queuing theory model M/G/1, where M represents random arrivals, and G some general distribution function of service times (time spent at the first position in the queue until merging or crossing the intersection) and 1 represents one approach lane.

Using the queuing theory, as shown above, this delay can be modelled with some variant of Pollaczek - Khintchine equation

$$D_q=\frac{\rho E(u)(1-C_u^2)}{2(1-\rho)} \tag{4.18}$$

where

E(u) is average service time ($1/Q$)

$C_u = \frac{\sqrt{Var(u)}}{E(u)}$ is coefficient of variation of service times distributions

Var (u) is variance of service time distribution.

The total delay of minor street vehicle can be expressed as

$$D = D_q + E(u) \tag{4.19}$$

Generally, the average service time is the reciprocal of capacity ($1/Q$). Using this expression the following model of the Pollaczek-Khintchine equations is derived:

$$D = \frac{1}{Q}\left(1 + \frac{\rho}{1-\rho}C\right) \tag{4.20}$$

where

$$C = \frac{1+C_u^2}{2} \tag{4.21}$$

This is a general formulation and the problem is to estimate the value of C. Only the extremes can be defined:

- Regular service: Each vehicle spends the same time at the first position. Thus, it gives $Var(u) = 0$, $C_u^2 = 0$ i $C = 0.5$.
- Random service: the times vehicles spend at the first position are exponentially distributed. This gives the following solution for M/M/1 system: $Var(u) = E(u)^2$, $C_u^2 = 1$ and $C = 1.0$.

Using the capacity equation which is a function of the critical gap and follow-up time (these values are related with intersection geometry) geometric delay is indirectly included in the Pollaczek-Khintchine model. Using various types of distribution of service time one gets various delay equations. For example Troutbeck in [20], using the approximation of C values according to applying service time distribution, suggested the following expressions for calculating average delay

$$D = D_{\min}\left(1 + \frac{\gamma + \varepsilon\rho}{1-\rho}\right) \tag{4.22}$$

where

γ, ε are constants

$\rho = q_m/Q_m$ is the degree of saturation

Q_m = minor stream capacity

$D_{min} = 1/Q_m$ average service time

This is a practical approximation of Pollaczek-Knitchine equation (4.18) where parameter C is expressed as $(\gamma + \varepsilon)$ and D_{min} is equivalent for $1/Q_m$.

These terms are the function of the critical gap, follow-up time and the headway distribution. Using different distributions of major stream headway results in different delay equations.

4.3.2 Time dependent delay models

Each of the above solutions (given by conventional queuing theory) can be used only for steady state (stationary conditions) solutions i. e. they are only applicable where the degree of saturation is less than 1. When the degree of saturation approaches one the average delay approaches infinity.

Mathematical solutions for time dependent problem have been developed by Newell [21] but this was to complicated for practical use. Kimber and Hollis in [22] developed heuristic approximate solution for time dependent problem. This is a hybrid model which for the estimation of delays in under-saturated conditions uses Pollaczek-Knitchine formula, for oversaturated conditions it uses deterministic approach, and in transition from unsaturated to saturated conditions it uses a coordinate transformation system.

Although less complicated than Newell theoretical solution, it is still impractical for engineering use because it needs the detailed data that are not usually collected at intersections such as queue length at the beginning of analyses and the variations of flow in the peak period.

A simpler equation can be obtained by using another approach to coordinate transformation method. The starting point of this method is that the steady state solution is acceptable for sites with a low degree of saturation and the deterministic solution is satisfactory for sites with a high degree of saturation (say 2 or more).

The steady state solution is

$$D_s = D_{\min}\left(1 + \frac{\gamma + \varepsilon\rho_s}{1-\rho_s}\right) \tag{4.23}$$

The deterministic solution is given by

$$D_d = D_{\min} + \frac{2L_0 + (\rho_d - 1)q_m T}{2q_m} \text{ za } \rho_d > 1 \text{ and}$$

$$D_d = 0 \text{ in any other case} \tag{4.24}$$

where

L_0= initial queue length,

T = time of analysis,

q_m = entry capacity.

These equations are illustrated as diagrams in Figure 4. 4.

For a given average delay the coordinate transformation method gives a new degree of saturation ρ_t, which is related to the steady state degree of saturation ρ_s and the deterministic degree of saturation ρ_d, such that

$$\rho_d - \rho_t = 1 - \rho_s = a \tag{4.25}$$

Equation (4.25) ensures that the transformed equation will asymptote to the deterministic equation and gives a family of relationships for different degrees of saturation and periods of operation.

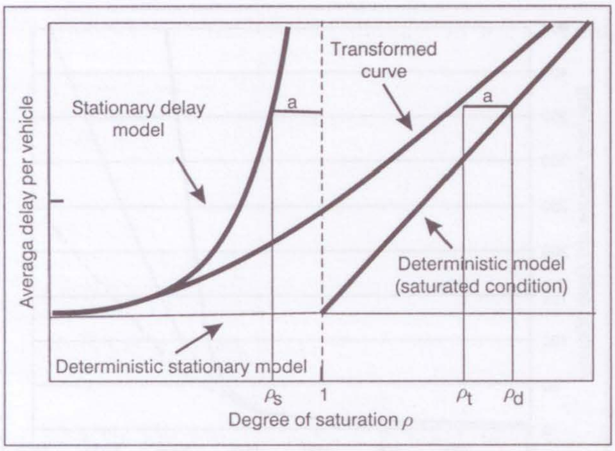

Figure 4.4 - Coordinate transformation method

The rearranging of the equations (4.23) and (4.24) yields two equations for ρ_s and ρ_d as a function of delays D_d and D_s

$$\rho_s = \frac{D_s - D_{\min} - \gamma D_{\min}}{D_s - D_{\min} + \varepsilon D_{\min}} \text{ and}$$

$$\rho_d = \frac{2(D_d - D_{\min}) - 2L_0 / q_m}{T} + 1 \quad (4.26)$$

Rearranging by putting $D=D_s=D_d$, $\rho=\rho_s$ one can obtain the equation for delay (still complicated). A simpler equation was developed by Akcelik and Troutbeck [23]. From (4.25) they expressed a by (4.23) and then by equalization of degree of saturations values and rearrangement they obtained the following equations for delay which can be used in time dependent conditions

$$D - D_{\min} = \frac{1}{2}\frac{L_0}{q_m}\frac{(\rho-1)T}{4} + \sqrt{\left[\frac{L_0}{2q_m} + \frac{(\rho-1)T}{4}\right]^2 + \frac{TD_{\min}(\gamma + \varepsilon\rho_s)}{2}} \quad (4.27)$$

Solution for M/M/1 system can be derived if ε is set to 1, γ is set to 0 and D_{min} is set to $1/q_m$

$$D = \frac{1}{q_m} + \frac{T}{4}\left\{(\rho-1) + \sqrt{(\rho-1)^2 + \frac{8\rho}{q_m T}}\right\} \quad (4.28)$$

This equation can be used to estimate the average delay under oversaturated conditions. This formulation was used for delay calculation in HCM 1997 and the same formulation is used HCM 2000 in the following format

$$D = \frac{3600}{q_m} + 900T\left\{(\rho-1) + \sqrt{(\rho-1)^2 + \frac{(q_m/3600)\rho}{450T}}\right\} + 5 \quad (4.29)$$

where constant of 5 sec/veh represents the average geometric delay in seconds.

4.3.3 Impact of capacity estimation on delay estimation

In the calculation of the capacity of an intersection it is assumed that the headways in major stream have some specific distribution. This distribution can be changed to another distribution and a new equation for capacity can be developed as shown in Chapter 1. The capacity models estimates are independent of the order in which the major stream headways arrive. The intersection would have the same capacity if the gaps offered to the minor street vehicles were ordered from the smallest to the largest as if they were ordered from the largest to the smallest, but the delay would be significantly greater in the former case. So, the delay models are quite sensitive to distributions of major stream headways. They can result in significant differences in results as shown in Figure 4.5. The distribution of queue lengths in major stream which depends on the proportion of free vehicles (α) has significant impact on results, too. This is shown in Figure 4.6.

From the delay equation one can see that the capacity or degree of saturation $\rho = \frac{q}{Q_m}$ is the key variable for the estimation of average delay. For the same traffic volume in the major stream, the average delay for minor street vehicles significantly change for small changes in capacity in the near-saturated conditions. A small mistake in the capacity estimation could result in great mistakes in the estimation of average delay which is shown in Figures 4.5 and 4.6.

Thus, we can see, although the capacity models which use deferent headway distributions give almost the same results, the delay models, for small differences in capacity can result in significant differences of delay estimates. Another factor that has a significant impact on the average delay estimation is the duration of the peak period T. This sensitivity is shown by Figure 4.7.

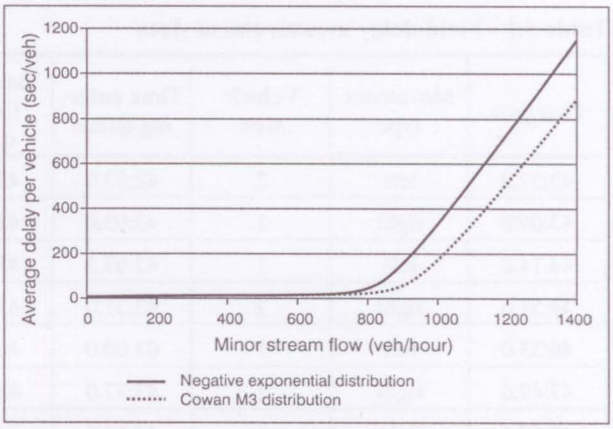

Figure 4.5 - Impact of headway distribution on the average delay results

Note: Major stream flow is 900 (veh/hour)
Same values of t_c and t_f are used

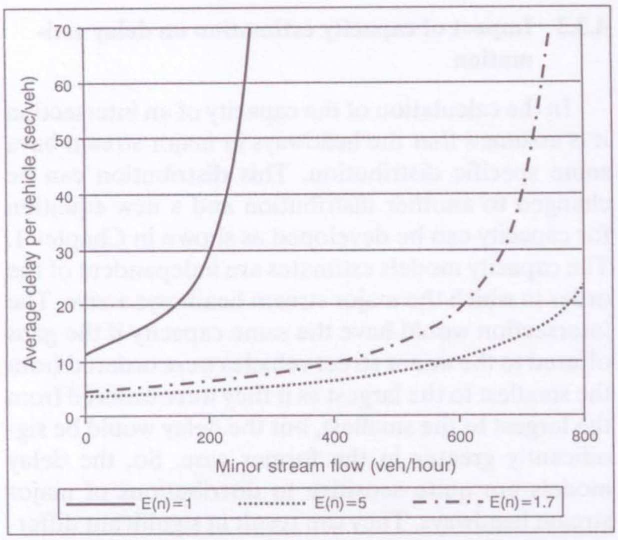

Figure 4.6 - Impact of average number of vehicles in a queue on the average delay results

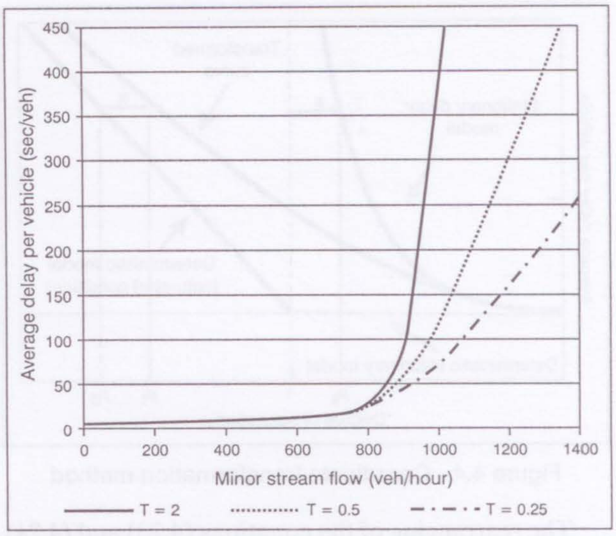

Figure 4.7 - Relationship between the duration of peak period and average delay

5. DELAY MODEL TESTING

All of the mentioned models can estimate delay under steady state condition. Only few models (Newell theoretical model, Kimber-Hollis model...) and some form of the delay model used in HCM can estimate delays for time-dependent conditions. The complexity of the theoretical models and Kimber-Hollis model limits their use. So, for evaluation and use in this research HCM delay model was selected.

5.1 FIELD DELAY MEASUREMENT

For the delay model testing it is necessary to measure the field delay. At the analyzed intersection the delay was measured for every vehicle from the field data based on the intervals of 10 minutes. The 10-minute intervals were used to reduce the variance. In order to determine all the delay components the following data were derived: pass time of each vehicle, movement direction, lane usage, vehicle type, time a vehicle enters queue, time a vehicle comes to the first position of queue, time a vehicle exits a queue.

The individual delay of each vehicle can be derived from the data shown in Table 5.1. The difference between the time vehicle comes to the first position in the queue and the time it enters a queue is the queue delay and the time a vehicle spends at the first position until departure is service delay. Total delay is a sum of these components.

The delay data were sorted according to the departure time of the vehicle to keep it consistent with measurements of the flow rate.

In order to be consistent with the delay models which estimate the average delay of minor stream vehicles during a given period in this study the delay data

Table 5.1 - Field delay measurement data

Pass time	Movement type	Vehicle type	Time entering queue	Time when forst in queue	Time exiting queue	Queue delay	Service delay	Total delay
42:57.0	left	2	42:53.0	42:53.0	42:56.0	0	3	3
43:09.0	right	1	43:03.0	43:03.0	43:07.5	0	4.5	4.5
44:14.0	left	1	43:03.5	43:10.0	43:16.5	6.5	6.5	13
44:51.0	right	1	43:31.0	43:31.0	43:38.0	0	7	7
46:33.0	left	1	03:00.0	46:20.0	46:31.5	0	11.5	11.5
47:49.6	right	1	43:37.0	43:40.0	44:13.5	3	33.5	36.5
49:09.6	right	1	44:12.0	44:15.0	44:18.0	3	3	6
49:52.5	right	1	44:13.0	44:19.5	44:24.5	6.5	5	11.5
49:59.5	right	1	44:20.0	44:25.0	44:26.5	5	1.5	6.5

were derived from field data by averaging all of the individual delays experienced by all of the minor stream vehicles that departed the intersection during a given interval.

5.2 Model testing

The selected HCM model was tested against the field data using the regression analysis. Figure 5.1 shows the delay estimation results while using the recommended values of the capacity model parameters. Each data point represents an average result for a 10-minute interval.

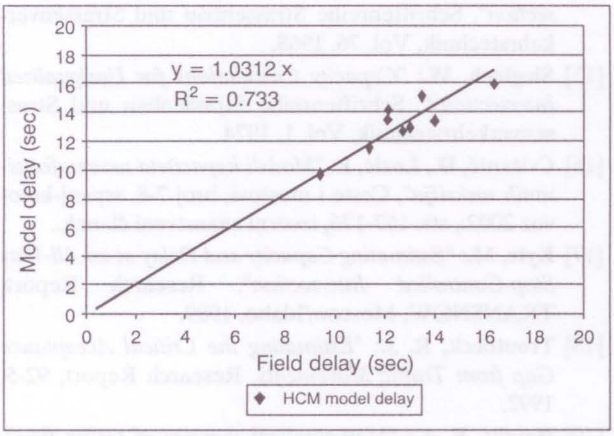

Figure 5.1 Delay model testing

The results show good correlation between field and estimated delay (0.878). MAE is 0.72 seconds, and MAPE is 6%. It is worth noting that at the analyzed intersection the saturation was not big. For a more intensive flow, i. e. values of ρ near 1, small differences in the capacity could result in great differences in delay estimation. So, it could be expected that testing of delay models in more saturated conditions would result in an inferior estimate. It could be stated that the delay estimation for the near-saturated conditions (ρ>0.8) are unstable. For a slightly different estimation of capacity we get significantly different value of delay. For the under-saturated condition this model gives good estimates for intersections with normal traffic and geometric conditions. Until a more robust model is developed, this model is acceptable for practical engineering use, especially because in the case of over-saturation something has to be changed in traffic (signalization) or geometric (additional lanes) layout of the intersection.

6. CONCLUSION

Based on the various model considerations and test results it can be concluded that using the Cowan distribution, which better describes the real traffic flow, results in overestimation of the real capacity. This is probably the consequence of the assumption that drivers are homogenous and consistent in the combination with the better description of headways in major stream flows. Besides, numerous traffic investigations have to be conducted for different traffic and road conditions in order to determinate the proper values of parameter a.

Therefore, in Croatia, it is the simplest to use the Harders model where the unique model parameter is traffic flow q, which is the commonly measured parameter. The capacity and delay model which describe interaction of more conflicting flows using impedance effect and weighting factors for conflicting flows (using negative exponential distribution of headways) showed good estimates of capacity and delay. Also, the possibility of using parameters from HCM methodology for description of prevailing road and traffic conditions in Croatia has been shown.

Finally, based on the consideration of various models characteristics and the performed tests it can be concluded that the capacity and delay procedures according to HCM 2000 methodology give fairly good results for the intersection with normal (common) traffic and road conditions.

Dr. sc. **DRAŽEN CVITANIĆ**
e-mail: drazen.cvitanic@gradst.hr
Mr. sc. **DEANA BREŠKI**
e-mail: deana.breski@gradst.hr
Sveučilište u Splitu, Građevinsko-arhitektonski fakultet
Matice hrvatske 15, 21000 Split, Republika Hrvatska
Mr. sc. **BORIS VIĐAK**
e-mail: boris.vidjak@igh.hr
Institut građevinarstva Hrvatske, Poslovni centar Split
Matice hrvatske 15, 21000 Split, Republika Hrvatska

SAŽETAK

PREGLED, TESTIRANJE I PROVJERA VALJANOSTI MODELA KAPACITETA I ZAKAŠNJENJA NESEMAFORIZIRANIH RASKRIŽJA

U ovom radu razmatraju se postojeći modeli kapaciteta i razine usluge nesemaforiziranih raskrižja. Parametri raznih razvijenih modela procijenjeni su na temelju prikupljenih podataka na cestama u zemljama gdje su modeli razvijeni. U ovom članku se razmatraju dva osnovna problema. Prvo se razmatra pitanje pouzdanosti i primjenjivosti modela općenito, a zatim primjenjivost parametara razvijenih modela za prevladavajuće uvjete prometa i prometnica u Hrvatskoj.

U uvodnom dijelu prikazane su osnove teorije prihvaćanja vremenskih praznina koja predstavlja temelj za razvoj modela kapaciteta i razine uslužnosti (zakašnjenja). Prikazani su različiti izrazi za proračun kapaciteta ovisno o primijenjenoj razdiobi vjerojatnosti vremenskih praznina u prioritetnom toku. Opisane su metode za proračun kapaciteta složenih modela s interakcijom više od dva prometna toka različitog ranga prioriteta. Prikazani su rezultati testiranja različitih modela kapa-

citeta. U drugom dijelu članka dan je pregled postavki teorije repova te razvijenih stacionarnih i vremenskih zavisnih modela zakašnjenja. Prikazane su pretpostavke te pouzdanosti i ograničenja modela. Također je testiran model zakašnjenja koji se koristi u "Highway capacity" manual metodologiji.

KLJUČNE RIJEČI

nesemaforizirana raskrižja, kapacitet, zakašnjenje, razina uslužnosti

LITERATURE

[1] *Pravilnik o osnovnim uvjetima kojima javne ceste izvan naselja i njihovi elementi moraju udovoljavati sa stajališta sigurnosti prometa*, N. N. br. 110, 2001.

[2] *Highway Capacity Manual* (HCM 1985). Transportation Research Board, Special Report 209, Washington, DC.

[3] *Highway Capacity Manual* (HCM 1994). Transportation Research Board, Washington DC.

[4] *Highway Capacity Manual* (HCM 2000). Transportation Research Board, Washington DC.

[5] **Abou-Henaidy, M., S. Teply,** and **J. D. Hunt**: "*Gap Acceptance Investigations in Canada*", Proceedings of the Second International Symposium on Highway Capacity. Australian Road Research Board- Transportation Research Board, Vol. 1, 1997, pp. 1-20.

[6] **Ashworth, R. & Bottom C. G.**: "*Some observations of Driver gap Acceptance Behaviour at a Priority Intersection*", Traffic Engineering & Control, Vol. 18(12), 1997.

[7] **Kittelson, W. K.** and **Vandehey, M. A.**: "*Delay effect on driver gap acceptance characteristics at two-way stop-controlled intersections*", Transportation Research Record 1320. Transportation Research Board, Washington DC, 1991.

[8] **Kyte, M. Tian, Mir, Hameedmansoor, Kittelson, Vandehey, Robinson, Brilon, Bondzio, Wu, Troutbeck**: "*Capacity and Level of Service at Unsignalized Intersections, Final Report, Volume 1 - Two way Stop-Controlled Intersections*", Final Report for National Cooperative Highway Research Program Project 3-46, 1996.

[9] **Brilon, W.** and **M. Grossmann**: "*The New German Guideline for Capacity of Unsignalized Intersections, : Intersections without Traffic Signals II* (Ed.: W. Brilon), Springer Publications, Berlin, 1991.

[10] **Hansson, A.** and **T. Bergh**: "*A New Swedish Capacity Manual/CAPCAL*", Proceedings 14th Australian Road Research Board Conference, Vol. 14(2), 1988, pp. 38-47.

[11] **Cowan, R. J.**: "*Useful Headway Models*", Transportation Research, 9(6), 1975, pp. 371-375.

[12] **Drew, D. R.**: "*Traffic Flow Theory and Control*", McGraw-Hill Book Company, New York, 1968.

[13] **Buckley, D. J.**: "*Road Traffic Headway Distribution*", Proceedings, 1st ARRB Conf., Vol. 1(1), 1962, pp. 153-186.

[14] **Harders, J.**: "*The Capacity of Unsignalized Urban Intersection*", Schriftenreihe Strassenbau und Strassenverkehrstechnik, Vol. 76, 1968.

[15] **Siegloch, W.**: "*Capacity Calculations for Unsignalized Intersections*", Schriftenreihe Strassenbau und Strassenverkehrstechnik, Vol. 1, 1974.

[16] **Cvitanić, D., Lozić, I.**: "*Modeli kapaciteta nesemaforiziranih raskrižja*", Ceste i mostovi, broj 7-8, srpanj-kolovoz 2002., str. 167-176, izvorni znanstveni članak..

[17] **Kyte, M.**: "*Estimating Capacity and Delay at an All-Way Stop-Controlled Intersection*", Research Report, TRANSNOW, Moscow/Idaho, 1989.

[18] **Troutbeck, R. J.**: "*Estimating the Critical Acceptance Gap from Traffic Movements*, Research Report, 92-5, 1992.

[19] **Haight, F. A.**: "*Mathematical theories of traffic flow*", Academic Press, London, 1963.

[20] **Troutbeck. R. J.**: "*Average Delay at an Unsignalized Intersection with Two Major Streams Each Having a Dichotomized Headway Distribution*", Transportation Science, 20(4), 1986, pp. 272-286.

[21] **Newell, G. F.**: "*Applications of Queueing Theory*", 2nd Ed. Chapman and Hall Ltd., London, 1982.

[22] **Kimber, R. M.** and **E. M. Hollis**: "*Peak Period Traffic Delays at Road Junctions and other Bottlenecks*". Traffic Engineering & Control. Vol. 19(10), 1978.

[23] **Akcelik, R.** and **R. J. Troutbeck**: "*Implementation of the Australian Roundabout Analysis Method in SIDRA*", Proceedings of the International Symposium on Highway Capacity U. Brannolte (ed.) Highway Capacity and Level of Service, Karlsruhe, A. A. Balkema, Rotterdam, 1991, pp. 17-34.