A review of critical gap estimation approaches at uncontrolled intersection in case of heterogeneous traffic conditions

Harsh Jigishbhai Amin+; Akhilesh Kumar Maurya

Indian Institute of Technology Guwahati, Guwahati, India

Article Info	Abstract
	The gap acceptance study is an indispensable study in the determination of capacity and delay of the individual moment at uncontrolled intersection and critical gap is one of the chief parameter allied with this study. Most of the studies related to critical gap estimation have been carried out in developed countries where traffic is homogeneous and rules of priorities as well as lane disciplines are voluntarily followed. However, in India, priority rules are less honored which consequently create more clashes on intersections and modelling of such behavior becomes complex. Even though very limited studies have been reported in such case. This paper presents the study of gap acceptance behavior of drivers on an uncontrolled intersection. In order to examine the gap acceptance behavior of drivers, 4-legged uncontrolled intersection located in a semi-urban area of Ahmedabad in the state of Gujarat was selected. Video recording technique was adopted for collection and analysis of the data. Different existing methods are used to calculate the critical gap for through and right turning movements from the minor road. Gaps are analyzed for three different vehicle types. A comparison between different existing methods of critical gap determination is carried out and their critical analysis is presented. The result shows, except clearing behavior approach, all other methods are failing to represent the actual scenario of the critical gap under mixed traffic conditions as these all methods had been developed for homogeneous traffic conditions.

+ Corresponding author. Indian Institute of Technology Guwahati, Guwahati-781039, India.
E-mail address: aminharshj@gmail.com.

Introduction

The efficiency of the road network is elucidated in terms of safety of the driver, average speed as well as the capacity of the road section. The overall traffic flow of the complete transport system basically depends on the performance of individual intersection. The capacity of two-way stop controlled (TWSC) intersection plays the chief role in the assessment of road network capacity. In the case of semi-urban and urban scenario, TWSC intersections are the elementary locations for traffic hazards, delay and congestions. Accounting, financial criteria, installation of signals becomes infeasible at all unsignalized intersections to diminish these effects. Thus, in India, traffic signals are installed solely on major roads. However, in a country, which is already developed, TWSC intersections are typically restricted by yield and stop sign boards that fix up the priority of assorted maneuvering. Effective application of priority rule has made it doable to cross the intersection with least clashes. On the contrary, in India the majority of intersection doesn't have those sign boards. Though they exist, drivers do not follow those designated signs and enter into the intersection's region even on-coming vehicle is about to crash. All these non-standard circumstances fabricate very complex travel behavior at TWSC intersection. The majority of the movement on these types of intersections are uncontrolled, therefore, this form of intersections are tagged as uncontrolled intersections in India. In spite of those difficulties on a non - standard intersection, very limited number of studies have been reported. This paper attempts to put the attention on some of the important methodologies and their suitability to the Indian conditions.

The capacity and service times at minor streets of uncontrolled intersections rely upon the possibilities to possess enough gap between vehicles of the upper prioritized streams to cross the conflict spaces securely. These possibilities may be a function of vehicle flow rate on the major streams, individual drivers' as well as vehicle characteristics that illustrate each individual gap acceptance behavior. Before beginning the introduction of the critical gap, the term of gaps ought to be illuminated. Highway capacity manual (HCM, 1985, 1994, 2000) described it as the time, in seconds, from the front bumper of the second of two successive vehicles to reach the starting point of the front bumper of the first. Although many studies have considered both gap and headway as the same, headway is the time interval between front bumpers of the vehicle to the front bumpers of following a vehicle. In conjunction with the term of gap comes another term lag. It is defined as the time interval between the arrivals of vehicles at a stop line of minor road and the arrivals of the first vehicle at upstream side of the conflict zone (Polus, 1983). Gap acceptance is vital objective characteristic of the performance of the driver and its relevancy in studies regarding capacities, delays, and road safety at unsignalized intersections. HCM (2000) defined the gap acceptance as the process by which a minor street vehicle accepts an available gap in conflict stream to complete his/her maneuver. Critical gap is one of the foremost necessary parameters associated with this study. It is defined as the minimum time between successive major stream vehicles, in which minor street vehicle can make a maneuver (HCM 2000).

Most theories related to gap acceptance behavior of drivers presumed that drivers are consistent and uniform which are employed for TWSC intersections. A consistent driver is anticipated to act the same manner at all moments and at all similar scenarios. The driver isn't anticipated to discard a gap and then afterward accept a smaller gap. For an uniform population, all drivers are anticipated to act in exactly the similar way. It is, definitely, unreasonable to expect drivers to be consistent and uniform. These presumptions of drivers being both consistent and uniform for either approach are obviously not truthful. Catchpole and Plank (1986) and Troutbeck (1988) have recommended that the entry capacity would reduce if drivers are heterogeneous. However, if drivers are consistent, then the capacity would be inflated. If drivers are presumed to be both more uniform and consistent together, rather than more realistically inconsistent and heterogeneous, than the changes in the forecast would be negligible. As a result, the overall impact of these presumptions is marginal and, for easiness, consistent and uniform driver behavior is presumed in most theory.

There is some inconsistency in the outcomes of numerous research associating to gap and lag acceptance. Miller (1972) examined lag and gap as one crowd data and pointed that the analysis of lag is less reliable than that of the gap. Daganzo (1981) pointed that the mean critical gap was significantly shorter than the mean critical lag. Gattis and Sony (1999) noticed majority of the critical lag values were lesser in comparison with critical gap value. However, the capacity approximation equation for uncontrolled intersection given by HCM (2000) admitting gap and lag as a same group data. A research by Sturr and Taub (1990) advised that 65 years are a crucial period after which eyesight becomes progressively poor under circumstances of degraded illumination. However, report on median gap acceptance across all age terms in regard to night and day is inconsistent. Kyte et al. (1994) found that the critical gap does not fluctuate with approaching traffic speed. However, in Alexander et al. (2002) observed that the speed of the conflicting traffic was the variable that has the most significant impact on the mean accepted gap size. Further results inferred that the calculation of opportunities for smashes were less precise for aged drivers (40–64 years) in comparison with younger drivers (18–29 years) (DeLucia et. al., 2003). Aged drivers were having issues to recognize, understand and decide the safety of a gap as compared to younger one (Laberge et. al., 2006). Therefore, aged drivers may suffer more difficulties at uncontrolled intersections due to abated visual capabilities, which involves depth perception and motion sensitivity. An evaluation of the critical gap in Indian condition is complex due to the lack of lane discipline, non-adherence to priority and traffic rules, different vehicular characteristics, improper parking, inappropriate geometric design of the junction and improper road marking. However, very few of the researchers conducted their studies on uncontrolled intersection that are prevalent in India.

This paper is organized in three sections excluding Introduction and Conclusion. Procedures of data collection and extraction are described in Section 1. 2nd Section discusses nine different estimation methodologies. Further, Section 3 gives the discussion on the results. And the conclusion of the paper along with directions for further research is presented in last.

1. Data collection

This paper presents a study of gap acceptance behavior of the driver on uncontrolled intersection. In order to analyze the gap acceptance behavior of the driver, the four-legged uncontrolled intersection located in semi-urban area of Ahmedabad in the state of Gujarat was selected. This intersection was located in plain terrain with adequate sight distances for each movement. There were no bus stands and parking nearby this intersection. Incoming or outgoing vehicle behavior was not impacted by the upstream/downstream signals. The total traffic volume on major road varies 1300 to 2400 vehicles per hour and on minor stream it varies from 700 to 1400 vehicles per hour. Data were collected during the peak hours of a typical weekday (9 A.M to 12 A.M). Video recording technique was adopted for data collection and five video cameras had been used simultaneously to collect the parameters like vehicle arrival rate/time, accepted and rejected gap/lag time, speed and type of conflict vehicles, waiting time of minor street vehicle at stop line, driver's age as well as gender, occupancy, etc. Out of five, three video cameras were placed on the terrace of the building situated near the intersection for gathering the various vehicle movements on different approaches of this intersection, accepted and rejected gap/lag time, speed and conflicting vehicle type, waiting time, etc. Moreover, two cameras placed on the shoulder and median of the minor road by using 5 feet high tripods to gather the age as well as gender of the drivers of minor road. Figure 1 (a) & (b) shows the locations of all cameras and birds view of the selected site respectively. The recorded tape was replayed on the big screen to extract and analyze the gap acceptance data. All minor road vehicles were divided into three categories (car, 2-wheeler & 3-wheeler) and the gap data were analyzed vehicle category wise for two lower priority movements; through and a right turn from minor stream.
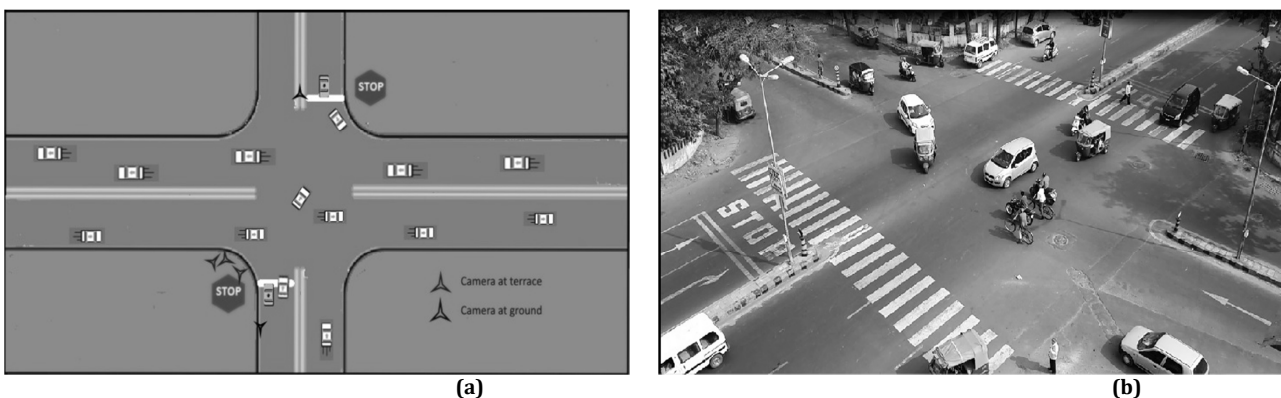

(a)   **(b)**

Figure 1 – (a) Locations of five video cameras on the selected site, (b) Birds view of the selected site located at Ahmedabad

2. Estimation methods for critical gap

2.1. Raff method

Raff method (1950) is based on macroscopic model, and it is the earliest methodology for estimating the critical gap which is employed in several countries due to its simplicity. According to Raff method, a critical gap is the time at the sum of the cumulative number of accepted gaps F_a and rejected gaps F_r is equal to 1. The original Raff theory uses solely lag data. Therefore, this approach thought of as statistically wasteful by some previous literatures. There are two ways to remedy this disadvantage. First is combining the gap and lag data based on the notion that there is no statistical significance between lag and gap. Second, as an alternative approach is to detach the lag and gap data into ''lag-only'' and "gap-only" curves (Miller, 1972; Fitzpatrick, 1991). The only disadvantage attributed to the present methodology is its sensitiveness to the traffic volumes under which they have been evaluated.

2.2. Greenshield method

Greenshield has introduced a bar chart to represent the total number of rejections and acceptances for each range of gap. The horizontal axis represents the range of gap size, in seconds, and the vertical axis of the bar chart represents the number of accepted and rejected gaps by positive or negative values respectively for a certain gap-range. The range of the critical gap size is the gap range having an equal number of acceptance and rejection, and the mean of this range is a critical gap. The result obtained with this methodology is very stricken by traffic volume of the major street and little sample size might distort the outcomes (Gattis & Low, 1999).

2.3. The lag method

An estimation of the critical gap from traffic observations with under-saturated conditions is highly complicated. In this situation, one simple approach could be based on lags. A lag is the time from the arrival of the minor vehicle at stop line until the arrival of the next major vehicle at the upstream face of the conflict point. Drivers and results are assumed to be a consistent and independence of the minor street vehicle arrival time or the traffic situation on the major street respectively. The downside attributed to the current technique is its observation periods as for each interval, a sufficiently giant sample ought to be needed. This technique needed terribly long observation periods because with low major street traffic flow it takes a while to enough smaller lags, and with large major street volume most minor street vehicles have to queue before they can enter the conflict zone. Afterwards, despite a huge number of drivers' decisions have been analyzed, there will be only few samples that may be used for this estimation procedure. An additional problem could be that the critical value for the lags might be systematically different from that for the gaps (Brilon, Koenig, & Troutbeck, 1999).

2.4. Harder's method

Harder has developed a method in 1968 and became more popular in Germany. This method is almost similar to lag method; however it takes solely gap time, whereas lag methodology uses lag time. Same type of assumptions is needed as those mentioned for lag methodology for practical applications. The major drawback of this methodology is the curve, which has real properties of cumulative distribution function of the critical gap, generating from this method may not be gradually increasing over the time or it may float. Therefore, to correct these values floating average procedure has to be adopted, or large sample size is needed to avoid this impact (Brilon, Koenig, & Troutbeck, 1999).

2.5. Acceptance curve method

From empirical and theoretical considerations, if the dependent variable is a binary variable, the shape of the response function will be curvilinear. The dependent variables of this response curve are the cumulative probability of accepting a gap of a specific length. The x-value corresponding to the 0.5 probabilities may be defined as a critical gap size. The main downside of this methodology is that the development of acceptance curve bias (or lag acceptance bias) produces a more or less distorted. This bias is introduced when data from drivers that reject multiple gaps are included. Drivers who want long gaps will often reject the lag and a number of other gaps before getting an appropriate gap, whereas drivers with low acceptance thresholds are more likely to accept the first gap offered to them. Considering all accepted and rejected gaps will produce a gap acceptance curve in which the percentage acceptance of a given gap size will be somewhat less than the percentage of minor street drivers prepared to accept a gap of that size. The resulting impact of this bias is that the reported essential gap is somewhat larger than the actual essential gap (Gattis & Low, 1999). Another downside is that if the data samples are less or if the number of time intervals are not stuffed with sufficient empirical values, accumulative curve might not be potential or could also be floated. Thus to avoid this impact higher gap interval ought to be needed to plot the curve and consequently, the results would be inaccurate.

2.6. Logit method

A couple of approaches are developed that may be summarized as logit models as they provide similarities to the normal logit models of transportation planning. In each case the models cause a function of the logit type. Logit model is basically a weighted linear regression model with a mathematical form as shown in Equation (1).

$$P(a) = (1 + e^{-(\beta_0 + \beta_1(i))})^{-1} \qquad (1)$$

Where $P(a)$ is the probability of accepting a gap of size i; β_0 and β_1 are regression coefficients. A fitted linear line can be plotted on the chart to see the time that offers 0.5 probabilities of acceptance of a gap size. This technique is often accustomed to verify the influence of different independent attributes in the critical headway, such as the waiting time, Avg. speed, etc. It ought to be noted that with this formulation a driver who accepts a gap smaller than one previously rejected cannot be outlined as "inconsistent," since the other independent variables will make a case for that behavior. Therefore, all gaps ought to be needed, not only the maximum rejected by each driver (Brilon, Koenig, & Troutbeck, 1999). The outcome obtained by this technique incorporates a sturdy dependency on major street volume (National Research Council, 1996).

2.7. Ashworth's method

A series of accepted gap is delineated by empirical distribution functions, however, the distribution function of the critical gap ought to be laid on the left side of the distribution function of the series of accepted gaps. By assuming normal distributions for t_c and t_f, Ashworth (1968) found that the average critical gap may be evaluated from mean of accepted gap $E(t_a)$. If the mean and variance of this distribution are E(ta) and Var(ta), then Ashworth's technique provides the critical gap based on Equation 2.

$$E(t_c) = E(t_a) - q_p Var(t_a) \tag{2}$$

Where q_p is the main stream flow in units of the vehicle per second and $E(t_c)$ denotes the mean of the critical gap (National Research Council, 1996). This is a very sensible explanation, and that can be used to give satisfactory outcomes in the office or the field. The major drawback of this methodology is that critical gap obtained by this method is highly correlated with major street traffic volume.

2.8. Wu's model

The theoretical background of Wu´s model (2006) is the probability equilibrium between the rejected and the accepted gaps. The main advantage of this model is, it provides a true average of the critical headway, and it does not need any predefined distribution function of critical gap and assumptions relate to the consistency and homogeneity of drivers whereas MLM needed predefined distribution function as well as these assumptions. This method yields the empirical distribution of the critical gaps, which can be helpful for microscopic simulation. Wu model produces similar results for the mean critical gap as obtained from Troutbeck MLM when using maximum rejected gap instead of all rejected gaps and if considering all rejected gaps, mean critical gaps would be shorter than previous results. Still, it has a limitation: in the observation data, the minimum accepted gap should be smaller than the utmost rejected gap. This can be happening only just in case of little sample size.

2.9. Clearing behavior approach

This unique methodology is given by Ashalatha and Chandra (2011), and this is the only approach which was developed for the condition relevant to Indian. They have taken clearing time to integrate the clearing behavior of the driver and make it differs from other approaches. Clearing time is defined as the time taken by a lower priority vehicle from the instant it moves from its stopped position to the instant its tail end reach the end of the conflict area. They defined the term of the critical gap as the gap corresponding to the intersection point of cumulative curves of gap acceptance and clearing time. This intersection point provides a situation when clearing time and available gap to a driver are just equal. This is the absolute minimum time needed by driver to enter the intersection safely. Researchers have presented this new concept with strong theoretical backgrounds. However, still, this method has some limitations. A critical gap is not only the function of accepted gaps, but it is the function of rejected gap too, and that is proved by ample empirical and practical researches. This phenomena will produce a higher estimation of the critical gap as compared to actual critical gap.

3. Discussion of results

The nine approaches represented above were used to determine the critical gap for through and right turns from a minor road (two wheelers, three wheelers and cars) at four-legged intersection located at Ahmedabad. The result of the analysis is given in Table 1. This result shows many variations in the critical gap parameters estimated by different methodology. The values are also quite low in the majority of cases. Though in heterogeneous traffic condition, this parameter is expected to be on the lower side but not as like this. As a lack of firm lane discipline and with wide variation in sizes of numerous types of vehicle, the vehicles are found to stand side by side. Smaller size vehicles can squeeze through any offered gap between the large size vehicles and enter into the intersecting space in a random manner. Even while accepting a gap, vehicles are found to enter the conflicting space abreast and after crossing the on-coming traffic they move as a single. In extreme cases of messy behavior, small sized vehicle like a 2-wheeler from a minor street may be found moving opposite to the mainline traffic slowly and exiting through a smaller gap available upstream of the intersection space. These distinctive and unusual features of mixed traffic make the gap acceptance quite diverse from what have been observed under homogeneous traffic conditions and this impact shows in Table 1 where a majority of methods are failing to represent the actual critical gap. Except clearing behavior approach, all other methodologies are established under the developed countries to analyze the homogeneous traffic behavior.

Table 1 - Comparisons of the critical gap (in seconds) estimated by nine important methods

	Vehicle Types	Raff	Wu	Logit	Ashworth	Lag	Harder	Acceptance curve	Clearing behavior	Greenshield
Through movement	2-Wh	2.30	2.55	2.70	2.45	2.60	2.70	2.40	4.80	2.25
	Car	2.65	2.97	3.10	2.95	3.04	3.20	2.80	5.00	2.50
	3-Wh	2.25	2.58	2.90	2.55	2.80	2.75	2.45	4.65	2.50
Right turning movement	2-Wh	2.20	2.45	2.60	2.30	2.45	2.50	2.35	4.70	2.25
	Car	2.55	2.70	3.10	2.80	2.95	3.10	2.75	4.95	2.75
	3-Wh	2.20	2.40	2.54	2.30	2.50	2.45	2.63	4.50	2.25

Therefore, results obtained by these approaches are erroneous and thus cannot apply under the conditions relevant to Indian. As compared to through movements from a minor road, right turning movements from a minor road were more chaotic as majority of the vehicle was crossing the intersection area in a zig-zag manner. Careful analysis of the field situation exposed that more than 60% of right turning from minor stream vehicles forced opposing vehicles to slow down.

Hence, major vehicle has to wait up to the time when the movements of lower priorities are clear. As a result, in the majority of case, critical gaps obtained for right turning movements are quite less as compared to through movements. In addition, careful analysis of the field situation revealed that the majority of vehicle accepts the gap in two stages in India. In the majority of cases, drivers were not bothering about the gaps in the far lane when they are exploring the gaps in the near lane. Once, they accept the gaps in the near lane and reach up to the median; they start searching the suitable gaps in the far lane. This condition creates more chaos at median separated uncontrolled intersection. However, this two stage gap acceptance process increases the entry capacity of an unsignalized intersection but only when there is sufficient storage space for one or more vehicles at center median.

Conclusion

This paper presents a detailed review of studies carried out in the past related to estimation of the critical gap. Further, this study also highlighted the difficulty in measuring critical gap on uncontrolled with heterogeneous and weak lane discipline traffic which is observed in India. In order to demonstrate the inefficiency of developed methodologies for critical gap estimation at Indian uncontrolled intersection, traffic movement data is collected from an uncontrolled four-legged intersection located in semi-urban area of Ahmedabad in the state of Gujarat. The critical gaps are calculated for three different vehicle types and two different maneuvers using nine different methodologies developed in the past. Estimated critical gaps are found to be relatively low with a wide deviation among the results obtained by these nine different approaches. This is due to the inborn fault of these methodologies to account for the heterogeneous traffic conditions as they evolved under fully homogeneity. However, concept of clearing behavior approach which has been evolved for mixed traffic conditions produces quite reasonable results for both movements. This approach takes the concept of clearing time to incorporate the actual driving behavior of drivers. This estimation procedure is quite simple, and it is applicable for mixed traffic condition as well as homogeneous traffic and similarly it is applicable for both over-saturated as well as under-saturated traffic conditions. The result obtained with this methodology is not highly affected by traffic volume on Major Street as cumulative distribution curve of clearing time and accepted gap act in a contradictory manner. Hence, the overall impact of traffic volume on the result is minimal. The estimation of critical gaps is not an end in itself. Critical gaps are used in capacities and delay estimation models. The preciseness of capacity estimation is largely determined by accurate estimation of critical gap parameters. Except clearing behavior approach, all other methods are really failing to represent the realistic value of the critical gap at four legged median separated uncontrolled intersections. Thus, this technique will assist as a simple tool of the working engineers and researchers in determining different tactics under heterogeneous traffic conditions.

The present study was taken up with the objective of initiating research on uncontrolled intersections in India. Very limited concerted efforts have been made earlier to analyze actual traffic flow at four leg uncontrolled intersection under mixed traffic situations. The critical gap suggested here are for three categories of vehicles only and a particular geometry of the intersection. Further research may be conducted for estimation of the impact of various parameters (like occupancy, age of the driver, sex of the driver, speed of on-coming traffic, size of on-coming vehicles, number of rejections etc.) on critical gap parameter, especially in the case of mixed traffic conditions.

References

Alexander, J., Barham, P., & Black, I. (2002). Factors influencing the probability of an incident at a junction: results from an interactive driving simulator. Accident Analysis & Prevention, 34(6), 779–792.

Ashalatha, R., & Chandra, S. (2011). Critical gap through clearing behavior at unsignalised intersection. Ksce Journal of Civil Engineering, 15(8), 1424-1434.

Brilon, W., Koenig, R., & Troutbeck, R. J. (1999). Useful estimation procedures for critical gaps. Transportation Research, 33(A), 161-186.

Catchpole, E. A., & Plank, A. W. (1986). The Capacity of a Priority Intersection. Transportation Research Board, 20B(6), 441-456.

Daganzo, C. (1981). Estimation of gap acceptance parameters within and across the population from direct roadside observation. Transportation Research, 15(B), 1-15.

DeLucia, P. R., Bleckley, M., Meyer, L. E., & Bush, J. M. (2003). Judgments about collision in younger and older drivers. Transp. Res. Part F: Traffic Psychol. Behavior, 6, 63-80.

Fitzpatrick, K. (1991). Gaps accepted at stop-controlled intersections. Transportation Research Record, 1303, 103-122.

Gattis, J. L., & Low, S. T. (1999). Gap acceptance at typical stop controlled intersection. Journal of Transportation Engineering, 125(3), 201-207.

HCM. (1985, 1994, 2000). Em Highway Capacity Manual, SR 209. Transportation Research Board, National Research Counsil, Washington D.C.

Kyte, M., Kittelson, W., Tian, Z., Brilon, W., Troutbeck, R., & Mir, S. (1994). New measurements for saturation headways and critical gaps at stop-controlled intersections. Proc. 2nd Int. Symposium on Hwy. Capacity at Sydney, Australian Road Research Board, 345-364.

Laberge, J. C., Creaser, J. I., Rakauskas, M. E., & Ward, N. J. (2006). Design of an intersection decision support (IDS) interface to reduce crashes at rural stop-controlled intersection. Transp. Res. Part C: Emerg. Technol, 14, 36-56.

Miller, A. J. (1972). Nine estimators of gap-acceptance parameters, In: Traffic Flow and Transportation. Newell, G. F. (Ed), American Elsevier Publ. Co, Inc., New York, 215-235.

National Research Council. (1996). NCHRP Web Doc 5 Capacity and Level of Service at Unsignalized Intersections: Final Report Volume 1 - Two-Way-Stop-Controlled Intersections. Washington, DC: The National Academies Press.

Polus, A. (1983). Gap acceptance characteristics at unsignalised urban intersection. Traffic Engineering and Control, 24(5), 255-258.

Raff, M. S., & Hart, J. W. (1950). A volume warrent for urban stop signs. Eno Foundation for Highway Traffic Control. Saugatuck, Connecticut.

Sturr, J. F., & Taub, H. A. (1990). Performance of young and older drivers on a static acuity test under photic and mesopic luminance conditions. Human Factors, 32(1), 1-8.

Troutbeck, R. J. (1988). Current and future Australian practices for the design of unsignalized intersections. In: Intersections without Traffic Signals, Springer Publications, Berlin.

Wu, N. (2006). A new model for estimating critical gap and its distribution at unsignalised intersection based on the equilibrium of probabilities. Proceeding of the 5th International Symposium on Highway Capacity and Quality of Service. Yokohama, Japan.