NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

UNIVERSITY OF ALBERTA

MODIFIED WHITING DELAY MODEL FOR SIGNALIZED INTERSECTIONS

BY

HUNG-CHUAN KUA

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN TRANSPORTATION ENGINEERING

DEPARTMENT OF CIVIL ENGINEERING

EDMONTON, ALBERTA

FALL 1990

ISBN 0-315-65046-X

University of Alberta
Edmonton
Canada T6G 2G7

Department of Civil Engineering

220 Civil/Electrical Engineering Building,
Telephone (403) 492-4235
Fax (403) 492-0249

October 11, 1990

Faculty of Graduate Studies and Research
University of Alberta
2-8 University Hall
Edmonton, Alberta

To Whom It May Concern

Copyright Release Letter

I hereby authorized Hung-Chuan Kua to reproduce figures from the Canadian Capacity Guide for Signalized Intersections for the purpose of his Master of Science thesis.

Sincerely,

Stan Teply
Editor
Canadian Capacity Guide for
Signalized Intersections

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR : HUNG-CHUAN KUA

TITLE OF THESIS : MODIFIED WHITING DELAY MODEL FOR SIGNALIZED INTERSECTIONS

DEGREE : MASTER OF SCIENCE

YEAR THIS DEGREE GRANTED : 1990

....

2, SOLOK TEMBAGA TIGA
GREEN LANE, PENANG 11600
MALAYSIA

Date: Oct. 11, 90

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

THE UNDERSIGNED CERTIFY THAT THEY HAVE READ, AND RECOMMENDED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH FOR ACCEPTANCE, A THESIS ENTITLED MODIFIED WHITING DELAY MODEL FOR SIGNALIZED INTERSECTIONS SUBMITTED BY HUNG-CHUAN KUA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

(S. Teply, Supervisor)

(J.J. Bakker)

(W.J. Sproule)

(J.M. Whiting)

Date: Oct. 11, 90

To my wife, Helen

ABSTRACT

This thesis examines the current analytical techniques for delay assessment at signalized intersections in Canada, Australia and the United States. The thesis discusses the reliability of these techniques, with particular focus on the Canadian model (Whiting delay equation) and on the inter-relationships of its parameters. The implications of using an appropriate evaluation time parameter for more accurate delay predictions, and in turn the development of a method to determine such parameter values and an alternate model of providing those delay predictions are discussed. The relevant information for this study comes from the literature, transportation reports, existing models, on-site surveys and computer simulations.

The review of the Whiting delay model uncovers three important points: (1) the random overflow delay component of the model cannot be simultaneously random and time-dependent as model results indicate, (2) the absolute value of the signal cycle time influences the occurrence of random overflow, and (3) the prediction of expected delay is more useful than that of the delay experienced. In addition to the derivation of an evaluation time formula, which better reflects the congestion period, an alternate random overflow delay model is developed to replace that in the Whiting equation.

The two equations produce comparable results. However, the alternate approach seems to explain the predicted delay in easier and more understandable terms.

EXECUTIVE SUMMARY

INTRODUCTION

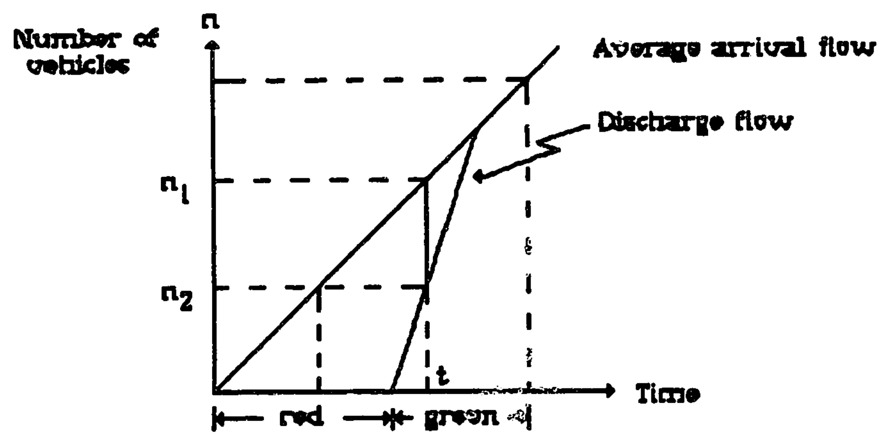
The main objective of this research was to review the Whiting overflow delay equation, and to determine the relationship between the evaluation time and the congestion period. After a detailed examination of the equation, the research was expanded to develop a better approach to predict random overflow delay.

This thesis describes the findings of the research. The thesis begins by explaining the theories of delay and the current methods for analytical signalized intersection delay assessments. The derivations and limitations are examined to identify the influential parameters of the overflow delay. The thesis then describes the Whiting overflow delay equation and its related questions. The relationship between the evaluation time and the congestion period is revealed through the derivation of an equation. It is followed by discussions of the development and verification of the new overflow delay prediction model.

The relevant information for this study comes from the literature, transportation reports, reviews of delay theories and existing models, queueing diagrams, field measurements and computer simulations.

THEORY OF DELAY

Delay is a useful tool in evaluating signal operations. Delay is defined as the additional time a driver has to spend at an intersection compared to the time in which there is unimpeded access. It may be caused by either the traffic signal or any other traffic. Generally, delay is determined by means of a queueing model, or queueing diagram:



A queueing diagram showing a queueing situation within a signal cycle of a lane of traffic at a signalized intersection

The difference between the number of vehicles arrived and the number of vehicles discharged in time interval t, is the number of vehicles in queue at time t. The time difference between the arrival and the discharge of a vehicle, will be the delay experienced by the vehicle. Therefore, the total delay is equivalent to the area enclosed by the average arrival and discharge curves, the average delay is equal to the total delay divided by the total number of cars discharged in a single cycle.

There are two major techniques for delay assessments: field measurements and analytical estimates. The latter method is more convenient and popular because delay predictions are possible even when signal systems are still in design stage, as the technique does not require the actual existence of a signal system.

DELAY PREDICTION MODELS

Three types of delay are usually estimated in analytical prediction of delay. They are:

1. Uniform delay - this type of delay results from the interruption of traffic flow by traffic signals.
2. Continuous overflow delay - experienced when the arrival flow is consistently greater than the capacity of the lane. A growing queue forms.
3. Random overflow delay - caused by fluctuations in vehicle arrivals. It may occur during some cycles under generally uncongested operations.

Although each of the three types of delay is significant within a certain range of degree of saturation, all three of them occur over the entire range of degree of saturation. Degree of saturation is the ratio of the travel demand to the travel supply. A total overall delay is the sum of all three types of delay. The average overall delay is the total overall delay divided by the number of vehicles passing through the intersection during the analysis period.

Three overall delay models have been developed and are currently used in many countries. They are named after their countries of application: the Canadian, Australian and American Delay Models.

In Canada, Canadian Capacity Guide for Signalized Intersections recommends the Whiting delay equation as the analytical delay estimation technique. This equation consists of a steady state uniform delay and a time-dependent overflow delay model. By introducing an empirical modifier into the overflow delay term, whiting developedan equation that accounts for all three types of delay. The introduction of a simple random overflow delay modifier is a major contribution of Whiting formula, since for low and high degree of saturation the unifrom and overflow models work well.

Unlike the Canadian formula, the Australian delay equation has a minimum degree of saturation before it applies the overflow delay term. The minimum degree of saturation is determined based upon the capacity per cycle of the traffic lane. The Australian formula also incorporates the effect of the absolute signal cycle time has on the uniform and random overflow delays and places much more emphasis on the random overflow delay than other models.

The American delay equation has a tendency of increasingly overpredicting delays in oversaturated conditions. The American model predicts stopped delay rather than overall delay. The difference between the two is that the stopped delay disregards the acceleration and deceleration delays. The model's assumption of stopped delay is always 77 percent of the overall delay is questionable. The correlation between the two is a function of the red interval and the deceleration delay, instead of just a constant. The American delay equation also assumes a fixed evaluation time of 15 minutes, regardless of the actual congestion period. This assumption of fixed evaluation time restraints the consideration of the effects of congestion period have on the overflow delay. Due to these assumptions, the American model loses its accuracy of predictions as the degree of saturation increases, and it diverges rapidly after the degree of saturation is greater than 1.0. It is recommended that predicted delays should only be used as an indication of the order of magnitude. This recommendation would also apply to the Canadian and Australian models.

The comparison of the three delay prediction models shows:

1. The significance of random overflow delay lies mainly between the degrees of saturation of 0.85 and 1.05.
2. The difference between the Canadian and Australian equations are negligible.

3. The three models are actually variations of one another, and can be generalized in one single equation.

REVIEW OF WHITING OVERFLOW DELAY EQUATION

The Whiting delay equation appears to be a good representation of actual traffic conditions. However, after further examination of the derivation principles of the Whiting delay equation, several major findings emerge:

1. There is a lack of definition of what, exactly, the evaluation time should be, and how it is related to the congestion period; all that is known is merely the evaluation time should reflect the congestion period. Evaluation time is the duration for a delay analysis. With continuous overflow delay being directly proportional to the evaluation time, the accuracy of delay predictions is dependent on the usage of the correct evaluation time.
2. The average continuous overflow delay predicted is the delay experienced by the vehicles discharged within the evaluation period. It seems to be of lesser value to predict delays of those discharged vehicles because delay already experienced has no use to the driver in trying to avoid time loss in the first place.

3. It is contradictory for the random overflow delay model of the equation to be both random and time-dependent at the same time.
4. The Whiting random overflow delay model neglects the effect that absolute value of cycle length has on the delay.

Besides the fact that Whiting's derivation uses an evaluation time that is not clearly defined, his random overflow delay equation has no specific theory to support it. Therefore, a different approach to the equation is required, together with a proper correlation of the evaluation time and the congestion period. This would improve the accuracy of delay predictions.

DEVELOPMENT OF NEW OVERFLOW DELAY PREDICTION MODEL (MODIFIED WHITING DELAY EQUATION)

Knowing that random overflow delay is time independent and cycle-length oriented, a model with probabilistic features should be chosen for the new prediction model. Because random arrival is usually represented by a Poisson distribution, it is used as the probabilistic delay prediction model. Applying the distribution, and simulating the conditions of random overflow at a signalized intersection, a model has been developed, which consists of two parts - one to predict the probability of random overflow and one to predict the resulting random overflow delay. Included in the new approach is also the principle of determining the delay expected instead of delay experienced. It appears that the new model is a more realistic representation of the actual traffic conditions.

The applications of the modified delay prediction model are:

1. to only individual fixed-time signalized intersections.
2. for all ranges of degree of saturation. At a low degree of saturation, the delay predicted is close to zero because the probability of overflow is close to zero. At high degree of saturation, the probability of random overflow is constant but is insignificant in comparison to that of the average continuous overflow.

The assumptions for the new modified model are:

1. Random arrival pattern is represented by a Poisson distribution.
2. Uniform discharge pattern is at a rate equal to the saturation flow.
3. Four vehicles overflow or underflow is used as a practical range.
4. Two consecutive cycles of random overflow are used as a maximum.
5. Vehicles stacking at the stopline.

It should be noted that it was not part of this research to review the uniform delay model. As a result no modifications have been attempted and this component of the modified equation was left as it was originally.

The relationship between the congestion period and the corresponding evaluation time was derived using the geometry of a queueing diagram. The evaluation time is redefined as the flow-persisting time, or the duration in which the arrival flow is consistently greater than the capacity. By applying the principle of conservation of matters and energy, and knowing the arrivals and the duration of the congestion, the evaluation time that allows prediction of the representative delay can be determined properly without any guessing.

VERIFICATION OF THE NEW MODEL

The new model was verified by using field data and computer simulations. Two delay surveys were conducted. One survey was done with regular out-of-step queue count and the other survey was made by means of vehicle trajectory reconstruction. Two sets of computer simulations were also made with a specially designed program.

The predictions obtained from the new model, using the conditions of the field surveys, were compared to those actual field measurements, computer simulated results, and those delays predicted from the Whiting delay equation. The comparisons show that results from the modified model correspond very well to those obtained analytically, as well as field measurements and computer-simulated results. Due to the limited number of field measurements available for further verification of the new equation, no further conclusions can be made at this date.

CONCLUSIONS

The objectives of this research were reached. An equation of how to obtain a proper evaluation time for delay assessment, and a new approach to random overflow delay prediction has been developed. Unlike Whiting overflow delay equation, where the random overflow model was derived empirically, the new model is derived based on a proven mathematical model. The modified model also predicts delay expected.

In conclusion, although limited field data and time did not allow sufficient model testing to prove the new model improves the quality of delay predictions, it can be said that the model presents a useful alternative delay-prediction technique that is mathematically justifiable.

Therefore, it is recommended that the new model should be thoroughly tested, and that it be adopted as another approach to delay predictions.

ACKNOWLEDGEMENT

The author wishes to express his sincere thanks to Professor S. Teply for his excellent supervision and guidance, Dr. W. Sproule, Dr. Whiting and Professor Bakker of the University of Alberta for their valuable suggestions and comments. The author also wishes to thank Dr. H.K. Kua and H. Mak for their assistance in preparation of this thesis.

TABLE OF CONTENTS

CHAPTER			PAGE
1	INTRODUCTION		1
2	DELAY AND TYPES OF DELAY		6
	2.1 Definition of Delay		6
		2.1.1 Vagueness in the Delay Definition	7
	2.2 Theory of Delay.		9
	2.3 Types of Delay		14
3	ANALYTICAL DELAY PREDICTION MODELS		20
	3.1 Basic Models		20
		3.1.1 Uniform Delay Model	21
		3.1.2 Continuous Overflow Delay Model	24
		3.1.3 Random Overflow Delay Model	26
		3.1.4 Comparison of the Models	29
	3.2 Delay Prediction Models: Current Models		32
		3.2.1 Canadian Delay Formula	33
		3.2.2 Australian Delay Model	36
		3.2.3 American Delay Model	38
		3.2.4 Generalized Delay Equation	41
4	WHITING OVERFLOW DELAY EQUATION: A REVIEW		45
	4.1 The Review		45
		4.1.1 Continuous Overflow Delay	45
		4.1.2 Random Overflow Delay	46
		4.1.3 Questions Related to Whiting Delay Equation	46
		4.1.4 Procedures of The Review	47
	4.2 Evaluation Time Versus Congestion Time		48
	4.3 Derivation Principles of Continuous Overflow Delay Equation		52
	4.4 Time Dependency of Random Overflow Delay		54
	4.5 Effects of the Absolute Value of Signal Cycle Time		61
	4.6 The Implications		68

5 MODIFIED WHITING DELAY EQUATION 70
5.1 Uniform Delay Model 70
5.2 Random Overflow Delay Model 70
5.2.1 General Model 70
5.2.2 The Probability Term 73
5.2.3 Expected Delay 79
5.2.4 Final Random Overflow Delay Model 82
5.3 Continuous Overflow Delay 84
5.4 Evaluation Time Versus Congestion Time 86
5.5 The Modified Delay Model 88

6 VERIFICATION OF THE MODIFIED MODEL 90
6.1 Delay Assessment Techniques 90
6.1.1 Computer Simulation 90
6.1.1.1 Program Structure 91
6.1.1.2 Simulation Results 92
6.1.2 Field Measurements 96
6.1.2.1 Vehicle Trajectory Reconstruction 96
6.1.2.2 Stopped Delay 98
6.1.2.3 Results 99
6.2 Discussions of Results 101

7 CONCLUSIONS 104

REFERENCES 107

APPENDIX A: TABLES 109

APPENDIX B: RANDOM OVERFLOW DELAY CALCULATIONS 110

APPENDIX C: LOCATIONS AND GEOMETRIC LAYOUTS OF SURVEYS 111

APPENDIX D: SURVEY RESULTS 114

APPENDIX E: FORTRAN PROGRAM LISTING 115

LISTS OF TABLES

TABLE		PAGE
1	Values of the calibration parameters in Australian, American and Canadian overflow delay model	42
2	Delay predicted from the American, Australian and Canadian delay formulas	43
3	Effects of evaluation time on random overflow delay	55
4	Skewing effects of Poisson distribution	59
5	Effects of cycle time on random overflow delay	67
6	Probability of overflow in consecutive cycles	77
7	Expected extra delay in the first cycle during a random overflow situation	80
8	Delay predictions from computer simulations	94
9	Delay predictions from field measurements	100
10	Delay predictions from various estimation models	101

LISTS OF FIGURES

FIGURE		PAGE
1	Time space diagram illustrating the basic concept of delay	6
2	Time space diagram showing a shockwave at the end of the queue of a lane of traffic at a signalized intersection	8
3	A queueing diagram showing the queueing situation within a signal cycle of a lane of traffic at signalized intersection	12
4	A queueing diagram showing the individual delay of vehicles in a lane of traffic at a signalized intersection	14
5	Time space diagram illustrating uniform and overflow delay	16
6	Queueing diagrams illustrating undersaturated and oversaturated conditions	17
7	A queueing diagram of a traffic lane with undersaturated flow at a signalized intersection	22
8	A queueing diagram of a traffic lane with continuous overflow delay	25
9	Overflow due to variations in the arrival flow in undersaturated conditions	27

10 Comparison of uniform delay model, Webster model and continuous overflow delay model ____ 29

11 The queueing diagram that is the basis of oversaturation overflow delay equation ____ 34

12 Comparison of American, Australian and Canadian delay equation ____ 44

13 Comparison of delays predicted by using Whiting delay equation with evaluation time equal to 60 and 120 minutes ____ 49

14 A flow diagram showing the peaks within a daily traffic flow ____ 51

15 A queueing diagram showing the duration of congestion and the time to the peak of the congestion ____ 52

16 Queueing diagram used for derivation of Whiting continuous overflow delay equation ____ 52

17 Probability curve showing the arrival per cycle at the stopline of a signalized intersection ____ 57

18 Probability curve showing the number of per cycle discharge from the stopline of a signalized intersection 57

19 A queueing diagram showing the effects of cycle time on occurrence of random overflow ____ 61

20 Probability curve showing the arrival per cycle at the stopline of a signalized intersection ____ 63

21 Probability curve showing the discharge per cycle from the stopline of a signalized intersection ________ 63

22 Poisson distribution applied to probability of discharge at signalized intersection ________ 69

23 Probability of arrival for overflowing vehicles ________ 74

24 Significance of individual probability of arrival to total probability of overflow ________ 74

25 A tree diagram illustrating the possible occurrence of random overflow in the first, second and subsequent cycles ________ 78

26 A time space diagram showing the queueing situation at a signalized intersection ________ 80

27 Queueing diagram used for derivation of the continuous overflow delay equation ________ 84

28 Queueing diagram used for relating evaluation time to congestion time ________ 87

29 A flowchart illustrating the logic of the simulation program ________ 93

30 Delay predicted using computer simulation ________ 95

31 The reference points in delay survey ________ 97

32 Principles of delay predictions by vehicle trajectory reconstruction ________ 98

33 Case 1: Comparisons of delay estimations from various models ____ 102

34 Case 2: Comparisons of delay estimations from various models ____ 102

CHAPTER 1 INTRODUCTION

"It may be sufficient to say that the operation is excellent or dreadful, but a number is nice. In truth, the ability to assign numerical values to different degrees of dreadfulness is important." These are Hurdle's words on signalized intersection operation in a paper titled 'Signalized Intersection Delay Models - A Primer for the Uninitiated' [1]. The quote shows the importance of being able to quantify the operational quality of a signal system. Such a need for the quantification becomes important in signalized intersection analysis. Through an analysis, the performance of a signal system can be evaluated, and the results can be used to compare before and after situations in signal system improvement, and justify the capital expenditure.

Delay, level of service, capacity and queue length are some of the criteria used for evaluating signal operation. Each of the criteria represents different characteristics of signal operation. Delay has become the most popular evaluation tool in signalized intersection analysis. Delays for individual lanes and the overall intersection represent the most powerful means of evaluating intersection performance and in turn provide a method for comparing intersection operations.

Delay experienced by a driver at a signalized intersection is defined as the additional time that the driver has to spend at the intersection compared to the time if the driver could go through the intersection unimpeded. It is often used as one of the governing factors in many engineering and economic decisions. For example, one of the goals for transportation engineers is to design and operate signal systems that minimize delay to the users, as delay is an indication of the drivers' perception of how well the intersection operates. In economic appraisal of an intersection improvement scheme, delay savings form the main benefit that justify capital expenditure.

There are two major of techniques for vehicular delay assessment at signalized intersections: field measurements and analytical estimates. Field measurements of delay involve actual site surveys, and include queue counts and vehicle trajectory reconstructions. They are useful for evaluation of signal operations; however, since they require site surveys, they are not applicable if a signal system is still in the design stage. The analytical estimation technique includes mathematical models that simulate the arrival, queueing and discharge of traffic at signalized intersections. They can be applied in either the design or operation stage since they do not require the actual existence of the signal system. As a result, the analytical delay estimation technique has become popular for signalized intersection delay assessment.

The Canadian Capacity Guide for Signalized Intersections [2] recommends the Whiting delay equation, which is also known as the

Canadian delay equation, as the analytical technique for signalized intersection delay assessment in Canada. The equation was derived using queueing theory and designed for the lane by lane analysis of individual fixed time signalized intersections. Delay equations have also been developed in Australia and the United States.

The Whiting delay equation can predict delay with reasonable accuracy if appropriate values are used for the parameters. However, the selection of the proper parameter values may present a problem. A good example is what to use for evaluation time. Evaluation time is the duration in which a delay assessment is done. It is stated in the Canadian Capacity Guide for Signalized Intersections that the evaluation time should reflect the congestion period, yet the exact relationship is unknown. Moreover, "since the delay for congested conditions depends greatly on the evaluation time, which is rarely determined accurately, the resulting delay should only be considered indicative of the order of magnitude rather than taken as an exact absolute value" [2]. The lack of understanding of this relationship has complicated the selection of the correct evaluation time which gives representative delay predictions of any congested conditions. It is necessary to know how the evaluation time is related to the congestion period, in order that an evaluation time which reflects the congested condition can be applied for the delay prediction. This research is motivated by this urge to gain a better insight into the accuracy of delay prediction; and in particular, designed to look for the relationship between the evaluation time and the congestion period.

The objective of this research is to derive an equation which relates the evaluation time to the congestion period. It is achieved by first reviewing Whiting overflow delay equation and then correlating the evaluation time with the congestion period. The review of the equation focuses on the derivation principles and the relationships between the parameters of overflow delay.

From the review, three additional questions have arisen.

1) Is the prediction of average continuous overflow delay for vehicles that arrive within the evaluation period more practical than that of vehicles which are discharged during that period ?
2) Can the random overflow arrival and delay function be random but time dependent ?
3) Is the absolute value of cycle time influenced by random overflow and the resulting delay ?

These questions have expanded the objective of this research to include an investigation of the possibility of a different approach to random overflow delay prediction.

This thesis summarizes the findings of the research. Chapter 2 discusses the queueing theory and the types of delay used in analytical estimation of delay. Chapter 3 presents the various models of analytical delay estimation. The approach employed in this research to achieve the objectives is explained in Chapter 4. In Chapter 5, the development of a modified model is detailed, while the verification of the model is shown in Chapter 6. Chapter 7 discusses the results and implications of this research.

CHAPTER 2 DELAY AND TYPES OF DELAY

2.1 DEFINITION OF DELAY

The concept of delay can best be explained by visualizing a situation where two vehicles are travelling towards a signalized intersection. One of the vehicles, A, passes through the intersection without having to slow down or stop, but due to interference of the traffic signal, the other vehicle, B, has to slow down and stop at the intersection. Then, it is said that vehicle B has experienced a delay while vehicle A has experienced none. The length of delay is the additional time that vehicle B has to spend at the intersection when compared to vehicle A. This concept is illustrated in Figure 1.

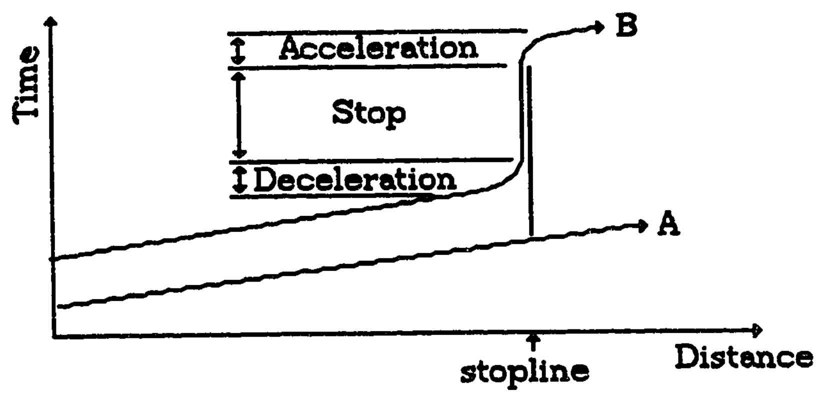


Figure 1 Time space diagram illustrating the basic concept of delay

In traffic engineering, the delay experienced by a driver at a signalized intersection is defined as the additional time the driver has to spend at the intersection compared to the situation of unimpeded access. Such delay may be caused by the traffic signals or by other traffic. Different factors lead to delays of different nature.

However, the actual delay experienced may be different from that shown in Figure 1. Delay may have already started at some point before the stopline. This shows the inherent vagueness of the delay definition.

2.1.1 Vagueness in the Delay Definition

In a paper titled 'Accuracy of Delay Surveys at Signalized Intersections' [3], Teply states that "since direct measurements of delay are not possible, surveys employ indirect techniques based on time-space concept or on queueing theory. A detailed examination of both approaches shows one of the major problems lies in the inherent vagueness of delay definition and different interpretations of the concept of delay." The vagueness of delay is due to several reasons. The two major factors are discussed in this section.

Varying travel speed is one of the causes of the vagueness in the delay definition. Speeds of vehicles approaching a traffic signal are rarely constant, especially when drivers can observe the change of signal from green to amber to red. Therefore, because of drivers' responses to changing signals, the arrival rate encountered at the stopline is not necessary the same as the rate at some point upstream of the signal. Such a situation may lead to an over or underprediction of delay due to the use of an inaccurate arrival rate. A similar problem also exists for vehicle discharge. In actual situation, due to the perception and reaction time of drivers and the acceleration of vehicles involved, the saturation flow is not obtained until a few seconds after the start of green interval. Saturation flow is the maximum rate of vehicle discharge observed during green interval. Therefore, the non definite starting and ending points of delay have caused difficulties in defining the reference points for delay predictions. Accuracy is subsequently affected.

Another reason to the vagueness of the delay definition is the varying arrival rate. In reality, vehicles do not join the queue at the stopline but at some point upstream. The rate at which vehicles are joining the queue is different than the average upstream arrival rate. This is because the end of the queue is travelling backwards. In traffic flow theory, this is called a "shockwave", as shown in Figure 2.

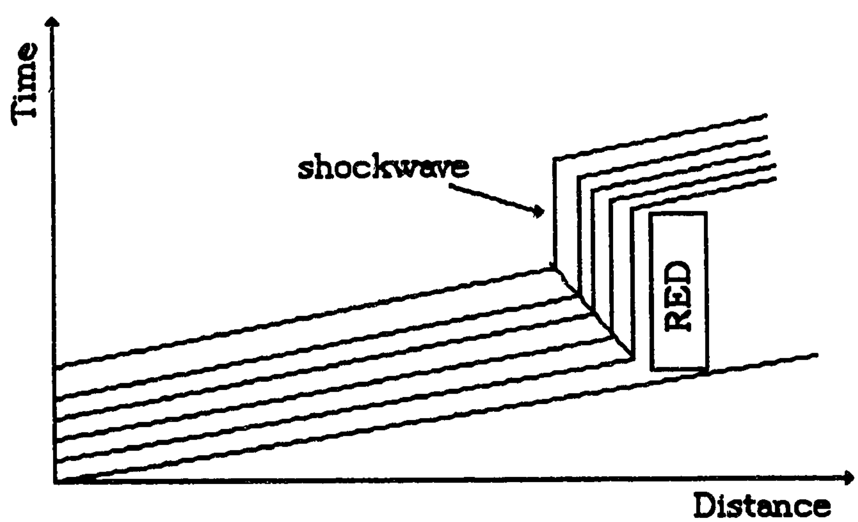


Figure 2 Time space diagram showing a shockwave at the end of the queue of a lane of traffic at a signalized intersection

With the shockwave travelling backwards, the rate of arrival at the end of the queue is greater than the upstream rate of flow. Therefore, if the upstream arrival rate is used for delay prediction, it can result in an underprediction.

2.2 THEORY OF DELAY

The delay theory applied in analytical estimation of delay uses a queueing model. It portrays the queueing situations of traffic at a signalized intersection. There are two approaches to the interpretation of a queueing model: probabilistic and deterministic.

A. Probabilistic concept

The queueing model views traffic as a stream of customers seeking service from a server that only provides service intermittently [1]. In this case, the server is the traffic signal. The customers who have to wait for service experience delay while those who obtain services immediately upon arrival experience no delay. The model assumes the customers arrive in a random manner with an average arrival rate, and they are discharged from the intersection in the order of arrival, at a maximum uniform rate when the signal light is green. A queueing model with these characteristics is called an M/D/1:(FIFO) system, where M represents a Markov or randomly distributed arrival; D stands for a deterministic or uniform discharge rate; 1 means one channel of arrival and service or only one vehicle can arrive and discharge at a time, and FIFO, is the abbreviation for first in, first out, which indicates services are in the order of arrival [4].

The characteristics of the M/D/1:(FIFO) system have limited the application of the queueing model to only individual fixed time signalized intersections. Individual in the sense that the arrival pattern at the intersection is random and not influenced by any upstream traffic pattern. This characteristic of only one vehicle being discharged at a time implies the concept of lane by lane analysis of delay.

B. Deterministic concept

The M/D/1:(FIFO) system is a good representation of the actual traffic condition. However, owing to the high level of uncertainty in the random arrival, it is impossible to consistently predict delay. If it is to predict delay consistently, the uncertainties have to be minimized.

As a result, assumptions are introduced into the queueing model to negate the vagueness caused by the lack of distinct starting and ending points for delay, and the continuous variations in the arrival and discharge rates. It is assumed that there is a constant average arrival rate, and the discharge rate changes instantaneously from zero to saturation flow at the onset of green interval.

The uncertain starting point of the maximum discharge flow complicates the nature of delay. To simplify the theory, the second assumption described above is made. Since this is a major assumption, a compensation is needed. The compensation is that an effective green instead of the actual green interval is used. The effective green is the period of the actual green interval where the discharge rate is equal to saturation flow. Having loss of time at the beginning of the green interval due to the perception and reaction time of drivers as well as acceleration time of vehicles, the effective green should be shorter than the actual green interval. However, since many drivers do make use of the amber period, the time lost at the beginning of green interval is usually offset. This extends the duration of the effective green interval. For the purposes of this research, the effective green interval is assumed to be equal to the actual green interval. It should be noted that the assumption of effective green being equal to the actual green is only valid in fully saturated conditions. Since this research focuses mainly on overflow delay, the assumption is justified. In Toronto, the effective green is found to be equal to the actual green interval, while in Edmonton, the effective green is taken to be equal to the green interval plus one second [6].

Applying the two assumptions to the M/D/1:(FIFO) system, the probabilistic queueing model applied for analytical estimation of delay at signalized intersection becomes a deterministic, or a D/D/1:(FIFO) system.

In graphical form, queueing model is known as queueing diagram. The diagram shows the time of arrival and discharge of every vehicle at a signalized intersection. This information allows the most important "measures of queue behaviour" [4] to be determined from the diagram. Figure 3 is a queueing diagram of a lane of traffic at a signalized intersection.

In Figure 3, the difference between the number of vehicles that arrive and the number of vehicles discharge in time t is the number of vehicles in queue at time t. W_i is the duration i^{th} vehicle has to wait after its arrival at the intersection before its departure. This waiting time is the delay i^{th} vehicle experiences.

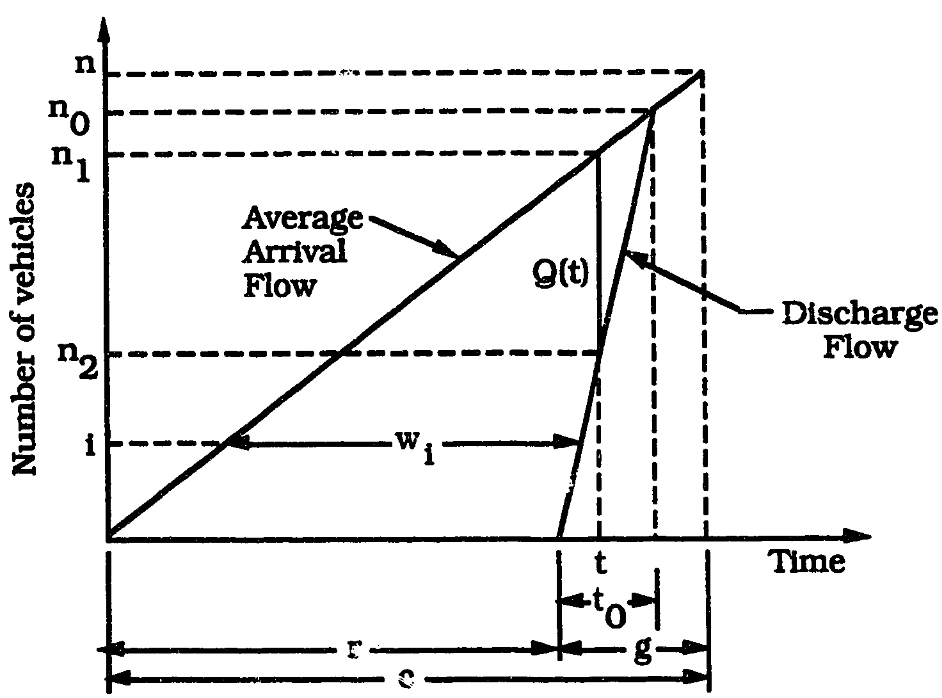


Figure 3 A queueing diagram showing the queueing situation within a signal cycle of a lane of traffic at a signalized intersection

where t_o = Time at which the queue which built up during the red interval has dissipated. Before t_o, the rate of discharge is s, and after t_o , it is q

n_1 = The number of vehicles arrived at the intersection in the traffic lane in time t

n_2 = The number of vehicles discharged from the intersection in the traffic lane in time t

$Q(t)$ = Queue length in the traffic lane at time t

w_i = Waiting time of i^{th} vehicle. The amount of time vehicle i has to wait at the intersection after arriving and before departing

r = Red interval (including amber period)

g = Green interval

Figure 3 is redrawn in another perspective to show the delay of individual vehicle. This is shown in Figure 4. The shaded segment represents the delay experienced by the first vehicle. This segment is of a height of 'one vehicle' and a width of 'w_1 time unit'. The delay for all the other vehicles can be represented by a segment of similar nature, i.e. a height of 'one vehicle' and a width of 'w_n', where n is the vehicle number. With the total delay being the sum of all individual delay, which is equivalent to the sum of the area of all the segments, it is equal to the area between the arrival and the discharge curves. The average delay is then the total delay divided by the number of vehicles discharged.

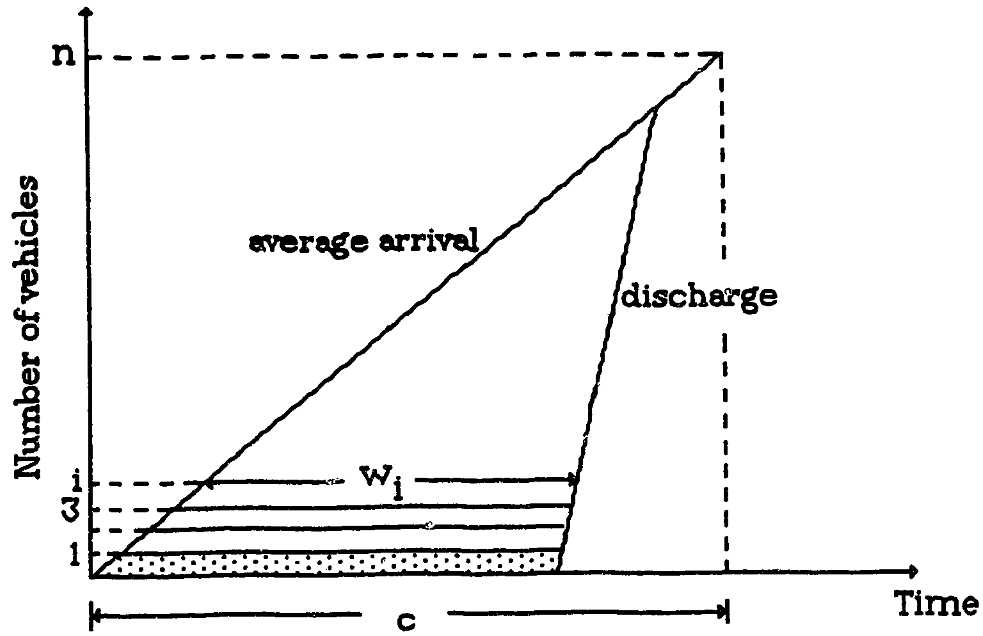


Figure 4 A queueing diagram showing the individual delay of vehicles in a traffic lane at a signalized intersection. The representation is for one signal cycle only.

Therefore,

a) Total delay = Area between the average arrival and discharge curves

b) Average delay = $\frac{\text{Total delay}}{\text{Number of vehicles discharged in a signal cycle, n}}$

2.3 TYPES OF DELAY

Different arrival and discharge patterns lead to different types of delay. Two types of delay are usually examined in delay assessments at signalized intersection. They are the uniform delay and overflow delay. These different types of delay are illustrated in Figure 5 and 6, which are time-space and queuing diagrams, respectively.

Uniform delay is the type of delay results from the interruption of traffic flow by traffic signals. Vehicles arriving during red interval experience the longest delay, as in the case of vehicle a to d in Figure 5a, while those arriving during that portion of green interval where the queue has already dissipated experience no delay, as shown by vehicles f and g. Uniform delay is a function of the cycle time and the degree of saturation. Degree of saturation is the average arrival flow to capacity ratio [5]. Figure 6a is a queueing diagram showing the queueing situation which results in uniform delay.

Overflow delay is the delay experienced by vehicles which are unable to be discharged within the signal cycle they arrive because the arrival flow is greater than the lane capacity. There are two types of overflow delay: random and continuous overflow delay. Random overflow delay is caused by the fluctuations in vehicle arrivals and may occur during an occasional cycle of uncongested operation. A situation where the average arrival flow is less than the capacity, is shown in Figure 5b. It is a function of the signal cycle time and the degree of saturation. Figure 6b is a queueing diagram showing the queueing situation which results in random overflow delay. Conversely, continuous overflow delay is experienced when the arrival flow is consistently greater than the capacity of the lane, thus leading to the formation of a growing queue, as shown in Figure 5c. Continuous overflow delay is a function of the degree of saturation, the absolute difference between the arrival flow and the capacity, and the duration of the congestion.

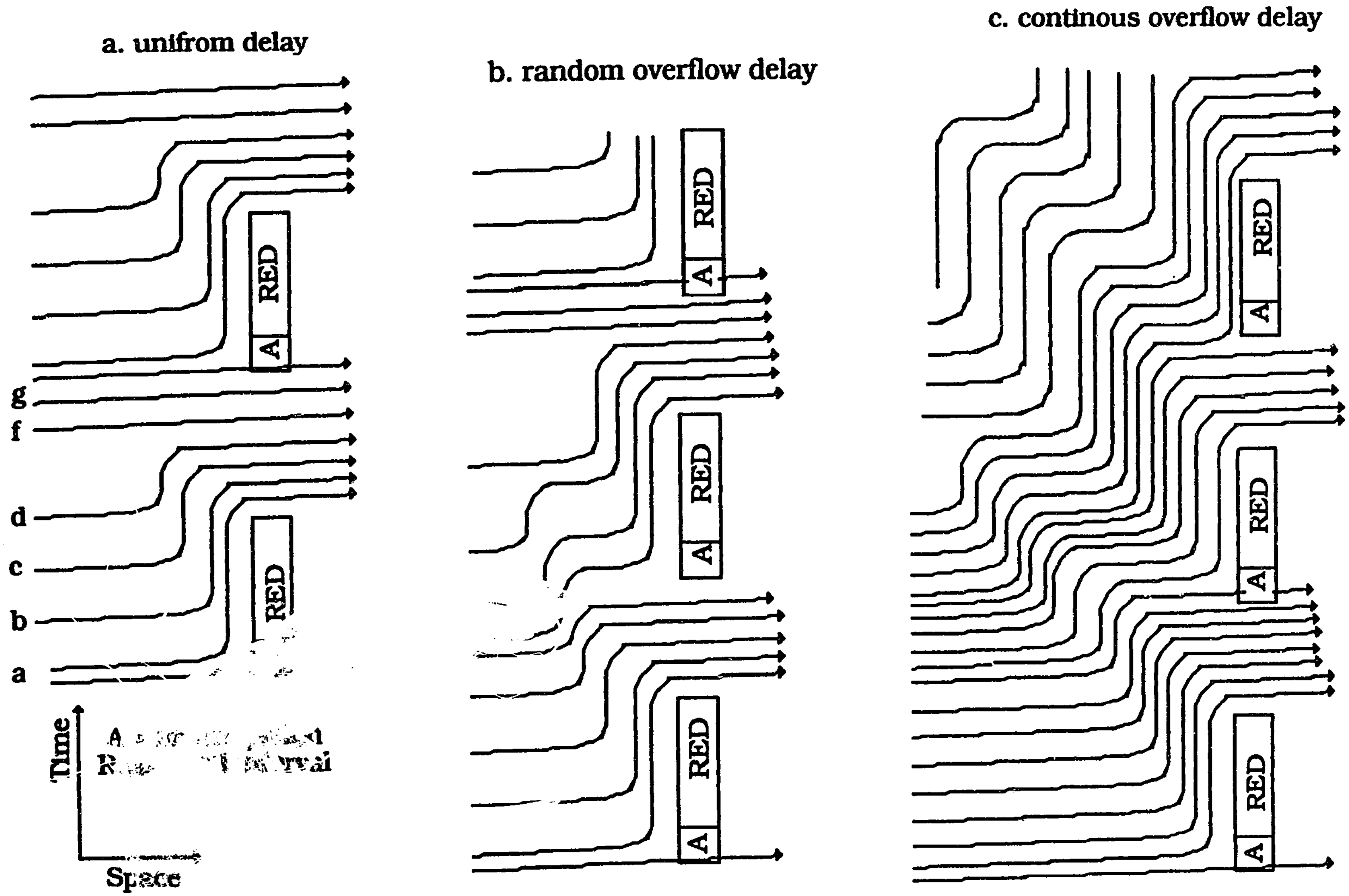


Figure 5a, b and c Time space diagram illustrating uniform and overflow delay (From [2])

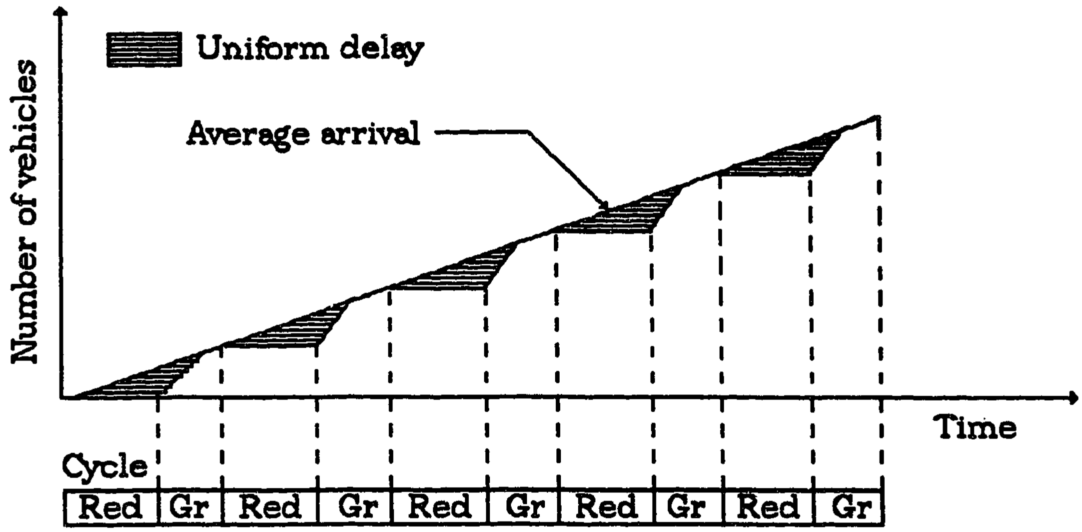


Figure 6a Queueing diagram illustrating undersaturated conditions (from [2])

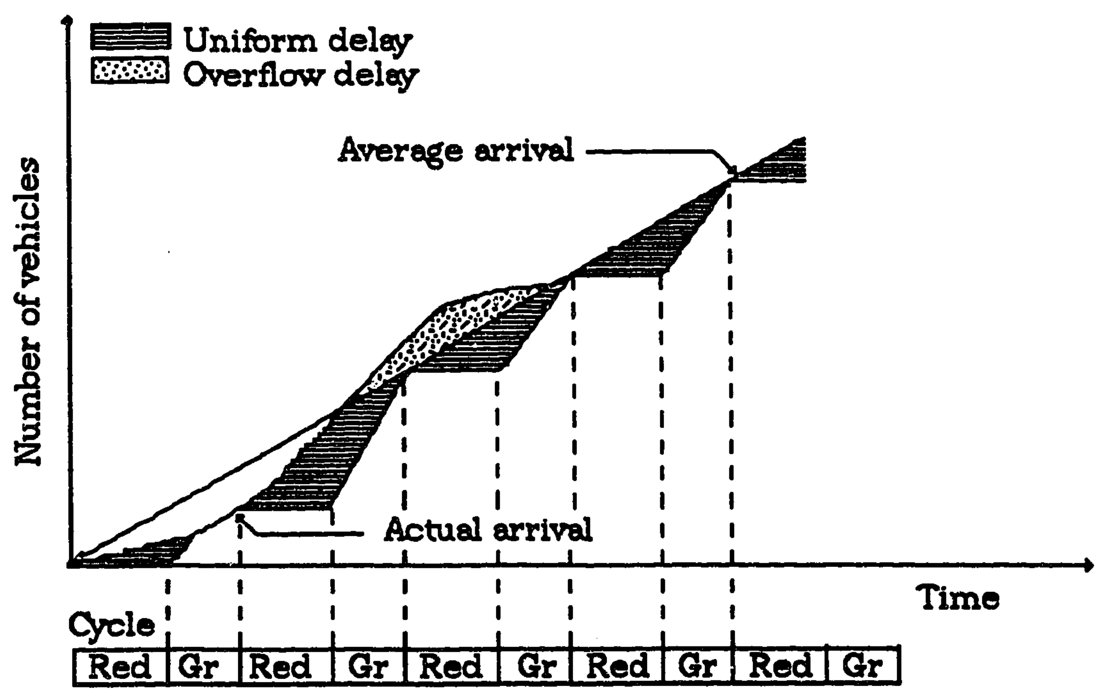


Figure 6b Queueing diagram illustrating capacity conditions with random overflow

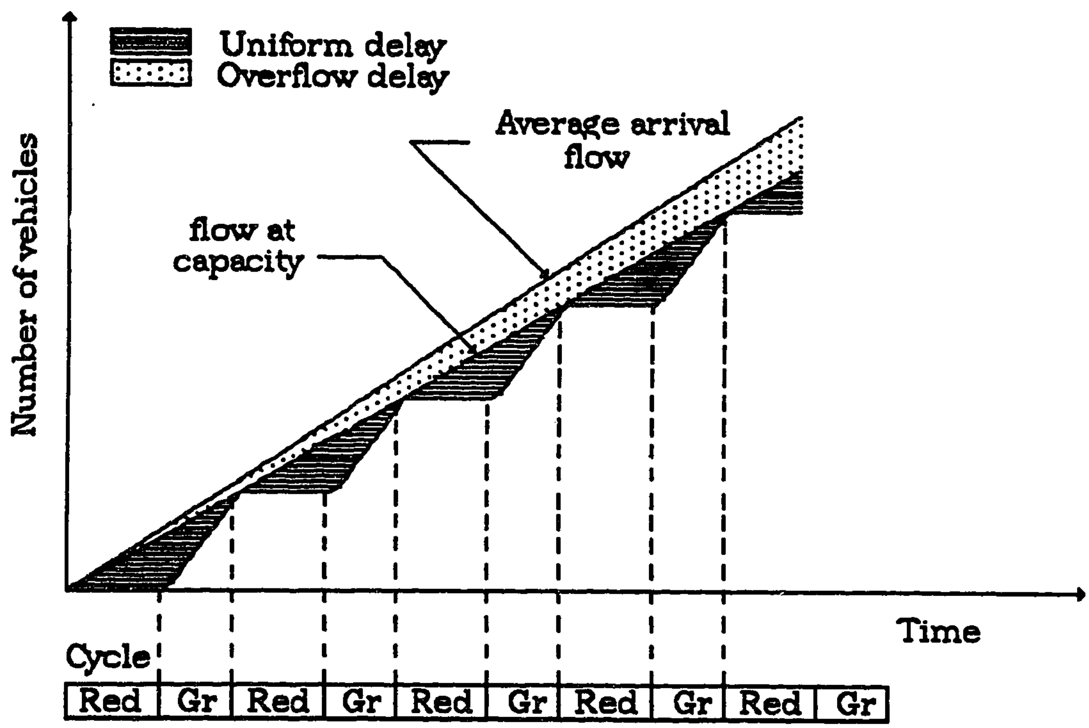


Figure 6c Queueing diagram illustrating oversaturated conditions (from [2])

In a growing queue situation, duration of congestion is important since the longer the congestion period, the longer the total overflow delay will be. As shown in Figure 6c, in which the longer the congestion period, the larger the area representing the total overflow delay. Total overflow delay is represented by the (dottedly) shaded area in Figure 6c. With the total overflow delay increases with congestion period at a rate greater than the number of vehicles increases with congestion period, a longer congestion period will result in a longer average continuous overflow delay.

The total overall delay is the sum of the total uniform delay and the total overflow delay. The average overall delay is the total overall delay divided by the number of vehicles which pass through the intersection during the congestion period.

CHAPTER 3 ANALYTICAL DELAY PREDICTION MODELS

The advantage of analytical delay predictions over field delay measurements has led to the development of various mathematical equations for signalized intersection delay estimations. These delay equations consist of various individual models which represent different types of delay. These individual models are of different mathematical theories. The uniform delay model is steady state and deterministic, while the continuous overflow delay model. Different individual models are then combined to form different delay equations that predict the overall delay.

In this chapter, three basic models which are used to predict different types of delay are discussed. In addition, three combined models used to predict the average overall delay are also discussed.

3.1 BASIC MODELS

In this section, three basic models used to predict three types of delay are discussed. The models are the uniform delay model, the continuous overflow delay model, and the random overflow delay model.

3.1.1 Uniform Delay Model

The uniform delay model is derived from a queueing model using undersaturated traffic conditions.

In graphical form, the queueing model which simulates uniform delay condition is a model with undersaturated traffic condition and has incorporated the assumptions discussed in Chapter 2. The queueing diagram is similar to that of Figure 6a. The assumptions of ignoring the randomness in the arrivals and the instantaneous rise of the discharge rate from zero to the saturation flow have resulted in the following conditions:

a) constant arrival rate, and

b) limiting the discharge rate to only

   i) zero during red interval

   ii) saturation flow during part of green where there is a queue

   iii) a rate equal to the arrival rate during part of green where there is no queue

The two assumptions result in a queueing diagram which consists of a series of identical arrival, queue and discharge patterns.

With the triangles in Figure 6a are identical for every cycle, any one of them can be used for the derivation of a mathematical equation which estimates the average uniform delay experienced during a signal cycle. Figure 7 is one of the triangles.

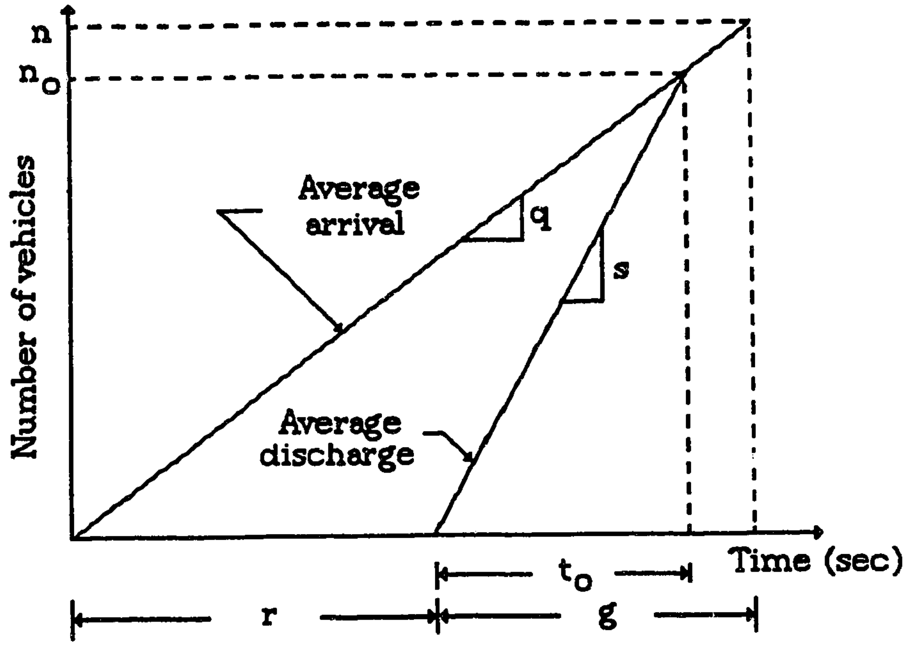


Figure 7 A queueing diagram of a traffic lane with undersaturated flow at a signalized intersection

Applying the principle which states that

$$\text{Average delay} = \frac{\text{Total Delay}}{\text{Number of vehicles in the system}}$$

the uniform delay model can be developed. Total delay is equal to the area between the arrival and the discharge curves. The derivation of the equation is as follows.

a) The average uniform delay is equal to the area between the arrival and the discharge curves divided by the number of vehicles discharge within the cycle.

$$D_U = \frac{\frac{1}{2} r\, n_o}{n} \tag{1a}$$

b) where $n_o = st_o = q(r + t_o)$ and $n = qc$

thus, $q\,r + q\,t_o - s\,t_o = 0$

$$t_o = \frac{qr}{s - q} \qquad (1b)$$

c) substituting Equation (1b) into (1a),

$$\Rightarrow \quad D_u = \frac{1}{qc}\left(\frac{1}{2}\, r\, s\, \frac{qr}{s - q}\right) = \frac{s\,r^2}{2\,c\,(s - q)} \qquad (1c)$$

d) However, as it is common practice to measure flows in terms of vehicles per hour, all flows in Equation (1c) are converted into hourly rate. Converting into hourly flow, where

i) saturation flow, s becomes S, and $S = C\,c/g$ where C is the hourly capacity

ii) arrival flow, q becomes V, and since

$$\frac{r}{c} = 1 - \frac{g}{c}$$

substituting into Equation (1c), yields

$$D_u = \frac{c\left(1 - \frac{g}{c}\right)^2}{2\left(1 - \frac{g}{c}\frac{V}{C}\right)} \qquad (1d)$$

where
- c = cycle time (sec)
- g = green interval (sec)
- V = hourly arrival flow (pcu/h)
- S = hourly saturation flow (pcu/h)
- C = hourly capacity (pcu/h), where $C = S\,g/c$
- D_u = average uniform delay (sec)

3.1.2 Continuous Overflow Delay Model

Equation (1d) was derived based on the assumption of an undersaturated traffic condition. However, there are times that oversaturation occurs over a long period of time. The presence of an arrival flow that is consistently greater than the capacity results in a queueing situation that is different from that shown in Figure 6a. The actual condition is one of Figure 6c. Figure 6c shows that there is continuous overflow delay in addition to the uniform delay. Therefore, in order to predict the delay correctly, a model that can predict the continuous overflow delay is required.

With the application of the same delay theory used in the derivation of the uniform delay model, an oversaturated delay model can be developed. As mentioned before, the average delay is the area between the arrival and the discharge curves divided by the number of vehicles discharged within the evaluation time. Applying this principle, an oversaturation delay equation can be derived from the geometry of a queueing diagram with oversaturated conditions. Figure 8 is a queueing diagram illustrating a traffic condition with continuous overflow. It should be noted that since continuous overflow occurs over a long period of time, the derivation of the equation in terms of hours is more practical.

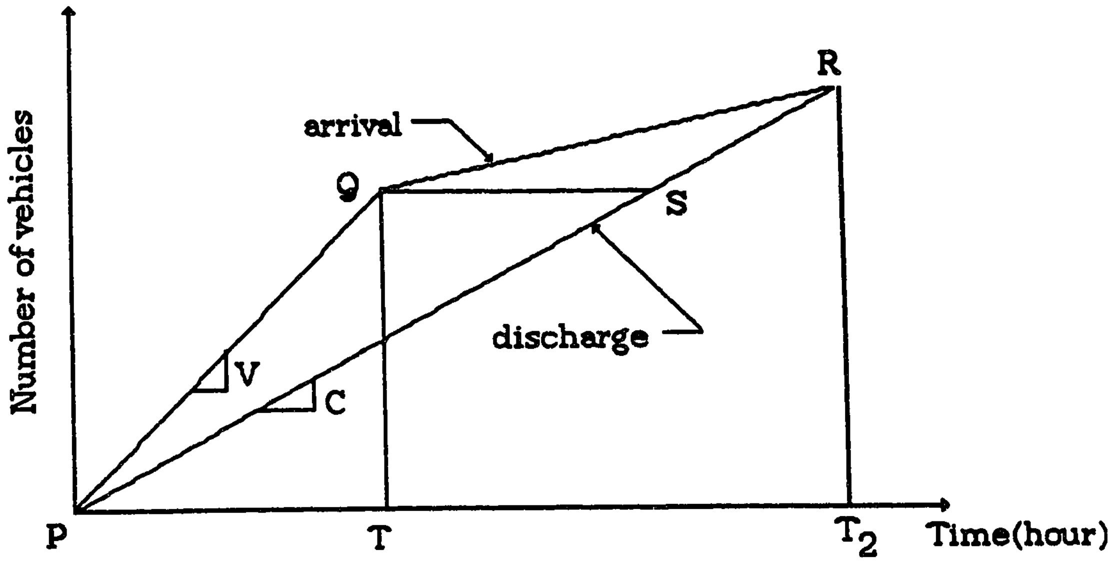


Figure 8 A queueing diagram of a traffic lane with continuous overflow delay. Area PQR representing the total continuous overflow delay

In Figure 8, it is noted that a vehicle which arrives at time zero experiences no delay, but the vehicle that arrives at time T must wait for a time equals to QS. Between time zero and time T, the horizontal distance between the arrival and the discharge curves varies linearly, so the average waiting time for those that arrive between time zero and T is half of the length QS. The same is true for those that arrive between T and T_2. Hence for everyone that is in the overflow queue that exists from time zero to time T_2, the average overflow delay is equal to half of QS. To calculate the average delay, first note that QU is

$$VT - CT = (V - C)T$$

Then, because QU is also equal to QS times the slope of the discharge curve, C, the average delay is

$$D_o = \frac{(V - C)T}{2C} \qquad (2)$$

The notations are similar to that of Equation (1d).

3.1.3 Random Overflow Delay Model

Equation (1d) and (2) are not ideal because in real life, vehicles do not arrive at an intersection uniformly, but rather in a random manner. This randomness in the arrival may cause some cycles of overflow even during undersaturated conditions. The additional delay is not included in the delay prediction of uniform delay or continuous overflow delay model. Therefore, an equation that can estimate this additional delay is needed.

At low arrival flow levels, the effects of randomness are not significant, but when the arrival flow approaches the capacity, the actual average delay experienced will be considerably longer than the average uniform delay, as predicted by Equation (1d). The reason is shown in Figure 9, in which the number of arrivals in the lane is the same during every cycle except in the second cycle, when some extra vehicles happen to arrive. At low arrival flow levels, even with extra arrivals, the total arrivals are still less than the lane capacity, thus causing no extra delay to the system as the extra arrivals are only utilizing the spare time of the green interval of the cycle. The spare green is the result of overdesign of the signal plan, where there is a longer green interval than is necessary. Conversely, at high arrival flow level, when the lane is almost saturated, it takes a long time for the extra queue that builds up in the second cycle to dissipate. The area between the arrival and the discharge curves is thus considerably larger than if the extra arrivals had not occurred. This extra delay is called the random overflow delay, and this signifies the effects of randomness in the arrival flows have on delay experienced at a signalized intersection.

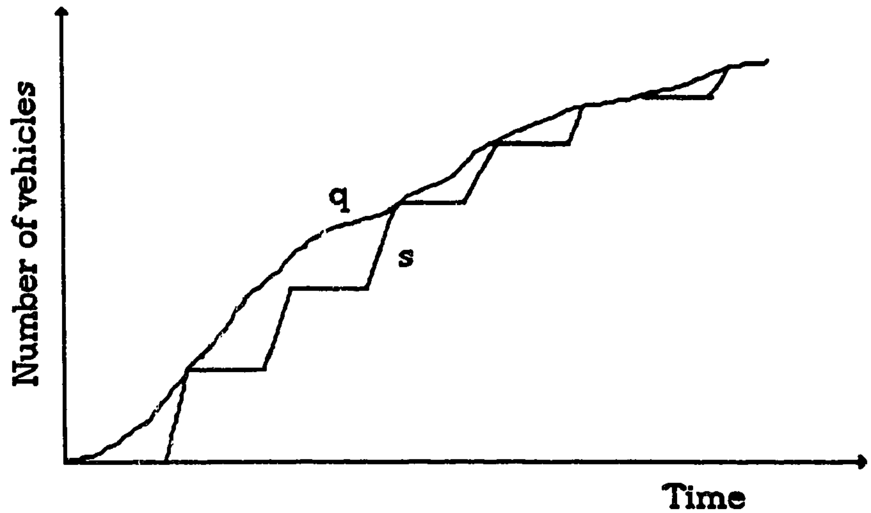


Figure 9 Overflow due to variations in the arrival flow in undersaturated conditions (from [1])

In order "to predict the random overflow delay, it is necessary to construct a stochastic model using the methods of probability theory" [1]. The usual stochastic model depends on four basic assumptions:

"1. The number of arrivals in a given time interval has a known distribution with a constant average, often Poisson, and the distribution does not change with the time of the day or the number of arrivals in any other time interval.

2. The headways between departures from stopline either have a known distribution with a constant mean, or are all the same.

3. Arrival flow is less than capacity.

4. The system has been running long enough to have settled into a steady state." [1]

In 1958, Webster [5] derived a steady state stochastic random overflow delay model. A steady state system is "a system that has operated for a sufficiently long time with the same average values of arrival and discharge and to have settle into a state that the system does not vary with time" and any changes in the system are all internal [7]. On the other hand, a stochastic system is one that "pertains to a process involving a randomly determined sequence of observations, each of which is considered as a sample of one element from a probability distribution", signifying randomness in the system [8]. Combining his model with the deterministic uniform delay model, Webster derived an equation for the average delay, D, where

$$D = \frac{c\,(1 - \frac{g}{c})^2}{2c(1 - \frac{g}{c}\frac{V}{C})} + \frac{(\frac{V}{C})^2}{2V(1 - \frac{V}{C})} - 0.65\sqrt[3]{\frac{C}{V^2}}\,(\frac{V}{C})^{(2 + 5g/c)} \qquad (3)$$

Notations and units are similar to that of Equation (1d).

The first term of Equation (3) is the deterministic uniform delay equation. The second term was derived by Webster to account for the additional delay due to the randomness in the arrival flow. The third term is a calibration term that was determined empirically; it reduces the delay predictions by ten to fifteen percent [9]. All three terms are applied using the units of passenger car unit per hour for flows, and seconds for time dimension.

3.1.4 Comparison of the Models

Figure 10 is a graphical comparison of the three basic delay models.

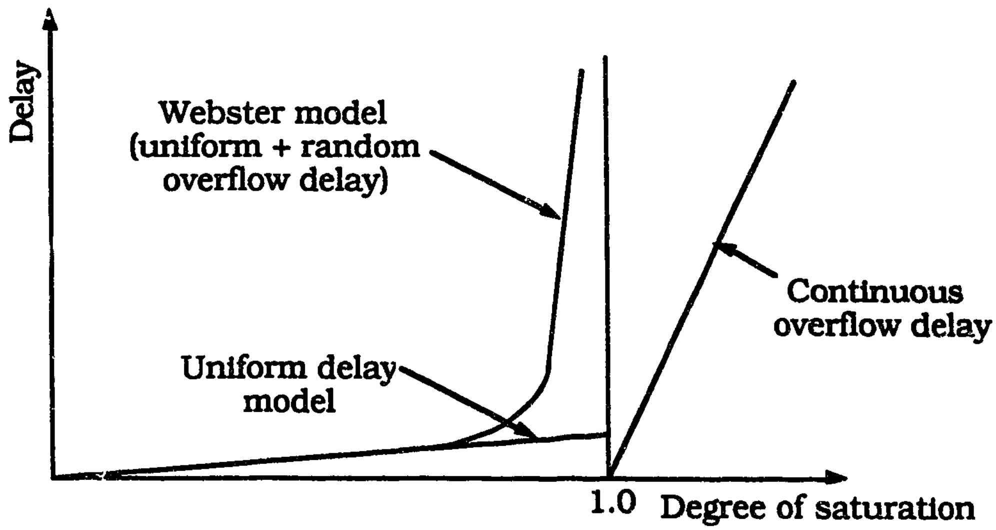


Figure 10 Comparison of Uniform delay model, Webster model and Continuous overflow delay model (from [1])

As shown in Figure 10, the random delay component of Equation (3) is small when compared to uniform delay that occur when the degree of saturation is small, but increases very rapidly as degree of saturation increases and approaches infinity as the degree of saturation approaches one. The reason for Webster model prediction of infinite delay is that in actual situation, it takes a long time for "a system with a degree of saturation close to one to settle into steady state. It simply takes a long time for such a queue to form, especially since vehicles are leaking away through the signal" [1]. Therefore, at low degree of saturation, any arrival within a cycle can be handled in the cycle itself. The constant repetition of arrival, queue, and discharge within the same signal cycle maintain a stable system. Conversely, at a high degree of saturation, where the degree of saturation is much greater than one, a growing queue will form. As the arrival will continuously cause the queue to grow, the state of the system at any time can be predicted. This characterizes a stable system. However, at a degree of saturation close to 1.0, the randomness in the arrival causes the queue to continuously and unpredictably grow and shrink. This continuous variation in the queue prevents the achievment of a stable system. As a result, the steady state conditions required by the Webster model is never achieved. The large delay predicted by Equation (3) will not be encountered.

The discrepancy in the Webster model is a result of the unrealistic assumption that the system is always in steady state, a state that is easily achieved at low and high degree of saturation but takes an infinite amount of time to achieve at degree of saturation close to 1.0 [1]. Thus, the inability of the system to achieve steady state at intermediate flow levels causes the failure of Webster model. In addition, since it is an assumption of the Webster model that the arrival flow is less than the capacity, the model also fails at high flow level, even though the system is in a steady state. In conclusion, the steady state model (Equation 3) is only useful for predicting delay at lightly loaded intersections, where the second and third element of the Webster delay equation do not matter and the arrival flow is less than the capacity.

With the Webster model applicable to only situations where arrival is less than capacity, there is a need for another model to predict the random overflow delay at situation where the average arrival is greater than the capacity. Random overflow does not present a problem when the overflow queue is long because error due to the ignorance of the effects of randomness will be small when compared with the estimated delay. However, this error can be significant if the intersection is only slightly oversaturated. For example, an overflow of eleven vehicles will not cause a significant difference in the average delay from that of an overflow of ten vehicles. In slightly oversaturated intersection, an overflow of one vehicle and overflow of two vehicles cause overflow delays that are very different.

In summary, the steady state queueing model works well when the degree of saturation is considerably less than 1.0. The deterministic queueing model works well when the degree of saturation is greater than one. However, at an intermediate level of degree of saturation, there is a problem. The steady state model predicts infinite delay, while the time dependent model predicts zero delay.

Although each type of delay is significant in a certain range of degree of saturation, they do occur over all ranges of degree of saturation. It would be desirable to have a model that can predict all three types of delay over all ranges of degree of saturation. This calls for a combination of the basic models. For example, a combination of the uniform delay model, the Webster model and the continuous overflow delay model would give a good overall delay model. However, as shown in Figure 10, it is obvious that the Webster model, which has incorporated the uniform delay model, and the continuous overflow models are "utterly incompatible" [1] at degree of saturation of 1.0. The combination of the two models results in a discontinuity. One predicting zero delay while the other predicting infinite delay at the same degree of saturation. Nevertheless, Figure 10 "gives an insight of what the actual delay function ought to be" [1]. Knowing that there is a model, the steady state queueing model, that works well when the degree of saturation is considerably less than one and another model, the deterministic queueing model, that works well when the degree of saturation is considerably more than one, it is logical to think that the actual delay function is one that lies between these two models. This is the basis for deterministic delay prediction models, that are currently being used.

3.2 DELAY PREDICTION MODELS : CURRENT MODELS

With the three models mentioned above as basis, three new models which estimate the average overall delay have been developed. These models which are a variation of each other are named according to the country of application: Australia, America (United States) and Canada. As they were derived based on the three previously discussed models, the constraint applied to the original models also applies to them. The applications of the models are limited to individual fixed time signalized intersections. The three models are discussed in the following subsections.

3.2.1 Canadian Delay Formula

The Canadian Capacity Guide for Signalized Intersections [2] recommends the Whiting delay equation as the analytical delay estimation technique for Canada. This equation consists of a uniform delay model and a time dependent overflow delay model derived by Whiting. Using the Webster steady state model and the continuous overflow delay model as the guides, Whiting concluded that the actual delay function is one that approximates both models at the extremes but in between them at a degree of saturation close to 1.0. This conclusion is consistent with the fact that the two models predicts delay with reasonable accuracy in lightly loaded or heavily oversaturated conditions, but over or underpredict when they are close to degree of saturation of 1.0.

With this conclusion, Whiting combined the uniform delay model and the continuous overflow delay model, and then introduced a modifier to join the two models at degree of saturation of 1.0. He derived a function that approximates the uniform delay model and the continuous overflow delay model at the extremes and is in between the two at degree of saturation close to one. Whiting's derivation of the continuous overflow delay equation is as follows [6]:

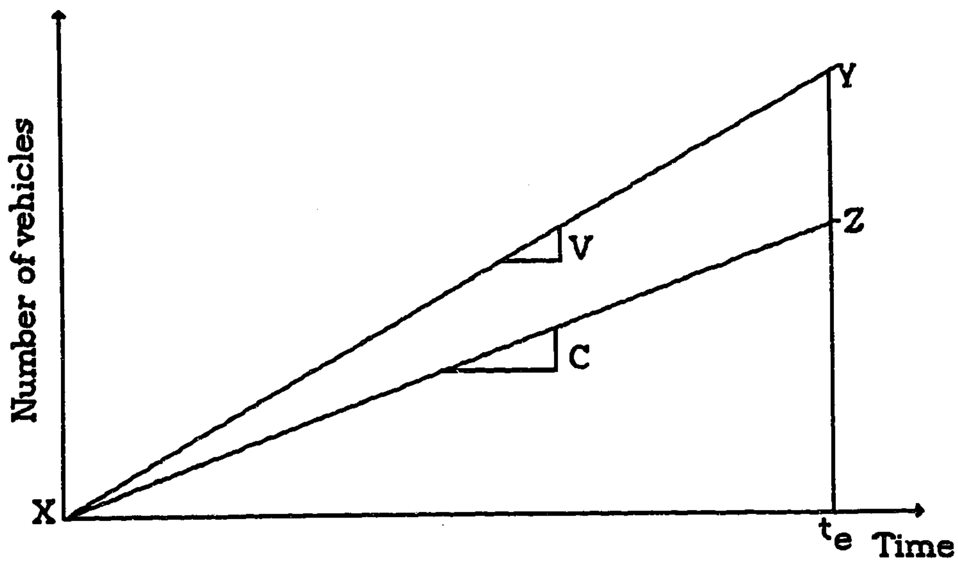


Figure 11 The queueing diagram that is the basis of oversaturation overflow delay equation (from [5])

With the

$$\text{Average Overflow delay, } D_o = \frac{\text{Total overflow delay within time period}}{\text{Number of vehicles discharged}}$$

Total overflow delay = area XYZ, and

number of vehicles discharged = Ct_e,

the average continuous overflow delay D_o is:

$$D_o = \frac{1}{C\,t_e}\left[\frac{1}{2}(V\,t_e - C\,t_e)\,t_e\right] \qquad (4a)$$

and converting the parameters into specific units:

t_e in minutes

V and C in pcu/h

D_o in seconds, yields

$$D_o = \frac{30t_e}{C}(V - C) \qquad (4b)$$

However, when V < C, then D_o < 0. Since

$D = D_u + D_o \implies D < D_u$ --- an impossibility.

Therefore, by spliting the (V - C) term into two parts and negating the effects of a negative value when V < C by using a mathematical manipulation of taking the square root of the square of a number:

$$(V - C) = 1/2 * 2 (V - C)$$

$$\implies \quad = 1/2 * [(V - C) + \sqrt{(V - C)^2}$$

$$\implies \quad = (V - C)$$

Thus, Whiting modified Equation (4b) into

$$D_o = \frac{15t_e}{C} \left[(V - C) + \sqrt{(V - C)^2} \right] \qquad (4c)$$

However, this equation only accounts for average continuous overflow delay. The combination of Equation (4c) and average uniform delay equation still neglects the additional delay due to random overflow.

When V = C, $D_o = 0$, which is not true because of randomness of arrivals. By introducing an empirical modifier into the overflow delay term, Whiting derived an equation that accounts for all three types of delay: uniform, random overflow and continuous overflow delay. The modifier is

$$\frac{240\ V}{t_e}$$

Thus, Equation (4c), the average overflow delay, becomes

$$D_o = \frac{15t_e}{C}\left[(V - C) + \sqrt{(V - C) + \frac{240V}{t_e}} \right] \qquad (4d)$$

Recalling that,

Average Overall delay = Average Uniform delay + Average Overflow delay

$$\Rightarrow \quad D = D_u + D_o$$

where D_u as in Equation (1d) and D_o as in Equation (4d), yields an equation that has three parts, with each being significant only in certain range of degree of saturation in which each type of delay is important.

with
- c = cycle time (sec)
- g = green interval (sec)
- V = hourly arrival flow (pcu/h)
- C = hourly capacity (pcu/h)
- t_e = evaluation time (min)
- D = average overall delay (sec)
- D_u = average uniform delay (sec)
- D_o = average overflow delay (sec)

3.2.2 Australian Delay Model

The Australian Delay Formula was modified by Akcelik into the present form by extending the application of the Miller formula [10] for overflow delay prediction by using the Transport and Road

Research Laboratory time dependent delay method [11]. The equation is as follow: (all notations are consistent with the Australian Road Research Board Report [12]).

$$D = \frac{qc\,(1 - u)^2}{2\,(1 - y)} + \frac{qT_f}{4}\left[z + \sqrt{z^2 + \frac{2\,(x - x_o)}{Qt_f}} \right] \qquad (5a)$$

$$x_o = 0.67 + \frac{sg}{600} \qquad (5b)$$

where

Q	=	hourly capacity (veh/h)
T_f	=	flow period, the time interval during which an average arrival flow persists (h)
x	=	degree of saturation, q/Q
y	=	flow ratio, q/s
z	=	$x - 1$
x_o	=	the degree of saturation below which the overflow queue approximates zero. Thus if $x < x_o$, the overflow delay is zero
sg	=	capacity per cycle (veh/cycle)
s	=	saturation flow per cycle (veh/cycle)
g	=	green interval (sec)
u	=	green time ratio (g/c)
qc	=	average number of arrivals in vehicles per cycle (q = flow in vehicle per second , c = cycle time in seconds)
D	=	average overall delay (sec)

Unlike the Canadian formula, the Australian delay equation has a minimum degree of saturation before the overflow delay term is applied. The Australian equation also places more emphasis on the per signal cycle flow patterns. It incorporates the effects of the absolute signal cycle time have on the uniform and the random overflow delay, as shown in Equation (5a) where the arrival per cycle is used and in Equation (5b) where the minimum degree of saturation before the overflow term comes into effect is evaluated based on the capacity per cycle of the traffic lane under evaluation. The Australian formula places three times more emphasis on the random overflow delay when compared to the Canadian formula.

3.2.3 American Delay Formula

The American delay formula was developed based on the same principles as the Canadian and the Australian formulas. The equation: (With notations consistent with the Highway Capacity Manual [13]). (Equation (6))

$$D = 0.38c \frac{(1 - \frac{g}{c})^2}{(1 - \frac{g}{c} X)} + 173X^2 \left[(X - 1) + \sqrt{(X - 1)^2 + \frac{16X}{C}} \right]$$

where

- D = average stopped delay per vehicle for the subject lane group (sec/veh)
- c = cycle time (sec)
- g/c = green interval to cycle time ratio
- X = degree of saturation
- C = capacity (veh/h)

However, due to a different calibration method, "the Highway Capacity Manual delay formula tends to overpredict the delay in oversaturated conditions. The overpredictions increase with the degree of saturation.

The Highway Capacity Manual formula predicts the stopped delay. Stopped delay is the time vehicles spent at the intersection when they come to a stop. The difference between overall delay and stopped delay is that the latter ignores the acceleration and deceleration delays. Since it is the stopped delay that is being measured in standard intersection delay measurements techniques, the Highway Capacity Manual recommends the prediction of stopped delay as opposed to overall delay. The Highway Capacity Manual conversion of overall delay to stopped delay is obtained by multiplying the overall delay formula by a factor of 0.77 (ie. it assumes that stopped delay is always 77% of overall delay) [13]. This assumption has been questioned. Teply has shown that the correlation between stopped delay and overall delay is not constant but a function of the red interval and the deceleration delay [3].

Moreover, unlike the Canadian and the Australian counterparts, the American delay equation assumes a fixed evaluation time of 15 minutes regardless of what the actual congestion period may be. This assumption of fixed evaluation time does not allow for the consideration of the effects of congestion period have on overflow delay. Overflow delay is very much dependent on the evaluation time and congestion period, since as previously explained in Chapter 2, the longer the congestion period, the longer the total overflow delay. With the total overflow delay increasing with congestion period at a rate greater than the number of vehicles increasing with congestion period, a longer congestion period will result in a longer average continuous overflow delay.

In addition, the American delay equation uses a second order degree of saturation. As a result, "the Highway Capacity Manual formula appears to produce a curve that does not have the fundamental characteristics of time-dependent delay formulation. For degree of saturation above 1.0, it diverges from the deterministic delay line and predicts very large delay values" [11]..

Besides the points mentioned above, the most significant difference between the American delay formula and the Canadian or the Australian formula, is that the accuracy of prediction of the former varies with the degree of saturation. For example, as stated in the Highway Capacity Manual [13],

Degree of Saturation	Accuracy of Predictions
$0.0 \leq V/C < 1.0$	Reasonable
$1.0 \leq V/C \leq 1.2$	Use with caution
$1.2 < V/C$	Not recommended

The divergence of the Highway Capacity Manual delay formula from the continuous overflow formula and subsequently the loss of accuracy with increasing degree of saturation are due to the fixed evaluation time as well as the fixed correlation between the stopped delay and the overall delay [11]. This may also be true for the Australian and Canadian equations. As it is stated in the Canadian Capacity Guide for Signalized Intersections [2], the delay predicted should only be used as an indication of the order of magnitude, because the credibility of the prediction decreases with increasing degree of saturation.

3.2.4 Generalized Delay Equation

Akcelik suggested that since the American, the Australian and the Canadian delay formulas can be treated as variations of one another, they can be written in a generalized equation [11]. He presents the following equation for the average overall delay, D is: (Equation (7))

$$D = \frac{c\left(1 - \frac{g}{c}\right)^2}{2\left(1 - \frac{g}{c}\frac{V}{C}\right)} + 900T\left(\frac{V}{C}\right)^n \left[\left(\frac{V}{C} - 1\right) + \sqrt{\left(\frac{V}{C} - 1\right) + \frac{m\left(\frac{V}{C} - x_o\right)}{CT}} \right]$$

where D = average overall delay (sec)

c = cycle time (sec)

g = green interval (sec)

V = hourly arrival flow (pcu/h)

C = hourly capacity (pcu/h)

T = flow period (h)

x_o = degree of saturation below which the second term

of the delay formula is zero. It can be expressed as
$x_0 = a + bsg$

m,n,a,b = calibration parameters

sg = capacity per cycle (s = saturation flow rate in pcu/sec and g = effective green in sec)

The generalized form of the delay equation predicts the overall delay as opposed to the stopped delay in the Highway Capacity Manual.

With different values applied to the calibration parameters, the Canadian, the Australian and the American Highway Capacity Manual delay equations can be obtained from Equation (7). (See Table 1)

Table 1 Values of the calibration parameters in Australian, American and Canadian overflow delay model. (From [11])

calibration parameters	m	n	a	b
Australian	12	0	0.67	1/600
American	4	2	0	0
Canadian	4	0	0	0

Assuming a situation with the following conditions, the three delay equations will be compared. The conditions are:

Capacity = 1000 pcu/h

cycle time = 100 sec

g/c ratio = 0.50

evaluation time = 15 min. (to be consistent with the American model)

Table 2 illustrates the difference between the delay predicted from the three equations.

Table 2 Delay predicted using the American, Australian and Canadian delay formulas

	Average Delay (sec)		
V/C	AMERICAN	AUSTRALIAN	CANADIAN
0.0	12.50	12.50	12.50
0.1	13.16	13.16	13.16
0.2	13.91	13.89	13.98
0.3	14.78	14.71	14.94
0.4	15.82	15.63	16.10
0.5	17.11	16.67	17.55
0.6	18.82	17.86	19.45
0.7	21.23	19.23	22.08
0.8	25.12	22.52	26.19
0.9	32.97	30.12	34.11
1.0	53.46	51.08	53.46
1.1	100.23	94.11	93.65
1.2	174.88	153.04	150.94

Figure 12 is a graphical representation of Table 2. It shows the followings:

a) random overflow delay is only significant when $0.85 < V/C < 1.05$ (note the range may vary slightly with the saturation flow [2].)

b) the range of significance of each type of delay

c) the difference between the Canadian and the Australian equations is so insignificant that it barely noticeable and the "divergence" [11] of the American Highway Capacity Manual at high degree of saturation.

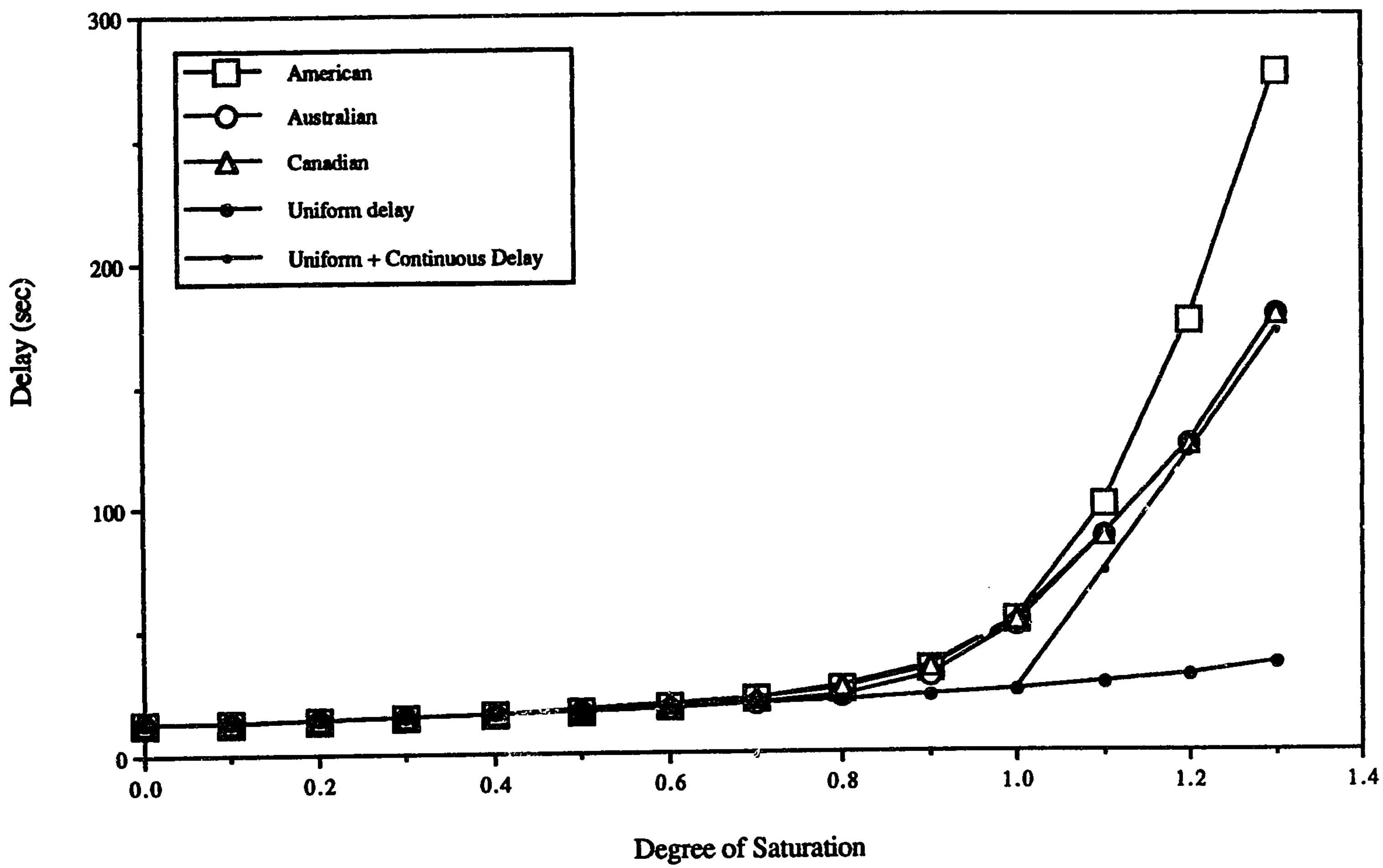


Figure 12 Comparison of American, Australian and Canadian delay equations

CHAPTER 4 WHITING OVERFLOW DELAY EQUATION: A REVIEW

This chapter discusses the review of the Whiting overflow delay equation. Questions are posed and procedures of the review are outlined in the first section. Detailed questions are defined and answered in the subsequent sections.

4.1 THE REVIEW

4.1.1 Continuous Overflow Delay

The following review shows that the Whiting overflow delay equation was derived based on the basic theory of delay as previously mentioned in Chapter 2. The queueing diagram used to represent the theory is a good representation of the actual oversaturated traffic situation. It shows what the key parameters - the arrival and discharge flows, saturation flow, capacity and congestion period. The diagram can also be used to show the effects these parameters have on the overflow delay. These key parameters and their effects are found to be reflected in the delay equation. The Whiting equation is found to be representative of the actual delay condition.

However, two limitations of the equation are identify. The derivation principle leads to the prediction of delay for vehicles discharged instead of for vehicles that arrived within the evaluation period as shown in section 4.3. The equation also fails to define the correct evaluation time - one that gives representative overflow delay predictions of a congestion period.

4.1.2 Random Overflow Delay

The random overflow delay equation is derived for the purpose of providing a smooth transition between the uniform delay and the continuous overflow delay equation. "This transition is not the result of any detailed analysis of queue behaviour but of intuitive ideas of what ought to happen" within this range of degree of saturation [1]. There is no specific theory employed for the derivation of the equation, although the equation seem to be a good numerical representation of the actual random overflowing situation.

From the study on a queueing diagram with random overflow, the parameters that are influential include the arrival and discharge flows, saturation flow, capacity and cycle time. These parameters and their effects on random overflow delay are reflected in the delay equation.

However, there are some differences between the influential parameters and those reflected in the equation. The equation is time dependent and excludes the effects of the absolute value of cycle time, while the queueing diagram shows that random overflow delay should be time independent and cycle time oriented.

4.1.3 Questions Related to Whiting Delay Equation

The previous overview of the theory of delay and Whiting delay equation has posed four questions.

a) What is the evaluation time that gives representative average continuous overflow delay prediction for a congestion period?

b) Is the prediction of the delay expected for vehicles that arrive within the evaluation time more practical than the delay experienced by vehicles that are discharged within the same time period?

c) Does the absolute value of cycle time have a significant impact on the occurrence of random overflow and the resulting delay?

d) Can the random overflow delay model be random while time dependent? If not, how can the random overflow element of Whiting formula be adjusted or replaced to make the formula independent of the evaluation time?

These questions are discussed in details in the following sections.

4.1.4 Procedures of Review

The following procedures are used in the review of the Whiting delay equation.

1) Determine the theory used for derivation of the equation.
2) Determine what the limitations of the theory are.
3) Study the queueing diagram that is representative of the delay condition to determine the influential parameters.
4) Examine if the most influential parameters are included in the delay equation.
5) Study how the parameters affect the overflow delay by using the queueing diagram.

6) Determine if the results from Step (5) are being reflected the same way in the equation.

The above procedures are designed to give information on the overflow delay equation that is crucial to the correlating of the evaluation time and congestion period.

Due to the fact that they are very different in nature, the random overflow and the continuous overflow delay equation will be reviewed separately.

4.2 EVALUATION TIME VERSUS CONGESTION TIME

The Whiting overflow delay equation is designed in such a way that the continuous and the random overflow delay part are dependent on the evaluation time. This is shown in Equation (4d) below, where average overflow delay, D_0 :

$$D_o = \frac{15\ t_e}{C}\left[(V - C) + \sqrt{(V - C)^2 + \frac{240V}{t_e}} \right] \qquad (4d)$$

Due to this characteristic of the equation, the overflow delays predicted using this equation are very sensitive to the evaluation time. As shown in Figure 13, the delay predicted using an evaluation time of

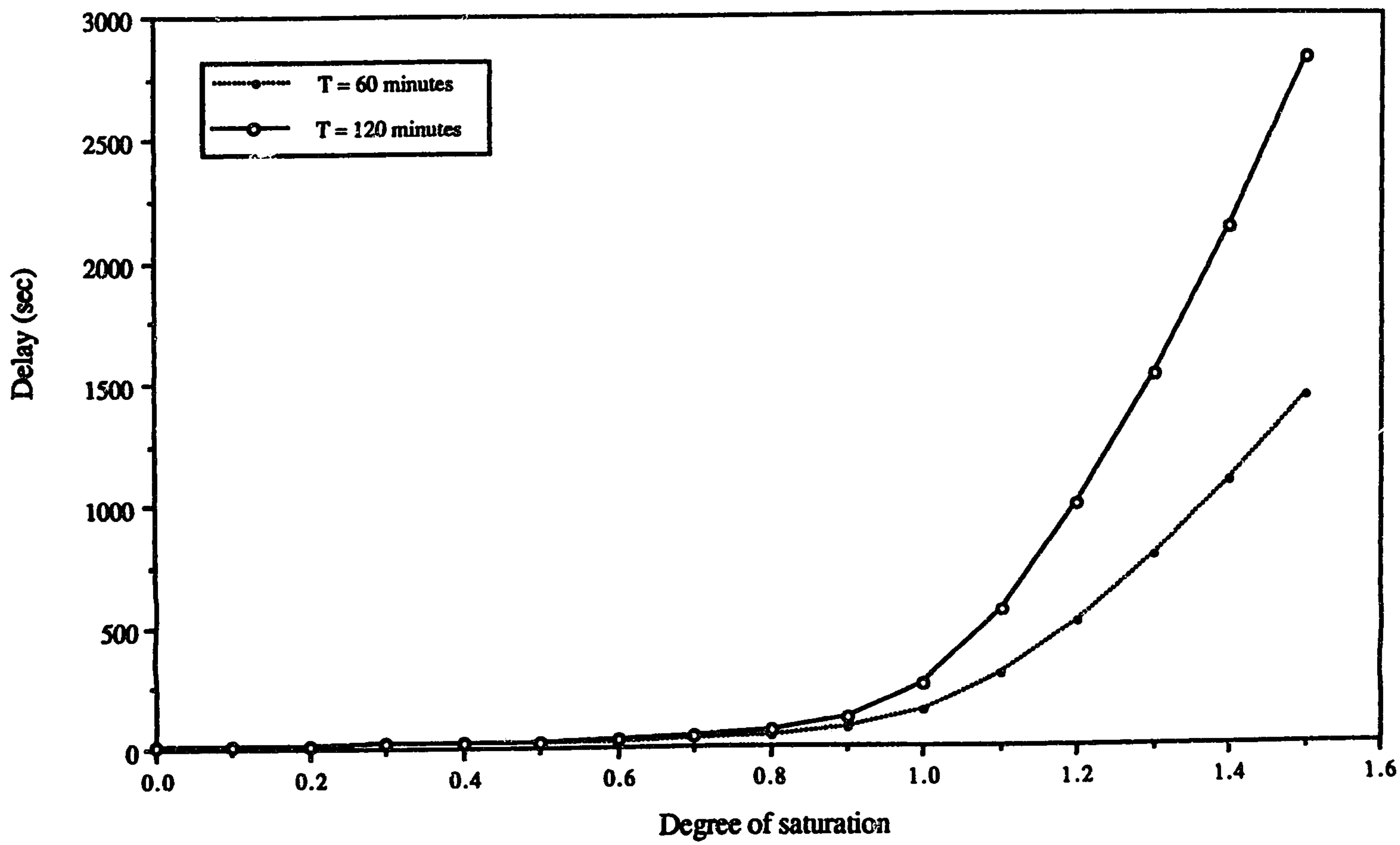


Figure 13 Comparison of delays predicted by using Whiting delay equation with evaluation time equal to 60 and 120 minutes

120 minutes is approximately twice of that using 60 minutes. This shows the evaluation time has a significant impact on overflow delay predictions. Therefore, if the representative overflow delay of a congestion period is to be estimated, the use of the correct evaluation time for delay prediction is necessary.

However, the correct evaluation time for practical applications is not immediately clear from the formula. The Canadian Capacity Guide for Signalized Intersection [2] states that the evaluation time should reflect the congestion period, yet there is no exact relation available. Moreover, the evaluation time is rarely determined accurately in the field. Should it be the period where there is overflow queue, t_c in Figure 14, or the period where the arrival is greater than the capacity, t_o in Figure 14, or the period where some average arrival rate persists. Figure 14 is a flow diagram illustrating the arrival flow over a period of time.

t_c in Figure 14 is the duration where there is an overflow queue at the intersection. This duration is also shown as t_c in Figure 15. t_o in Figure 14 is the duration where the average arrival flow is consistently greater than the capacity, as shown as t_o in Figure 15.

Whiting's derivation of the delay equation uses an evaluation time that is not clearly defined. It is only known that the evaluation time should reflect the congestion period, t_c. With the average continuous overflow delay dependent on the evaluation time, it is essential that one is able to determine the representative evaluation time of any congestion period for the purpose of delay prediction.

In conclusion, since the correct evaluation time is essential to the representative delay prediction of a congestion period, a relation that allows the determination of the corresponding evaluation time would be beneficial. Therefore, it is an objective of this research to find this relationship.

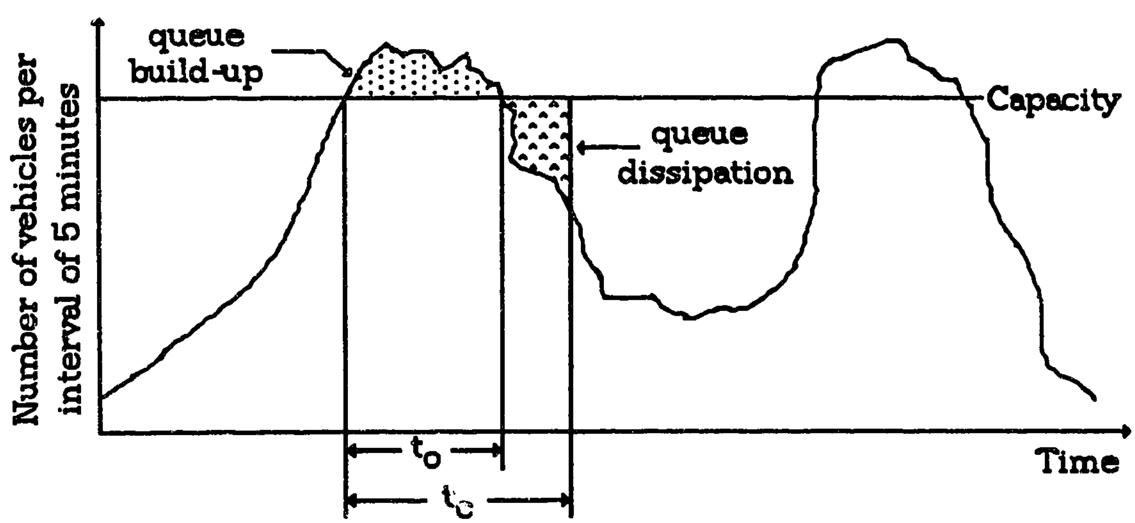


Figure 14 A flow diagram showing the peaks within a daily traffic flow

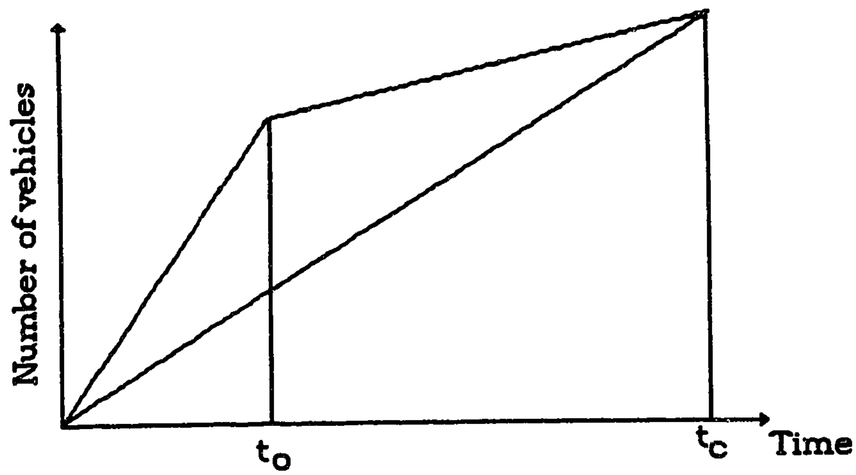


Figure 15 A queueing diagram showing the the duration of congestion and the time to the peak of the congestion

4.3 DERIVATION PRINCIPLES OF CONTINUOUS OVERFLOW DELAY EQUATION

The following is an overview of Whiting's derivation of the continuous overflow delay equation. (See Chapter 3 for details)

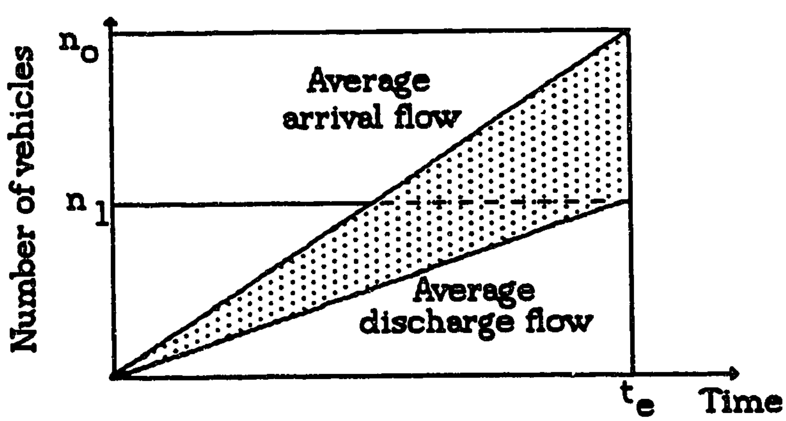


Figure 16 Queueing diagram used for derivation of Whiting continuous overflow delay equation

The shaded area is the total overflow delay experienced within the evaluation period. The average continuous overflow delay, D_0 is:

$$D_o = \frac{\text{total overflow delay experienced within } t_e}{\text{number of vehicles discharged within } t_e \text{ , } n_1}$$

Whiting's use of n_1 vehicles for delay calculation is based on the reasoning that $n_1 + 1$ to n_0 vehicles arrive within the evaluation period, and contribute to delay. However, since they are not discharged within the evaluation period, they are excluded from the average delay calculation. This reasoning leads to an interpretation that the average continuous overflow delay determined is the delay for vehicles discharged within the evaluation period, or it is the delay experienced by vehicles.

It appears that an equation which estimates the expected delay or the delay for vehicles arriving within the evaluation period is of more practical value. This is because a driver is more interested in knowing what is the expected delay than what is the delay experienced, since the latter is something that the driver already knows. This is especially true when drivers plan for the shortest route. Therefore, a derivation principle which results in a model that predicts the expected delay by a driver who is arriving at the intersection within a certain time period may be more practical and useful.

4.4 TIME DEPENDENCY OF RANDOM OVERFLOW DELAY

Random overflow occurs when the arrival is greater than the capacity during occasional cycles. This type of overflow leads to the occurrence of random overflow delay.

It has been a common practice to represent the randomness of traffic arrival patterns at a signalized intersection by a Poisson distribution [2]. The distribution is a single variable discrete probabilistic function used extensively in queueing theory. This function implies that the occurrence of an event is not influenced by another event [7]. Therefore, if an arrival pattern of traffic at a signalized intersection is represented by a Poisson distribution, then the number of arrivals in one cycle is independent of the previous cycles or subsequent cycles, with the constraint that the average of the arrivals in all cycles is equal to the average arrival per cycle used in the distribution. This characteristic implies a time independent system.

However, the Whiting overflow delay model is designed in such a way that the random overflow delay part is dependent on the evaluation time, as shown in Equation (4d) and the following example, where D_0 is the average overflow delay:

$$D_o = \frac{15\, t_e}{C} \left[(V - C) + \sqrt{(V - C)^2 + \frac{240V}{t_e}} \right] \qquad (4d)$$

Assuming a case where the arrival flow is equal to the capacity, V = C (this is to eliminate the effects of the continuous overflow delay portion of the equation on the delay estimations in which the evaluation time has a strong influence), then Equation (4d) becomes

$$D_o = \frac{15\ t_e}{C} \sqrt{\frac{240V}{t_e}} \qquad (8)$$

Assuming a case where the arrival flow equals to the discharge flow, where V = C = 1000 pcu/h, the significance of evaluation time on the random overflow delay predictions using Whiting equation is shown in Table 3.

Table 3 Effects of Evaluation Time on Random Overflow Delay

Evaluation Time (min)	Random overflow delay (sec)
60	56.9
45	49.3
30	40.2
15	28.4

This example shows that evaluation time has a significant impact on the estimated random overflow delay using Whiting overflow delay equation.

However, since the arrival pattern is random, it should be time independent. The following example illustrates the fact.

Figure 17 and 18 respectively depict the probability curves of the per cycle arrival and discharge at a signalized intersection. These Figures are obtained by using a computer program which was written as part of this research. The program simulates the arrival, queueing and discharge patterns of vehicles. The Figures show that the per cycle arrival pattern is not influenced by the evaluation time; and the effect the evaluation time has on the discharge is minimal. From the Figures, it can be seen that random overflow as well as the resulting delay are time independent.

A study on the actual number of arrivals and discharges in every simulated cycle has led to the following explanation of why random overflow delay should be time independent.

A new term called 'underflow' has been introduced. It is the situation where the actual number of arrival at the stopline is less than the discharging capacity of the lane. Note that the computer simulation program assumes vertical stacking of vehicles at the stopline. Since the randomness of arrivals is represented by a Poisson distribution, the characteristics of the distribution must apply to both the random overflow and underflow. It is a constraint of the Poisson distribution that the actual per cycle arrival flows average out to the mean arrival per cycle. That means there are cycles where the arrivals are greater than the average and there are cycles where the arrivals are fewer than the average.

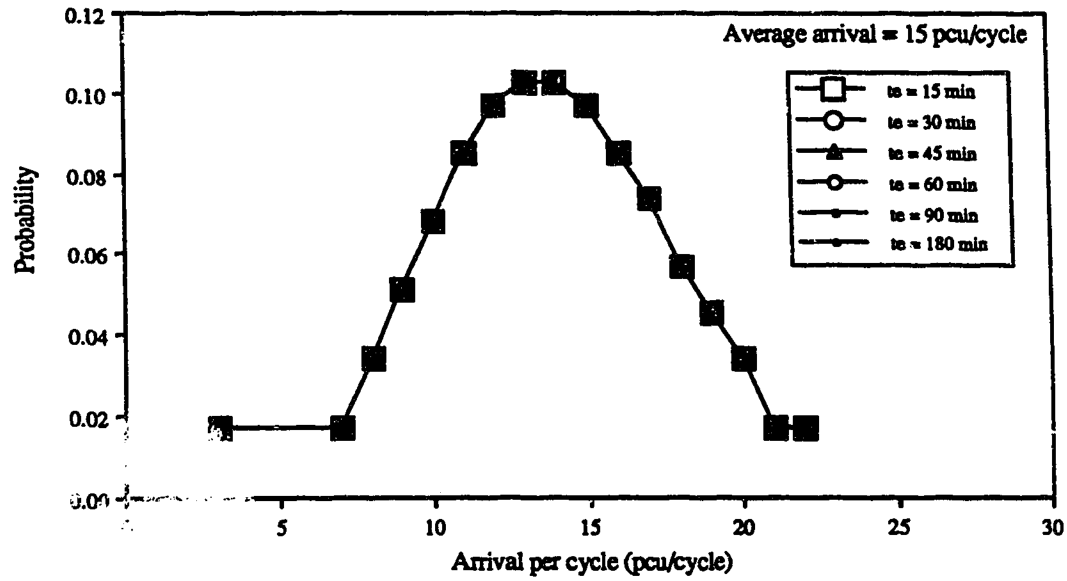


Figure 17 Probability curve of arrival per cycle at a signalized intersection

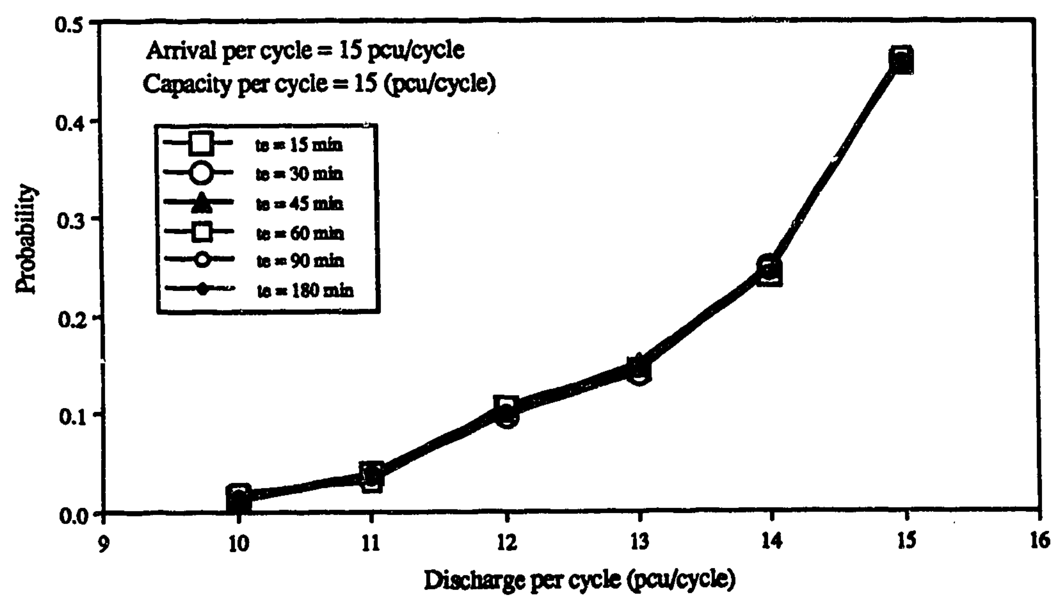


Figure 18 Probability curve of discharge per cycle at a signalized intersection

In other words, the higher than average arrival in one cycle will be offset by the less than average arrival in the subsequent cycles. Therefore, regardless of the degree of saturation, the arrival will average out. This makes the effect of the overflow and underflow in a cycle or cycles time independent.

It should be noted that throughout the discussion, only random overflow and underflow of vehicles are mentioned. This is because only the overflow and underflow of vehicles and not the delay that are offsetting. Random overflow delay is accumulative. This means that a cycle with underflow does not reduce the random overflow delay that has already incurred, even though it may reduce the number of vehicles experiencing random overflow from the previous cycle. Random overflow delay still remains an important delay component when arrival flow approaches capacity.

Table A in Appendix A is a listing of the probability of X number of vehicles that arrive in a cycle calculated using Poisson distribution. The arrival flow is taken to be equal to the capacity in order to eliminate the effects of continuous overflow delay part of the Whiting overflow delay equation. Table A shows that the probability of an underflow is consistently higher than that of an overflow. The reason of that is due to the nature of the Poisson distribution: skewed to the left [14]. Therefore, if there is a cycle of overflow, the overflow will be dissipated in the subsequent cycles. Even if the overflow in the first cycle is not dissipated in the second cycle due to an at capacity flow or another overflow, the system will eventually have some cycles of underflow that will dissipate the built up queue. This is known as the 'regression to mean', and is one of the constraints of the Poisson distribution. Therefore, as long as the average arrival flow and the capacity remain the same, the probability of overflow and underflow remains the same. Thus, as long as the average arrival flow and the capacity remain the same, the underflow will absorb the overflow, making the period of time the average arrival flow persists insignificant.

However, one may argue that the skewing effect of a Poisson distribution is significant only when the mean is close to zero. As stated by Harnett in his book 'Statistical Method' [14], "when the mean is not too close to zero, the shape of the Poisson distribution will often have a very symmetrical appearance'. However, considering the probability of arrival in Table A, it can be seen that the skewing effect is still significant up to an average arrival of forty vehicles per cycle. Table 4 shows the skewing effects of the Poisson distribution.

Table 4 Skewing effects of Poisson distribution

Probability of	Average arrival per cycle (V = C) 20	25	30	35	40
P(C-10 to C-1)	0.4652	0.4609	0.4538	0.4451	0.4357
P(C)	0.0888	0.0795	0.0726	0.0673	0.0629
P(C+1 to C+10)	0.4273	0.4246	0.4194	0.4126	0.4055

P(C-10 to C-1) in Table 4 is the sum of the probability of arrival of the capacity per cycle minus one vehicle compared to the capacity per cycle minus ten vehicle. P(C+1 to C+10) is the sum of probability of arrival of the capacity per cycle plus one vehicle compared to the capacity per cycle plus ten vehicle. P(C) is the probability of arrival of the at-capacity vehicle.

Table 4 shows that the probability of overflow is approximately 3 to 4% lower than that of the underflow. For the application in this research, it can be considered that the skewing effect of Poisson distribution is still be significant. Hence, the overflow in one cycle will always be offset by the underflow in the following or subsequent cycles. This makes the length of time an average arrival flow persists insignificant.

In conclusion, the above examples show that if the arrival flow pattern is represented by a Poisson distribution, then the constraints of the distribution must also apply to the model that represents the arrival flow. Therefore, it is questionable to assume the arrival to be random in nature but is time dependent. A model that is discrete, probabilistic and time independent may be a more realistic representation of the actual random overflow situation.

4.5 EFFECTS OF THE ABSOLUTE VALUE OF CYCLE TIME

The Whiting random overflow delay model does not include the effects of the absolute value of cycle time. It considers its significance only in terms of the green interval to cycle time ratio. As shown later in this Chapter, delay predicted using Whiting delay equation is not influenced by the absolute value of the cycle time.

However, as Figure 19 illustrates, the cycle time is influential in terms of the absolute value as well as in the green interval to cycle time ratio.

Example

Consider a situation where there are two signal plans with the same green interval to cycle time ratio but one of the plan has a cycle time that is twice of the other.

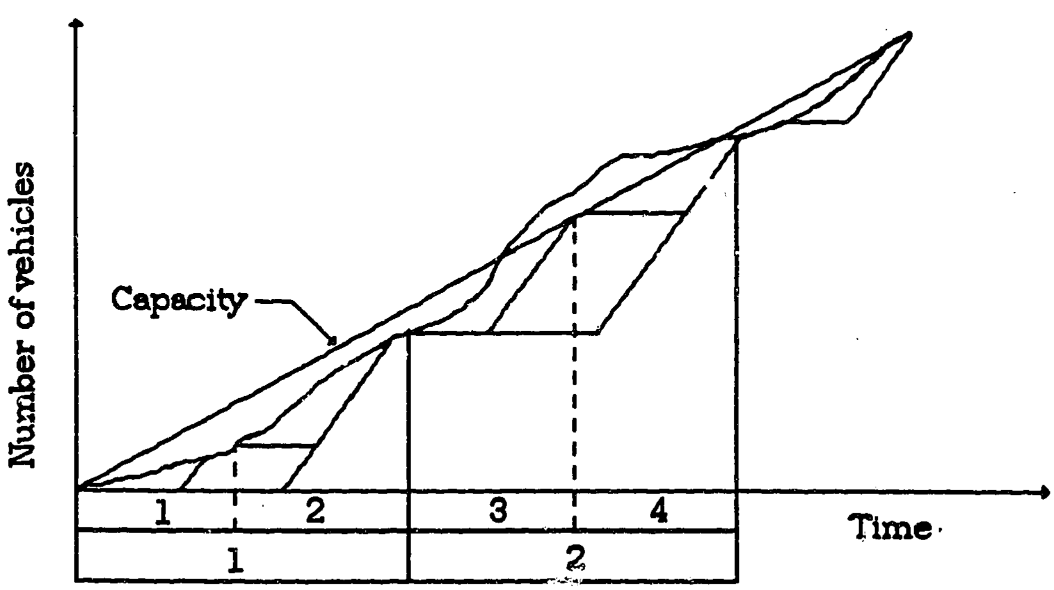


Figure 19 A queueing diagram showing the effects of cycle time on occurrence random overflow

The diagram shows that the random overflow which occurs during the third cycle of the signal plan with a shorter cycle time would not occur if a longer cycle time were used. Table 4 shows the same result. In Table 4, the probability of overflow for a shorter cycle time is 2.2% to 3% higher than that of a longer cycle time. For example, in a case of 20 arrivals per cycle compared to 40 arrivals per cycle, the latter has a cycle time twice of the former and has a probability of overflow of 0.4025 while the former has a probability of 0.4273.

The observation made from this diagram has led to a detail analysis on the subject.

Further testing of this subject involved a computer simulation to study the occurrence of random overflow with respect to the cycle time. The random overflow delay calculated by the Whiting delay equation was examined using the theory of probability. The reason for this testing is based on the fact that since random overflow delay is probabilistic in nature, and by including the elements of probability, a more realistic representation of the traffic conditions may be provided.

TEST A

Figure 20 shows that with shorter cycle time, the occurrence of greater than capacity arrival is more frequent than with longer cycle time.

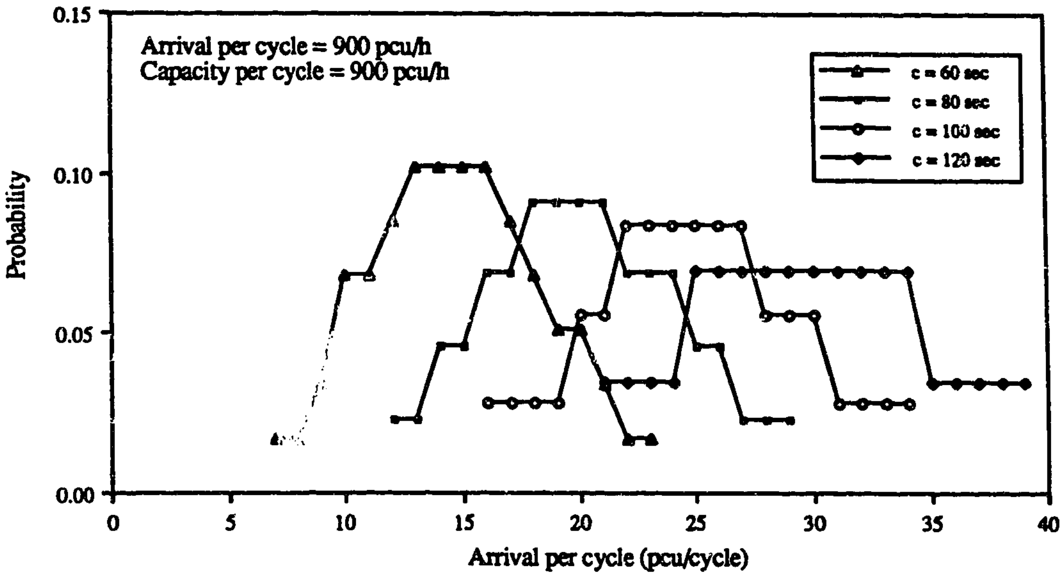


Figure 20 Probability curve showing arrival per cycle at the stopline of a signalized intersection

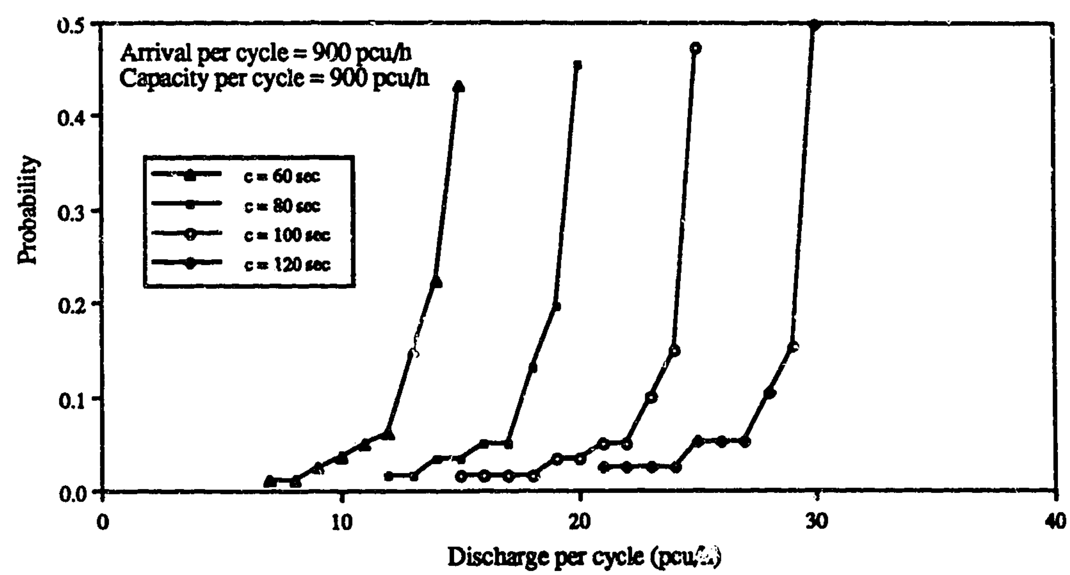


Figure 21 Probability curve showing discharge per cycle from the stopline of a signalized intersection

Figure 21 shows that a signal plan with longer cycle time experiences fewer at-capacity discharges. Therefore, it can be said that longer cycle time results in a lower probability of overflow.

In conclusion, Figure 20 and 21 show that cycle time influences the occurrence of random overflow delay in terms of both the cycle time and the green interval to cycle time ratio.

However, the question whether a probability of lower occurrence of random overflow with longer cycle time leads to shorter random overflow delay still remains. The question is answered in the following test.

TEST B

Since the randomness of arrival is represented by the Poisson distribution, the function is used for determining the probability of arrival of vehicles.

$$P(X) = \frac{m^X e^{-m}}{X!}$$

where P(X) = probability of x vehicles arrive

m = average arrival

X = number of arrivals

The counting interval used for this research is a signal cycle. This selection is based on two reasons. An hour interval is too large and does not show the effects of randomness. On the other hand, a counting interval of one second is too small for it only shows if there is a vehicle, and conceals the effects of randomness. A counting interval on the basis of a signal cycle does not have the above problems, and it is the ideal time period for the counting interval. Therefore, the average arrivals per cycle is used, making the Poisson equation

$$P(X) = \frac{V_c^X e^{-V_c}}{X!}$$

where V_C = average number of arrivals per cycle (pcu/cycle)

Thus applying the Poisson distribution, the probability of overflow is

$$P(OVERFLOW) = P(C+1) + P(C+2) + P(C+3) + ...$$

where C = capacity per cycle (pcu/cycle)

It is the summation of the probability of arrival of vehicles that cause overflow, where P(C+1) is the probability of arrival of the capacity per cycle plus one vehicle.

Since number of vehicles arriving is a function of probability, the delay that occurs as a result of the random overflow can be determined using the theory of probability. In probability theory, the average or "expected value, E(x) of a discrete random variable x is found by multiplying each value of the random variable by its probability and then summing over all values of x" [12] or

$$E(x) = \Sigma\ x\ P(x)$$

For the purpose of this research, the random variable x is the overflowing vehicles. Since every overflowing vehicle has a corresponding overflow delay, the average overflow delay can be determined using the principle of expected value of probability. Thus, the average overflow delay due to the randomness in the arrival is:

$$\text{Average overflow vehicle delay} = \sum \text{expected delay for vehicle x} * \text{probability of x arriving}$$

Applying the above principle, the random overflow delay of any traffic condition can be determined. Consider a case where a signalized intersection has five signal plans of the same green interval to cycle time ratio but different cycle time.

Assuming the following values for the parameters:

Arrival = discharge = 900 pcu/h

Saturation flow = 1800 pcu/h

green interval to cycle time ratio = 0.5

evaluation time = 60 minutes

cycle time = 60, 80, 90, 100, 120 seconds

The arrival and the discharge is made equal to eliminate the effect of the continuous overflow delay part of the Whiting overflow delay equation. Now, assume a case of overflow where the arrival is $1.2V_C$ vehicles in the first cycle is followed by arrivals of V_C or less vehicles in all subsequent cycles. The selection of $1.2V_C$ vehicles is based on the findings that a degree of saturation of 1.2 is usually the upper limit in overflow of arrivals [2,11,13].

With this information, the random overflow delay can be determined. Results are shown in Table 5. See Appendix B for details of calculations.

Table 5 Effects of cycle time on random overflow delay D_{RO}

cycle time (sec)	# cycles per hour	arrival or capacity per cycle	P(overflow)	probability	Whiting
60	60	15	0.452	9.8	60.0
80	45	20	0.447	14.7	60.0
90	40	22.5	0.444	16.9	60.0
100	36	25	0.441	19.9	60.0
120	30	30	0.432	25.6	60.0

Table 5 shows that signal plans with same green interval to cycle time ratio but of different cycle time have different probabilities of occurrence of random overflow. It also shows that the random overflow delays predicted using theory of probability vary with the cycle time, while the delays predicted by the Whiting overflow delay equation are independent of the cycle time.

The fact that the probability of overflow varies with cycle time is supported by the Canadian Capacity Guide [2]. It can be interpreted from Figure 22, a diagram taken from the Guide, that the probability of discharge increases with cycle time. With this fact and the observations made from Table 5, it is believed that random overflow delay is dependent on the absolute value of the cycle time, and that it should be incorporated into the delay prediction model.

It should be noted that the difference between the delay estimated by theory of probability and the Whiting equation may be due to an assumption of different overflowing situations. However, it was not an objective tp examine the accuracy of the overflowing situation used by Whiting, but merely to demonstrate that the absolute value of cycle time is crucial to the occurrence of random overflow and the resulting delay.

4.6 THE IMPLICATIONS

As a results of the above findings, the objective of this research is expanded to include modifications to the Whiting delay equation. The new approach involves a probabilistic instead of the traditional deterministic method to the estimation of random overflow delay. It determines the expected delay instead of delay experienced. It appears that the new approach may be a more realistic representation of the actual traffic conditions, and will improve the quality of delay prediction.

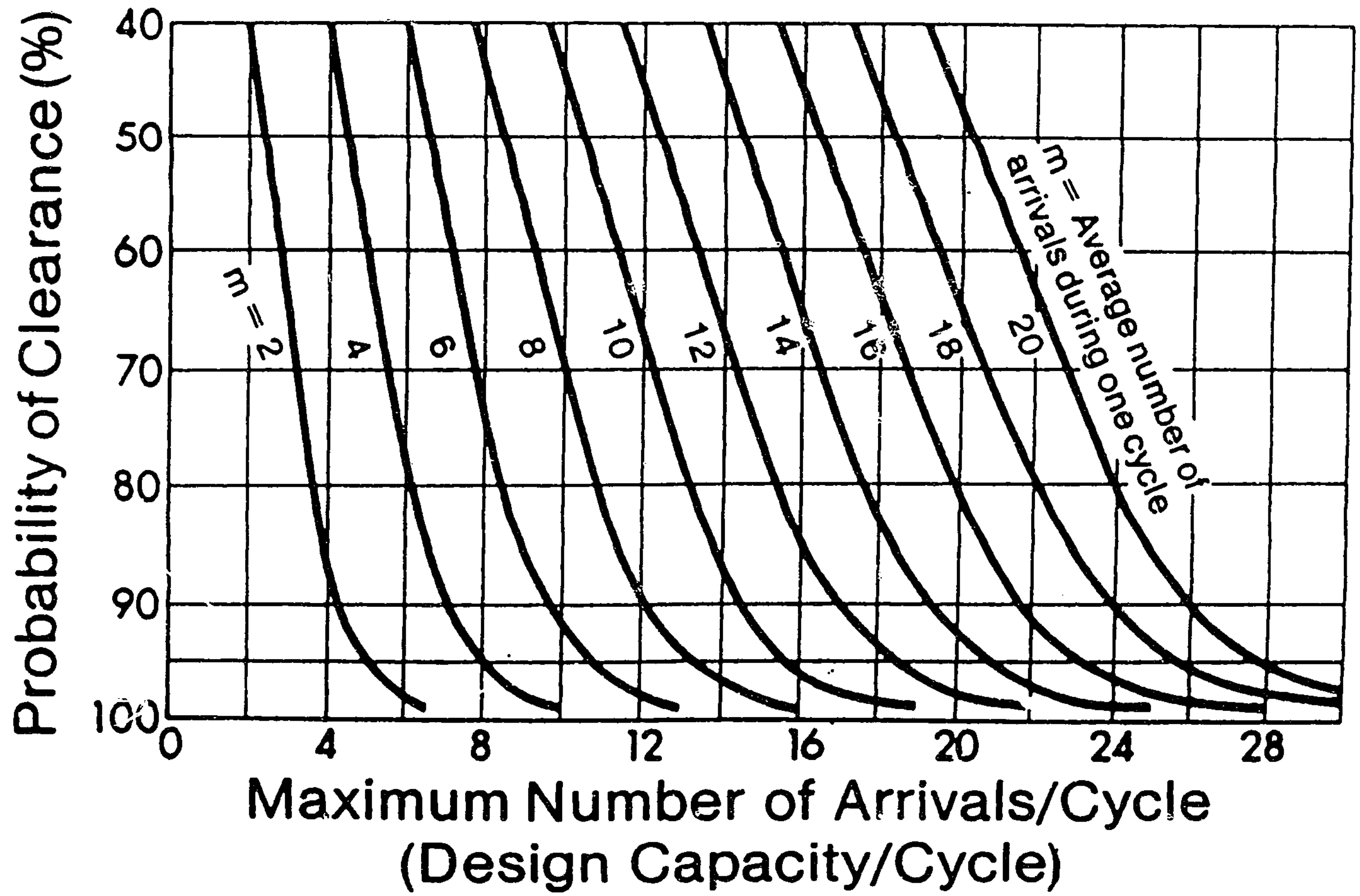


Figure 22 Poisson Distribution applied in the determination of probability of clearance at a signalized intersection [2]

CHAPTER 5 MODIFIED WHITING DELAY EQUATION

In this Chapter, the development of the modified Whiting delay equation is discussed. The Chapter includes the derivation of the equation which correlates evaluation time to congestion period and the development of the new random overflow delay prediction model. The delay equation is presented according to its three parts : uniform, random overflow and continuous overflow delay.

5.1 UNIFORM DELAY MODEL

It is not part of this research to review the uniform delay model, so no modification has been attempted. The uniform delay model applied in the modified equation is identical to that in the Whiting delay equation.

5.2 RANDOM OVERFLOW DELAY MODEL

5.2.1 General Model

The following conclusions can be made from referring to the discussion on the questions associated with the Whiting random overflow delay model in Chapter 4.

1. The arrival pattern of traffic at signalized intersections is assumed to be random in nature and best represented by Poisson distribution [2].

2. The occurrence of random overflow and the resulting delay is affected by the signal cycle time.

3. Average random overflow delay is time independent.

The modified random overflow delay model takes these three factors into consideration.

(1) Since arrival pattern of traffic at signalized intersection is represented by a Poisson distribution, the probability of arrival as well as the probability of overflow can be determined by using this distribution. The Poisson distribution is

$$P(X) = \frac{m^X e^{-m}}{X!} \qquad (9)$$

where $P(X)$ = probability of X vehicles arrive

m = average number of vehicles arrive

X = actual number of vehicles arrive

(2) Since random overflow delay has been shown to be cycle time oriented, the random overflow delay model can be derived in terms of arrival and capacity per cycle. The probability of random overflow occurring is determined by summing the probability of arrival of the vehicles which cause overflow. In mathematical form, random overflow occurs when $X > C_C$. Thus, the probability of random overflow:

$$P(X > C_C) = \sum_{X = C_C + 1}^{X = \infty} \frac{V_C^X e^{-V_c}}{X!} \qquad (11)$$

where $P(X > C_c)$ = total probability of random overflow

V_c = average arrival per cycle (pcu/cycle)

C_c = capacity per cycle (pcu/cycle)

X = actual number of arrivals in the cycle (pcu)

(3) A time independent probabilistic model is used for analytical estimation of random overflow delay because the arrival function is commonly represented by Poisson distribution, which is a probabilistic function, and the average random overflow delay is found to be time.

With the random overflow delay model being a probabilistic function, the principle of expected value of a random variable from a probabilistic distribution, previously mentioned in Chapter 4, can be applied to determine the average random overflow delay. The principle can be interpreted as the average delay is equal to the summation of all the products of the probability of a delay occurring and the delay expected if a delay does actually occur [7].

In this research, the variable is the delay expected by the x^{th} vehicle

$$D_{RO} = \sum_{X = C_c + 1}^{X = \infty} P(X)\, t_X \qquad (12b)$$

where D_{RO} = average random overflow delay (sec)

t_X = expected delay corresponding to the X^{th} vehicle (sec)

$P(X)$ = probability of X number vehicles arrive or that of t_X occurring

C_c = capacity per cycle (pcu/cycle)

Equation (12b) is the general random overflow delay model in the modified delay equation. Details of the development for each term of the equation are discussed in the following sections.

5.2.2 The Probability Term, P(X)

Probability of experiencing random overflow in a signal cycle can be determined by summing the probability of arrival for all vehicles which may cause overflow in that cycle, as shown in Equation (11). This is a summation of the probability of arrival of the capacity-per-cycle plus one vehicle to the infinite vehicle.

However, an examination of the individual probability will show that the summation is incorrect because in actual situation, the arrival of the infinite vehicle will not occur. Not only that, the individual probability of arrival for each overflowing vehicle approaches zero rapidly as the degree of saturation increases, as shown in Figure 23. The conclusion which may be drawn from these two findings is that the summation of the probability of arrival for the overflowing vehicles can be limited to a specific number of vehicles greater than the capacity.

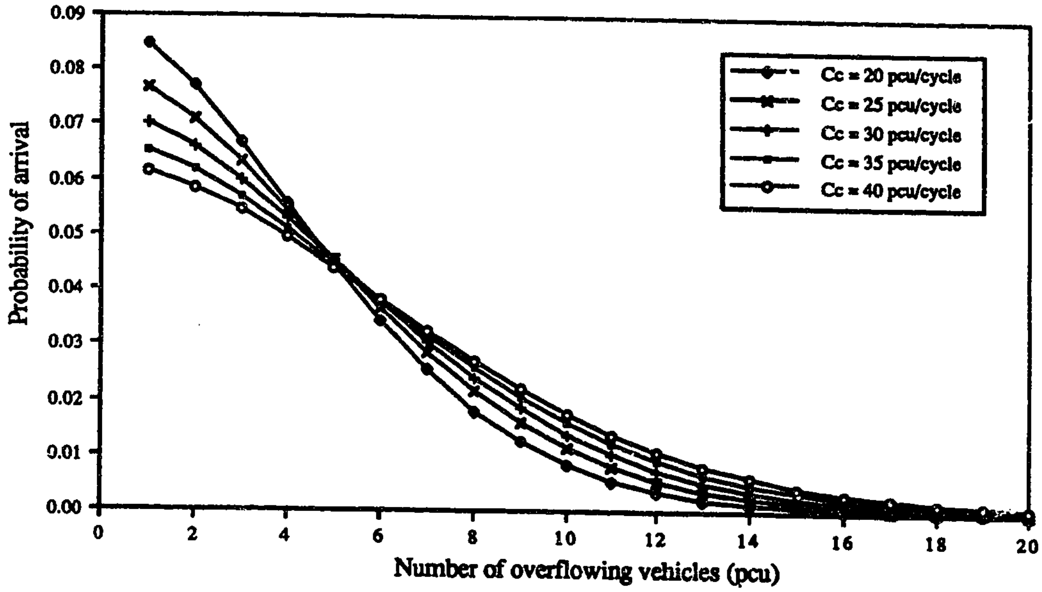


Figure 23 Probability of arrival for overflowing vehicles

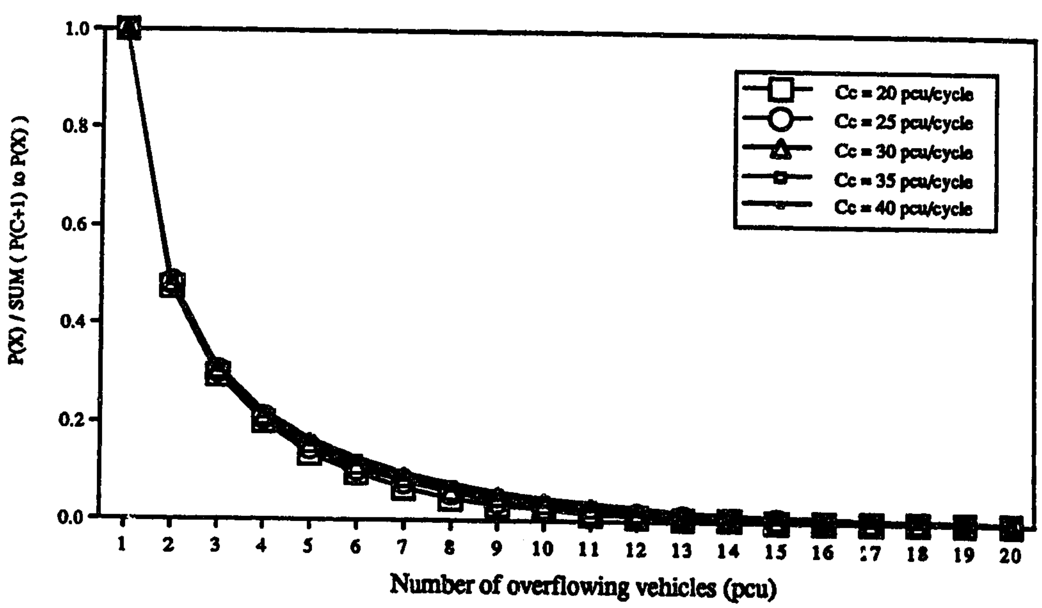


Figure 24 Significance of individual probability of arrival to total probability of overflow

Figure 24 is shows the ratio of the probability of arrival of an overflowing vehicle to the total probability of overflow. It shows that the difference between the probability of arrival of two successive overflowing vehicles approaches zero very rapidly as the number of overflowing vehicles increases. This means that since the difference between the probability is so small, including an extra overflowing vehicle will not have much impact from that if the extra vehicle is left out from the delay prediction. This assumption is used to determine the number of overflowing vehicles significant to random overflow delay predictions.

Figure 24 also shows that the significance of the probability of an extra overflowing vehicle relative to the total probability of overflow is negligible as of the fourth or the fifth ov flowing vehicle.

Therefore, for the purpose of this research, the fourth overflowing vehicle will be considered as the last overflowing vehicle that is still significant to the prediction of random overflow delay. This is the number used as the upper boundary for the summation of the probability of random overflow. Thus,

$$P(X > C_c) = \sum_{x = c_c + 1}^{x = c_c + 4} \frac{V_c^x e^{-V_c}}{X!} \qquad (13a)$$

where $P(X>C_c)$ = total probability of random overflow

V_c = average arrival per cycle (pcu/cycle)

C_c = capacity per cycle (pcu/cycle)

X = actual number of arrival

However, in actual traffic conditions, random overflow does not always occur in only one cycle. In some situations, it may occur in successive cycles. In order to have a model that is a good representation of the actual traffic condition, this factor must be incorporated into the random overflow delay prediction model.

Figure 25 is a tree diagram illustrating the possible occurrence of successive cycles of random overflow. It shows that there are infinite number of possibilities. For example, consider the worst possible case where there are four random overflowing vehicles in the first cycle. For random overflow to occur in the second cycle, there must be Cc - 3 to Cc + 4 vehicles arriving in the second cycle. For random overflow to occur in the third cycle, there must be Cc - 7 to Cc + 4 vehicles arriving in the third cycle, and so on. However, similar to the overflowing vehicles, the significance of the extra vehicle of underflow is limited after a few underflowing vehicles. Using the same technique as in the case for overflow, it is found that the limit of significance of underflowing vehicles is also four vehicles. This number is then used as the lower boundary for the summation of the probability of consecutive cycles of overflow.

Figure 25 also shows that there is a chance for the random overflow to occur in many successive cycles, so it is necessary to consider all the cycles. However, Table 6 shows otherwise. Table 6 tabulates the probability of consecutive cycles of random overflow for the worst possible case (Cc + 4 vehicles of overflow in the first cycle, and with the limit of four overflowing or underflowing vehicles in the subsequent cycles) .

Table 6 Probability of overflow in consecutive cycles

Number of cycles with random overflow	number of arrivals that will cause overflow	probability of arrival	probability of overflow
1	$C_c + 4$	0.0529	0.0529
2	$C_c - 3$ to $C_c + 4$	0.5879	0.0311
3	$C_c - 4$ to $C_c + 4$	0.5888	0.0080
4	$C_c - 4$ to $C_c + 4$	0.5888	0.0011

For the case with the highest probability of occurrence, the chances of overflow for three consecutive cycles is low, 0.80%. The omission of the possible occurrence of three or more consecutive cycles of overflow in delay predictions seems justified. It is found that for all possible cases of random overflow in the first cycle, with either 1 to 4 overflowing vehicles, the probability of random overflow in three consecutive cycles is 1.07%. It is insignificant.

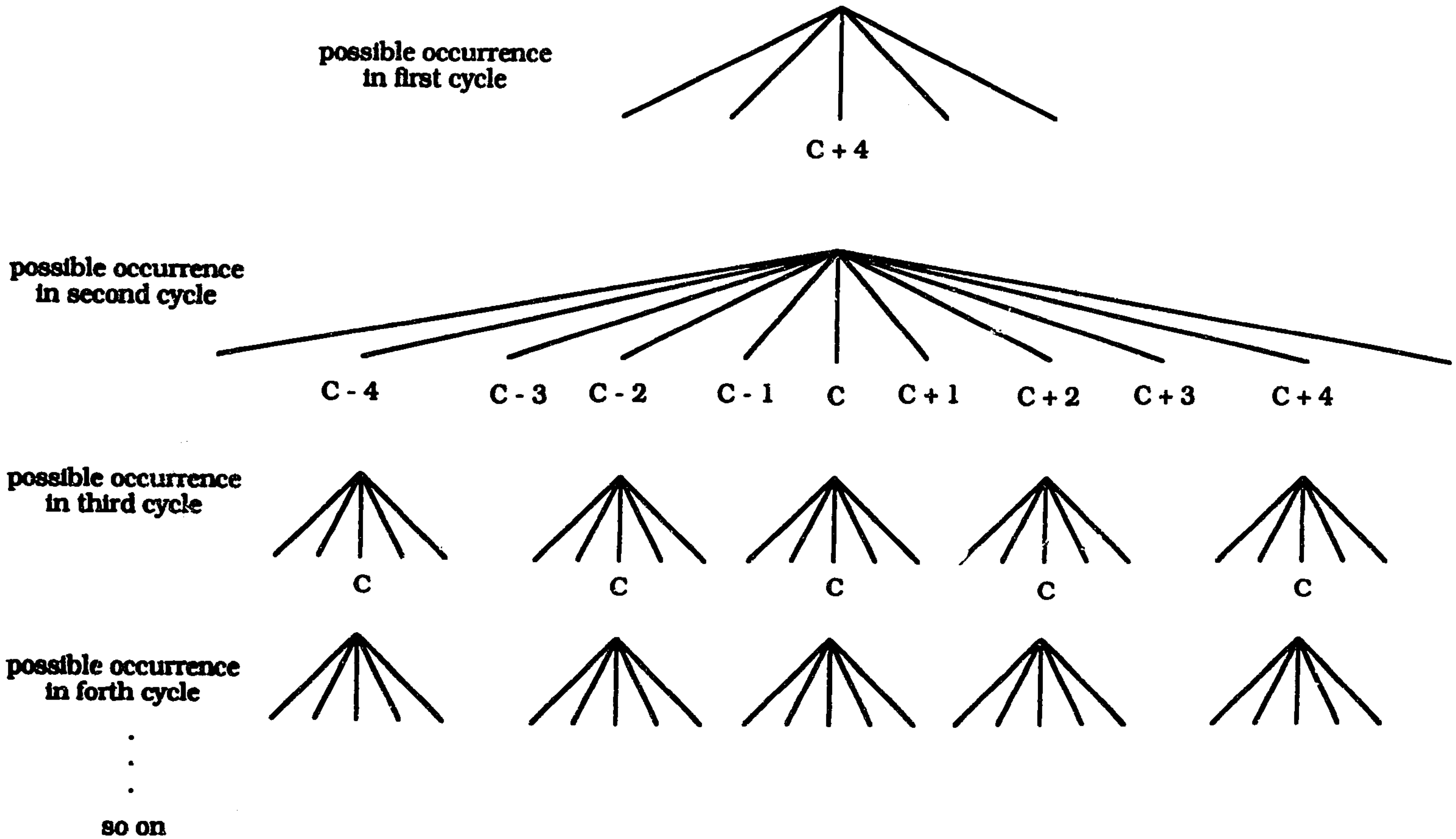


Figure 25 A tree diagram illustrating the possible occurrence of random overflow in the first, second and subsequent cycle

Considering the possible random overflow of either 1 to 4 vehicles in the first cycle and followed by a second cycle of random overflow of either 1 to 4 vehicles, the following equation can be derived. The total probability of random overflow in two successive cycle is

$$P(X>C_c) = \sum_{X=C_c+1}^{X=C_c+4} \frac{V_c^X e^{-V_c}}{X!} + \sum_{X=C_c+1}^{X=C_c+4} \frac{V_c^X e^{-V_c}}{X!} \sum_{Y=C_c-4}^{Y=C_c+4} \frac{V_c^Y e^{-V_c}}{Y!} \qquad (14a)$$

or

$$P(X > Cc) = P(X = {}_1C_{c4}) + P(X = {}_1C_{c4}) \; P(Y = {}_{-4}C_{c4}) \qquad (14b)$$

where $P(X>C_c)$ = probability of random overflow

V_c = average arrival per cycle (pcu/cycle)

C_c = capacity per cycle (pcu/cycle)

X = overflowing vehicle in the first cycle (pcu)

Y = overflowing vehicle in the second cycle (pcu)

5.2.3 Expected Delay

The prediction of expected delay are discussed according to the specific cycle of the vehicle arrivals.

Vehicles arrived in the first cycle of random overflow

Based on the assumption of a maximum of four overflowing vehicles in the first cycle of random overflow, a time space diagram illustrating the queueing situation at a signalized intersection can be produced (shown in Figure 26). This figure can be used to approximate additional delay experienced due to the random overflow.

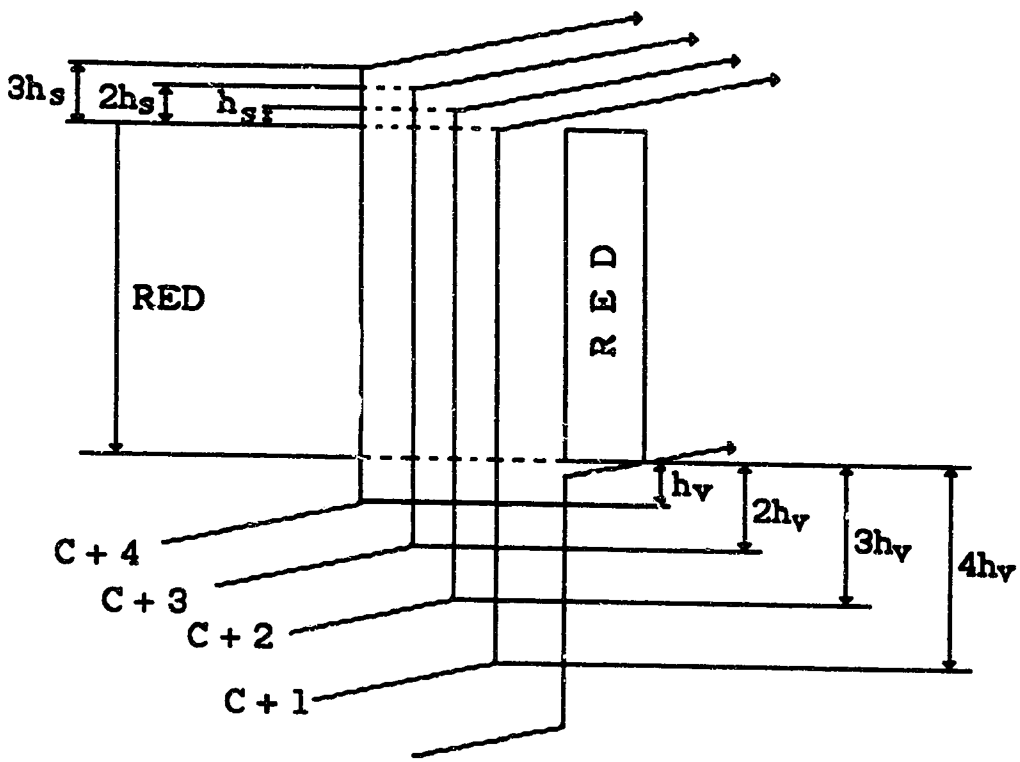


Figure 26 A time space diagram showing the queueing situation at a signalized intersection. The diagram is used for the determination of the excepted random overflow delay in the first cycle. The diagram is drawn based on actual queueing at an intersection. Note that the diagram is not to scale.

From Figure 26, the additional delay experienced by each overflowing vehicle can be determined, see Table 7.

Table 7 Expected extra delay in the first cycle during a random overflow situation

Vehicle number	Expected delay
C + 1	red + $4h_v$ + $0h_s$
C + 2	red + $3h_v$ + $1h_s$
C + 3	red + $2h_v$ + $2h_s$
C + 4	red + $1h_v$ + $3h_s$

C + 5 - it is the first vehicle in the second cycle where red = red interval (including amber) (sec)

C_c = capacity per cycle (pcu/cycle)

h_s = saturation flow headway (sec)

h_v = average per cycle arrival headway (sec),

$h_v = 3600/V_c$ The headway used for delay calculation is the actual arrival headway. However, as the actual arrival headway is not available, the average per cycle arrival headway may be used since for random overflow to occur, the average actual headway in the overflowing cycle has to be less than the average per cycle arrival headway. Thus, with the use of h_v, it already results in an overprediction of individual delay.

Averaging the expected delay, yields

$$\text{Average delay, } t_d = \text{red} + 2.5\,h_v + 1.5h_s \qquad (15)$$

<u>Vehicles arrived in the second cycle</u>

The calculation of the expected delays for vehicles that arrived in the second cycle of overflow is complicated. The complication is caused by the fact that the expected delays for the vehicles arrive in the second cycle of random overflow are dependent on the actual number of vehicles arrived in the second cycle and the average arrival flow. Since the actual number of arrivals in the second cycle is unknown, concise calculation of the expected delay for the second cycle is impossible. Therefore, for practical purposes, the average expected delay in the second cycle is assumed to be equal to that in the first cycle. This assumption is justified because in most cases, the saturation flow is greater than the arrival flow, and the additional delay experienced in the second cycle would be less than that in the first cycle. This occurs especially in the case where the vehicles arrived in the first cycle are assumed to have waited through an entire red interval or more. This is based on the fact that since h_s is less than h_v, at some particular time within a signal cycle, the number of vehicles that can be discharged from the queue is equal to or greater than the number of vehicles that actually arrive and queue. Any vehicle which arrives subsequent to this time can go through the intersection without delay. Therefore, the average extra delay for vehicles in the cycle where the overflow queue dissipated is less than that of the previous cycle of random overflow. As shown in Figure 26, the delay for vehicle 1 is greater than the delay for vehicle 2. The delay for vehicle 2 is greater than the delay for vehicle 3, and so on. The average expected delay will be less than that of the first cycle. The assumption that the expected delay in the second cycle is equal to that of the first cycle is a conservative measure and justifiable.

5.2.4 Final Random Overflow Delay Model

Combining the probability term and the expected delay term by using the principle of expected value, Equation (16) is developed.

$$D_{RO} = P(X={}_1C_{c4})\ t_d + P(X={}_1C_{c4})\ P(X={}_{-4}C_{c4})\ t_d \qquad (16)$$

where D_{RO} = average random overflow delay (sec)

$P(X={}_1C_4)$ = probability of having four vehicles more than capacity arrived in the first cycle

$P(X=-4C_4)$ = probability of having four vehicles less than or more than capacity arrived in the second cycle

t_d = average expected delay (sec)
= red + $2.5h_v$ + $1.5h_s$

red = red interval of signal plan (sec)

h_v = average per cycle arrival headway (sec)

h_s = discharge/saturation flow headway (sec)

Equation (16) is the probabilistic approach to random overflow delay prediction. It is a cycle-time oriented, time independent, and probabilistic model.

Applications

Equation (16) is only applicable to individual fixed time signalized intersection. It can be applied over all ranges of degree of saturation. At low degree of saturation, the probability of overflow is close to zero. The average random overflow delay is also close to zero. This can be seen from Table A in Appendix A. At high degree of saturation, the probability of random overflow is constant and significant. However, the average random overflow delay is insignificant when compared to the average continuous overflow delay, as shown in Table A in Appendix A and Figure 12 in Chapter 3.

5.3 CONTINUOUS OVERFLOW DELAY

The advantage of predicting overflow delay expected as compared to the delay experienced has led to the modification of the Whiting continuous overflow delay equation. The modification consists of a different derivation principle, one which allows for the estimation of the expected delay for vehicles arriving at the intersection within the evaluation period.

The main differences between Whiting's derivation of the continuous overflow delay equation and the modified approach are that the delay calculation includes the entire delay of all the vehicles arrive within the evaluation time (i.e. area QRS is also included). The average delay is obtained by the division of the number of vehicles contributing to the new total delay, that is Vt_e vehicles instead of Ct_e. As a result, some of the delay experienced after the evaluation period are included in the average delay calculation.

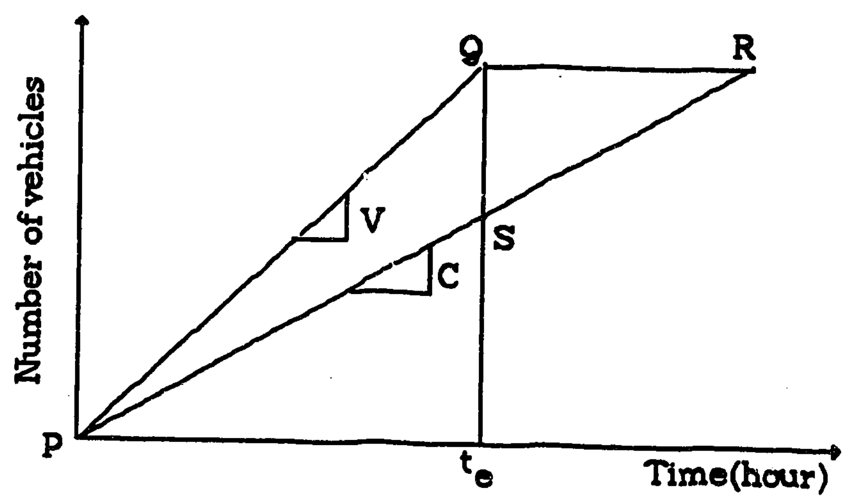


Figure 27 Queueing diagram used for derivation of the continuous overflow delay equation

The derivation of the modified continuous overflow delay equation is as follows:

With the difference in time between the arrival and the discharge curves varies linearly, the average delay for vehicles arrived during the evaluation period is half of the length of time QR. Therefore the average continuous overflow delay, D_{co} is:

$$D_{co} = \frac{1}{2} QR \quad = \frac{1}{2} \frac{QS}{C}$$

$$= \frac{1}{2} \frac{(V - C)\, t_e}{C}$$

converting t_e into minutes:

$$D_{co} = \frac{30\, t_e}{C} (V - C)$$

However, if $V < C$, then $Dco < 0$, and this is unrealistic. A modifier is needed. With some mathematical manipulations, the following equation can be written

$$D_{co} = \frac{15 t_e}{C} \left[(V - C) + \sqrt{(V - C)^2} \right] \quad (17)$$

Equation (17) is used to predict the average continuous overflow delay expected for vehicles that arrive at the intersection within the evaluation period. The equation is applicable to only individual fixed time signalized intersection. It is derived based on the same assumptions as in Whiting equation. It differs from Whiting only in the sense that Whiting equation is derived based on the vehicles discharged within the time period.

Note that Equation (17) is identical to the Whiting equation. Although the Whiting's and the modified approach are different, they result in the same continuous overflow equation because of the geometry of the queueing diagram.

5.4 EVALUATION TIME VERSUS CONGESTION PERIOD

In the modified model, the evaluation time has been redefined as the flow persisting time. Flow persisting time is the duration in which the overflow arrival persists. It is denoted as t_p in Figure 28. This definition for the evaluation time originated by Akcelik [12]. The redefinition allowed the derivation of a relationship between the flow persisting time and the congestion period based on the geometry of a queueing diagram.

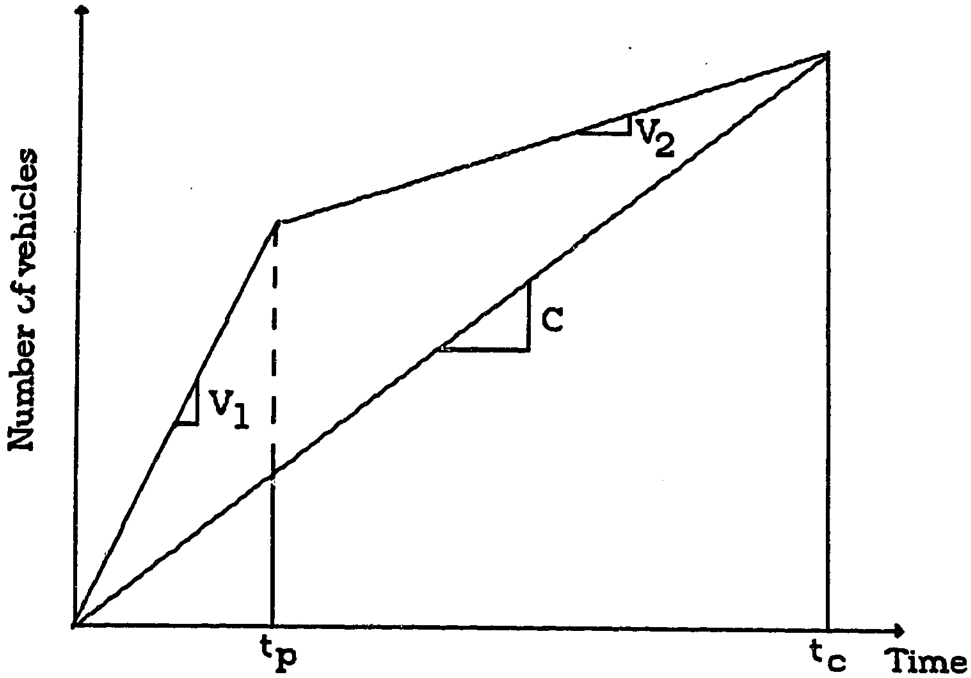


Figure 28 Queueing diagram used for relating evaluation time to congestion time

where V_1 = initial hourly arrival flow (pcu/h)

V_2 = final hourly arrival flow (pcu/h)

C = hourly capacity (pcu/h)

t_p = flow persisting time (min). The duration in which the arrival is greater than the capacity.

t_c = congestion period (min). The duration in which there is overflow.

By applying the principle of conservation of matters and energy, which says that the number of vehicles going into a system must be equal to the number of vehicles coming out from the system, if there is no source or sink inside the system the following relation is obtained.

$$V_1 t_p + V_2(t_c - t_p) = C t_c \qquad (18a)$$

rearranging,

$$t_p = \frac{C - V_2}{V_1 - V_2} t_c \qquad (18b)$$

Note that from the equation:

a) t_p is a function of V_1, V_2, C and t_c

b) V_1, V_2, and C are known as they are usually measured

c) t_c can be known from local conditions

The time at which the traffic reaches the peak in the congestion period can now be determined.

With the definition of the flow persisting time, the evaluation time term can now be :

a) any time period -- to evaluate the average continuous overflow delay for any time lapse within the congestion period.

b) flow persisting time -- to evaluate the representative average continuous overflow delay of the entire congestion period.

The introduction of flow persisting time allows one to determine the representative overflow delay of any congestion period without having to guess the correct evaluation time.

5.5 THE MODIFIED DELAY MODEL

By combining the average uniform, random overflow and continuous overflow delay equations, an equation for predicting the average overall delay expected at an individual fixed time signalized intersection can be determined. The equation is applicable over all ranges of degree of saturation. The equation is

$$\text{Average overall delay, } D = D_u + D_{RO} + D_{CO}$$

where D_u = average uniform delay (Equation (1b))

D_{RO} = average random overflow delay (Equation (16))

D_{CO} = average continuous overflow delay (Equation (17))

CHAPTER 6 VERIFICATION OF THE MODIFIED MODEL

This chapter discusses the verification of the modified delay equation. It explains the application of the techniques applied and compares the predicted delay using the various techniques.

6.1 DELAY ASSESSMENT TECHNIQUES

The modified Whiting delay equation is verified using computer simulated delay estimations and field measurements. The results of the modified model are also compared to that of the Whiting delay equation.

6.1.1 Computer Simulation

A computer program has been developed to simulate the arrival, the queueing and the discharge patterns of traffic at the stopline of a signalized intersection. The program is designed to generate the time of arrival and discharge of individual vehicles. The difference between the arrival and discharge times of each vehicle is the delay experienced by the vehicle. The summation of all individual vehicle delay yields the total delay experienced. The division of the total delay by the number of vehicles gives the average delay.

This simulation program was described in Chapter 4 to demonstrate that random overflow delay is evaluation time independent and cycle time oriented. Therefore, the use of this program for the verification of the modified delay model may be considered to be unsuitable. However, it should be noted that the modified delay model is not designed to fit the characteristics of the computer program, but the program is merely used to justify the assumed characteristics of the modified model. In other words, it is also used as a verifier in the previous Chapter. Application of this program for verification of the modified delay model is valid for it is used only for validation of the characteristics of the model in Chapter 4, and not used as a guide to the derivation of the equation.

6.1.1.1 Program Structure

To simulate a random arrival pattern, a negative exponential function and a pseudo-random number generator are used for generating the arrival headways of vehicles. " A pseudo-random number generator is a set of functions applied sequentially to generate random numbers. One common algorithm is to pick a starting number, called a "seed number". A sequence of mathematical operations are then applied to this seed number. The number produced is the random number and it is applied as the seed number for generating the next random number. " [15] These random numbers are then used for generating the arrival headways.

The generated vehicles are discharged from the stopline at a uniform rate equal to the saturation flow. The discharge of vehicles is only allowed during the green interval.

Using the negative exponential function, a vehicle, with a headway of 'h' seconds from the start of the simulation is generated. This is the time of generation. The vehicle is assumed to be at the stopline of the intersection. The program checks if the traffic signal is green; if it is red, the time counter increases and the checking process repeats itself. This continues until the signal is green, and the vehicle is discharged. The difference between the time of generation and time of discharge is the delay experienced by the vehicle. After the vehicle is discharged, a second vehicle is generated and goes through the similar process. This entire process continues until the end of simulation period. The program then sums the individual delay and determine the average delay by dividing the total delay by the number of vehicles simulated.

Figure 29 is a flowchart illustrating the logic of the program. The program is written in FORTRAN and runs on an IBM PC or compatibles. A FORTRAN listing and samples of input and output are included in Appendix E.

6.1.1.2 Simulation Results

Operation of two different intersections has been simulated to verify the modified delay equation. The two intersections are 87 Avenue - 109 Street and Hebert Road - St. Albert Trail. For each intersection, ten trials have been made. Each trial was simulated using a different seed number for the random number generator. The average of the delays simulated from the ten trials is then used to verify the modified delay equation.

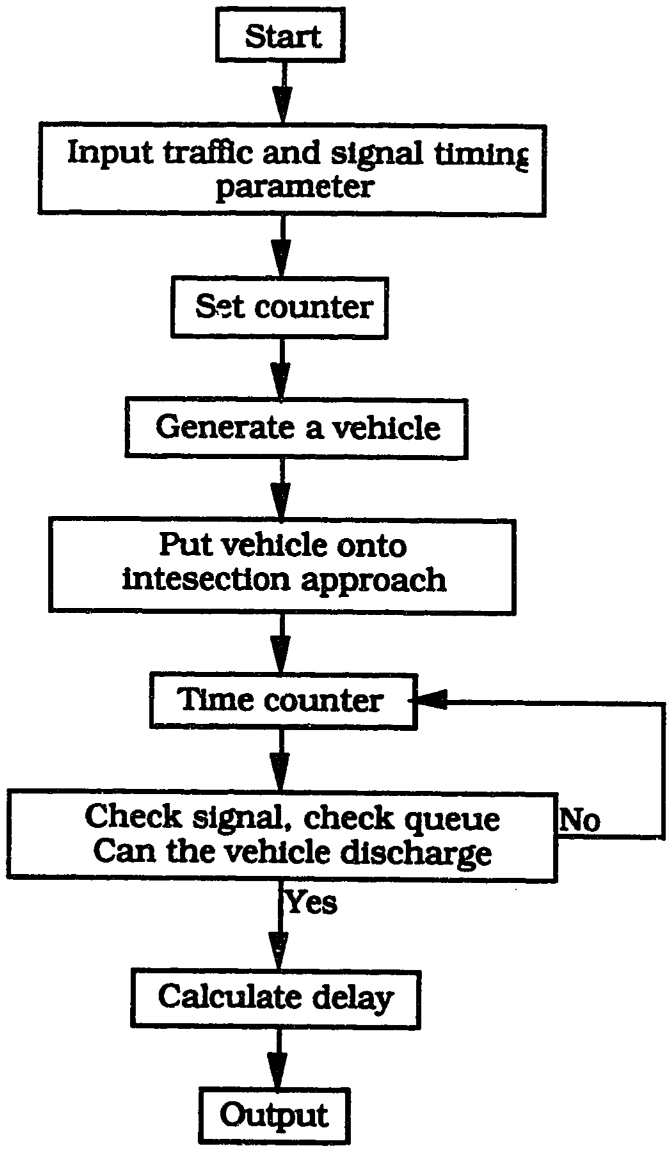


Figure 29 A flowchart illustrating the logic of the simulation program

The arrival and saturation flows and signal timing used for the simulations are those obtained from the field measurements made for verification of the modified equation. The results of the simulations are shown in Table 8.

Table 8 Delay predictions from computer simulations

	Case 1	Case 2
Arrival flow (pcu/h)	760	445
Saturation flow (pcu/h)	1700	1348
Cycle time (sec)	105	75
Effective green interval (sec)	45	25
Effective red interval (sec)	60	50
Capacity (pcu/h)	730	450
Evaluation time (min)	24	42
Number of cycles	14	34
Shortest overall delay (sec)	79.9	52.6
Longest overall delay (sec)	90.5	75.3
Average overall delay (sec)	85.8	63.9

where Case 1 : Hebert Road at St. Albert Trail

Case 2: 87 Avenue at 109 Street

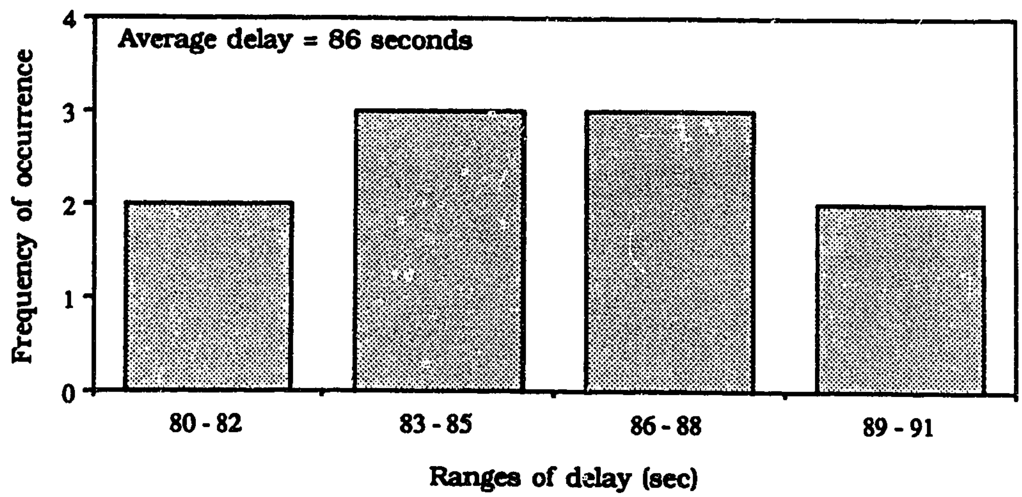


Figure 30a Delay predicted using computer simulation for Case 1: Hebert Road - St. Albert Trail

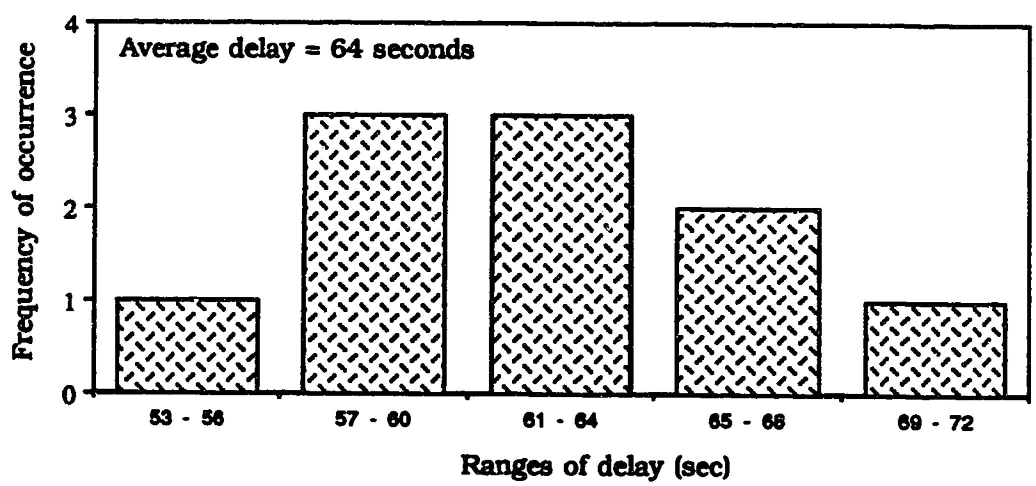


Figure 30b Delay predicted using computer simulation for Case 2: 87 Avenue - 109 Street

6.1.2 Field Measurements

Delays at two intersections were determined by surveys. For the intersection is Hebert Road - St. Albert Trail site; delay was determined using vehicle trajectory reconstruction technique [16]. At the 87 Avenue - 109 Street intersection, the delay was surveyed using the standard regular out-of-step stopped delay method. These two locations are selected based on the fact that there is very limited number of intersections in Edmonton that are operating with degree of saturation close to 1.0 at the time of survey. With these two intersections operating under the conditions similar to the simulated conditions, they were selected.

6.1.2.1 Vehicle Trajectory Reconstruction

In this survey, a traffic data time recorder is used to record the time of arrival and departure of each vehicle during the survey period. The recorder is a modified and re-programmed DATAMOS traffic classification counter. The time recorded is then compiled into a computer and the trajectory of each vehicle is reconstructed by using computer programs. These programs were written by Evans for his Master of Science degree at the University of Alberta [16]. The following description on the principles behind the vehicle trajectory reconstruction delay surveys are summarized in a portion of Evans' thesis.

In the field, two reference points are selected. The downstream reference point is at the stopline, while the upstream point is some distance away.

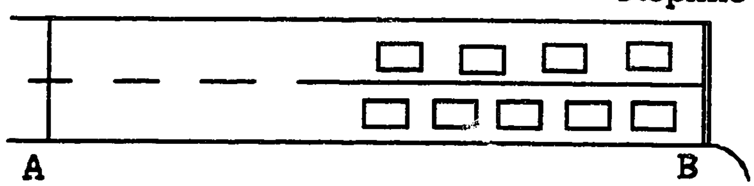


Figure 31 The reference points in delay survey (from [16])

The distance between A and B must be long enough so that the maximum queue at the intersection does not exceed this distance during most of the evaluation time. This requires some field observations before any actual measurements. If the distance is too short, the queue may form beyond the reference point, and vehicles that joint the queue upstream of the upstream reference point will not be recorded until some time after they cross the upstream reference point. This results in underpredicting the individual delay. If too long a distance is used, sight limitations of the surveyors may come into effect.

With the reference points set up, the surveyors can begin the survey. The surveyor watching the upstream point records the time that a vehicle crosses the reference point, while the person watching the stopline does the same when any vehicle crosses the stopline. The time data are then transferred into a computer. Using the compilation and processing programs written by Sabourin and Evans [16,17], the trajectory of each vehicle is reconstructed. The principle employed for the reconstruction is shown in Figure 32.

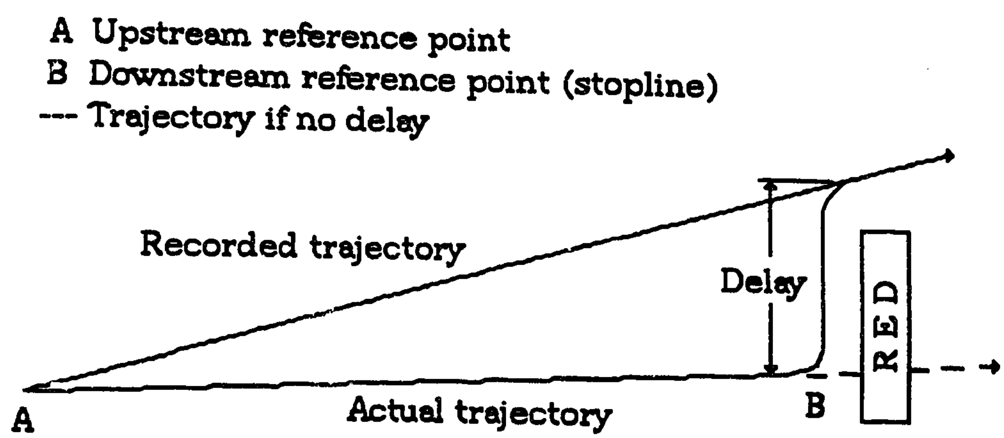


Figure 32 Principles of delay predictions by vehicle trajectory reconstruction (from [16])

By specifying the average free flow speed at the stretch of the road where surveys are done, the programs can then calculate the difference in time between the recorded time of discharge and the time of discharge if no interference existed. Average free flow delay speed is the traveling speed if no interference exists. The average of all the individual delays is equal to the average delay experienced at the intersection within the time period of evaluation.

6.1.2.2 Stopped Delay

The standard intersection stopped delay surveying technique is the most common means in delay surveys. The technique is the regular out-of-step queue count delay survey and is well explained in the Highway Capacity Manual [13] and Manual of Traffic Engineering [17]. The following is a summary of the technique.

The technique involves "the counting of number of vehicles stopped in the intersection approach at periodic intervals (such as every 15 seconds). This sampling, along with a volume count, provides estimates of the vehicle-seconds of stopped delay with considerable accuracy if the sampling is properly selected (not as even subdivision of the length of signal cycle) " [17].

The average stopped delay is converted into average overall delay using a technique outlined in the paper titled 'Accuracy of Delay Surveys at Signalized Intersections' [3] :

$$\frac{D}{D_s} = \frac{r^2}{(r - t_d)^2}$$

where
- D = Average overall delay
- D_s = Average stopped delay
- r = red interval
- t_d = deceleration delay

6.1.2.3 Results

Case 1 (Hebert Road - St. Albert Trial) : A vehicle trajectory reconstruction delay survey was conducted for a period of 24 minutes (14 cycles). The time of arrival, the time of departure, the cycle time, the red interval, the green interval and the arrival were recorded.

Saturation flow and the capacity are then calculated from the survey results.

Case 2 (87 Avenue - 109 Street) : A regular out-of-step stopped delay survey with ten second intervals has been conducted for a total time period of 42 minutes (34 cycles). The number of vehicles in queue is noted at the end of every interval. Then, the total number of vehicles in queue is counted; and dividing the total delay by the number of vehicles counted, the average stopped delay is obtained. The conditions for the two cases are tabulated in Table 9.

Table 9 Delay predictions from field measurements

		Case 1	Case 2
Type		vehicle trajectory	queue count
Measured arrival flow (pcu/h)		760	445
Measured saturation flow (pcu/h)		1700	1350
Cycle time (sec)		105	75
Effective green interval (sec)		45	25
Effective red interval (sec)		60	50
Capacity (pcu/h)		730	450
Evaluation time (min)		24	42
Number of cycles		13	34
Average overall	: measured	78.3	---
delay (sec)	: calculated	---	65.8

Case 1 : Hebert Road at St. Albert Trail

Case 2: 87 Avenue at 109 Street

6.2 DISCUSSIONS OF RESULTS

The Whiting and Modified Whiting delays are calculated using analytical delay equations. The actual, corrected and computer simulated delays are from the previous sections. Table 10 shows the results from various prediction models. The delays are also compared graphically in Figure 33 and 34.

Table 10 Summary of delay predictions from various estimation models

Case #		1	2
Degree of saturation		1.04	0.99
AVERAGE OVERALL DELAY (sec)	Actual	78.3	65.5
	Corrected	---	68.5
	Whiting	91.3	75.5
	Modified Whiting	90.9	61.8
	Computer simulation	85.8	63.9

Figure 33 and 34 are bar charts of the delay assessed from the various technique used. They show good compatibility between the delay assessed by using field measurements, computer simulations, Whiting delay equation and the modified delay equation.

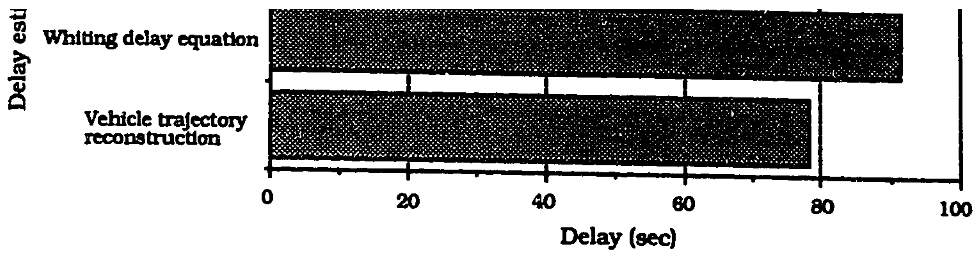


Figure 33 : Comparisons of delay estimations from various models
Case 1 : Hebert Road at St. Albert Trail

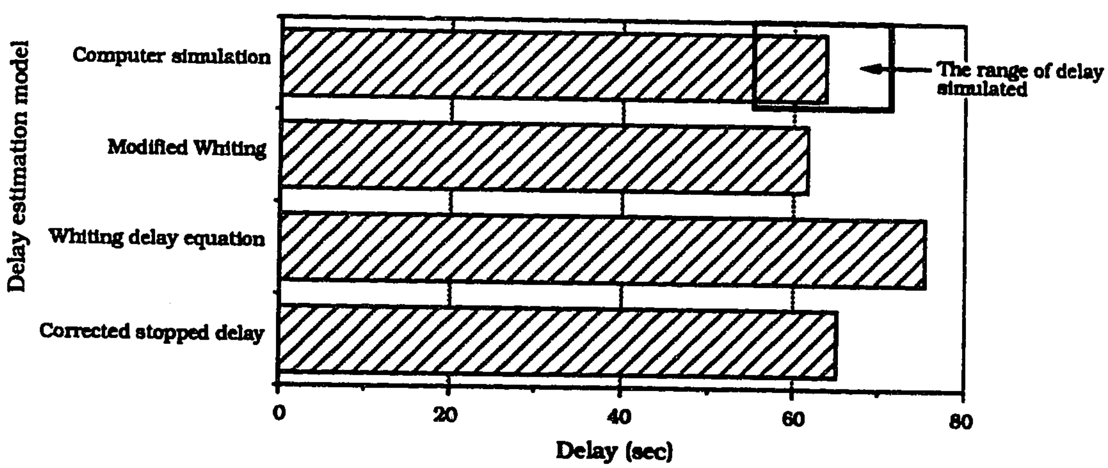


Figure 34 : Comparisons of delay estimations from various models
Case 2 : 87 Avenue at 109 Street

Case 1 (Hebert Road at St. Albert Trail) : The results show good compatibility among the three models. The differences are small. The results show that both the Whiting formula and the modified model are over predicting delay when compared to the simulation model and especially when compared the actual traffic conditions.

Case 2 (87 Ave. at 109 St.) : In this case, the results show a greater variation between the three models. The field measurements and the computer simulation model have predicted delays that are close to each another. The results are slightly higher than that of the modified model. However, the Whiting model prediction is significantly higher than that of the other three delay assessments. For practical purposes, the results obtained can be considered to be compatible.

The results show that the modified Whiting delay equation predicts delays with similar if not better accuracy than the Whiting delay equation. However, due to the limited number of field measurements available for the verification of the modified equation, no further conclusions can be made.

CHAPTER 7 CONCLUSIONS

Two objectives were set for this research. They are:

a) to identify the relationship between the evaluation time and the congestion period.

b) to investigate a probabilistic approach to random overflow delay estimations.

Both of these objectives were achieved.

Evaluation Time

In this thesis, the evaluation time in Whiting delay equation has been redefined as the flow persisting time. Flow persisting time is the duration the overflow arrival persists. This redefinition has made the representative delay predictions of any congestion period possible without having to guess what the correct evaluation time ought to be.

Delay Expected versus Delay Experienced

The modified Whiting overflow delay model uses a derivation principle that leads to the estimation of the delay expected foe vehicles arriving within the evaluation period.

Due to the geometry of a queueing diagram, the derivation principle which allows for prediction of expected delay, used in the modified equation, results in a continuous overflow delay equation that is similar to the Whiting equation, which is one that uses a derivation principle that predicts delay experienced. This is due to the geometry of the queueing diagram. Even though the estimation of delay expected is more practical, it is the same as the delay experienced predicted from Whiting equation. A numerical improvement has not been observed.

Probabilistic approach to delay estimations

The new approach to overflow delay prediction involves a probabilistic model for estimating the random overflow delay. The probabilistic model of random overflow delay differs from the Whiting model is that it is time independent, and it includes the effects of the absolute value of cycle time.

Even though the modified model appears to be more accurate representation of the actual traffic conditions, it does not predict delays that are much different from the Whiting delay equation. Therefore, it can be concluded that although random overflow is time independent and is influenced by the cycle time, a model that incorporates both of these factors has not improved the quality of delay predictions. Although the Whiting equation excludes the absolute value of cycle time, this research shows that the quality of delay predictions of Whiting delay equation is not affected. The Whiting random overflow delay model was derived empirically as a connector of the uniform delay model and continuous overflow delay model.

Recommendations

Due to the limited amount of field measurements available for verification, no definite conclusion can be made.

A suitable follow-up to this research would be a further verification of the modified delay model with more field measurements and computer simulations. The verification should include the testing of the random overflow model over a wide range of degree of saturation and evaluation time. Also, the validity of the flow persisting time equation should be tested.

Potential users should be surveyed for their preference of the delay prediction model. Do they prefer a more accurate but complicated, or a simple but less accurate model ?

Concluding Remarks

In conclusion, although the limited number of surveys and simulation trials has not proven that the modified model improves the quality of delay prediction when compared to the Whiting equation, it presents a thereotically cleaner approach. The modified model has three characteristics not present in Whiting's formula. Firstly, the random overflow delay prediction is based on widely assumed Poisson principle. The second characteristic is that the model allows for the prediction of delay that is representative of a congestion period without having to guess the evaluation time. The third improvements is that the model makes it possible to include the delay which will be experienced by drivers in the queue in congested systems.

REFERENCES

1. Hurdle, V., Signalized Intersection Delay Models: A primer for the Uninitiated, Transport and Road Research Report 971, Transportation Research Board, Washington D.C., U.S.A., 1984, p. 96 - 124.

2. Teply, S., (ed.) Canadian Capacity Guide for Signalized Intersections, ITE - District 7 and the University of Alberta, Edmonton, 1984.

3. Teply, S., Accuracy of Delay Surveys at Signalized Intersections, 69th Annual Meeting of Transportation Research Board, Washington, D.C., 1989.

4. Gerlough, D. and M. Huber, Traffic Flow Theory, Special Report 165, Transportation Research Board, National Research Council, Washington, D.C., 1975, pp. 117.

5. Teply, S., CIV E 614 Lecture Notes, University of Alberta, Edmonton, 1988.

6. Webster, F. V., Traffic Signal Settings, Road Research Paper, No. 39, Department of Scientific and Industrial Research, HMSO, London, 1958.

7. Gibra, I., Probability and Statistical Inference for Scientists and Engineers, Prentice Hall, Englewood Cliff, New Jersey, 1981, pp.100 - 102.

8. Stein, J., The Random House College Dictionary, Random House of Canada Limited, Toronto, 1975, pp. 379.

9. Robertson, D.I., Traffic Models and Optimum Strategies of Control - A Review, Proceedings of the International Symposium on Traffic Control Systems, Berkeley, California, 1979, Volume 1, pp. 262 - 288.

10. Miller, A. J., The Capacity of Signalized Intersections in Australia, Australian Road Research Board Bulletin, No. 3, Nunawading, March 1968.

11. Akcelik, R., The Highway Capacity Manual Delay Formula for Signalized Intersections, ITE Journal, vol 58 no. 3, Institute of Transportation Engineers, Washington, D.C., March 1988, pp. 23 - 27.

12. Akcelik, R., Traffic Signals: Capacity and Timing Analysis, Research Report ARR 123, Australian Road Research Board, Nunawading, 1981.

13 Transportation Research Board, Highway Capacity Manual, Special Report 209, Washington, D.C., 1985, pp. 9 - 31.

14. Harnett, D.L., Statistical Methods, Third edition, Addison Wesley Publishing Company, Inc., Reading, Massachusetts, 1982, pp. 197 - 204.

15 Cooper D. and Clancy, M., Oh Pascal An Introduction to Programming, W.W. Norton and Company, Berkeley, 1982, pp 225.

16. Evans, G., Vehicle Trajectory Reconstruction for Delay Measurements, Master of Science Thesis, University of Alberta, Edmonton, 1989.

17. Sabourin, P., DATAMOS User Manual, University of Alberta, Edmonton, 1986.

18. Box, P., Manual of Traffic Engineering, 4th Edition, Institute of Traffic Engineer, Arlington, 1976, pp. VI-6.

APPENDIX A TABLES

Table A Probability of X number vehicles arrived

Actual arrival,X	Average arrival per cycle 20	25	30	35	40
10	0.0058	0.0004	0.0000	0.0000	0.0000
11	0.0106	0.0008	0.0000	0.0000	0.0000
12	0.0176	0.0017	0.0001	0.0000	0.0000
13	0.0271	0.0033	0.0002	0.0000	0.0000
14	0.0387	0.0059	0.0005	0.0000	0.0000
15	0.0516	0.0099	0.0010	0.0001	0.0000
16	0.0646	0.0155	0.0019	0.0002	0.0000
17	0.0760	0.0227	0.0034	0.0003	0.0000
18	0.0844	0.0316	0.0057	0.0006	0.0000
19	0.0888	0.0415	0.0089	0.0011	0.0001
20	0.0888	0.0519	0.0134	0.0020	0.0002
21	0.0846	0.0618	0.0192	0.0033	0.0004
22	0.0769	0.0702	0.0261	0.0052	0.0007
23	0.0669	0.0763	0.0341	0.0080	0.0012
24	0.0557	0.0795	0.0426	0.0116	0.0019
25	0.0446	0.0795	0.0511	0.0162	0.0031
26	0.0343	0.0765	0.0590	0.0219	0.0047
27	0.0254	0.0708	0.0655	0.0283	0.0070
28	0.0181	0.0632	0.0702	0.0354	0.0100
29	0.0125	0.0545	0.0726	0.0428	0.0138
30	0.0083	0.0454	0.0726	0.0499	0.0185
31	0.0054	0.0366	0.0703	0.0563	0.0238
32	0.0034	0.0286	0.0659	0.0616	0.0298
33	0.0020	0.0217	0.0599	0.0654	0.0361
34	0.0012	0.0159	0.0529	0.0673	0.0425
35	0.0007	0.0114	0.0453	0.0673	0.0485
36	0.0004	0.0079	0.0378	0.0654	0.0539
37	0.0002	0.0053	0.0306	0.0619	0.0583
38	0.0001	0.0035	0.0242	0.0570	0.0614
39	0.0001	0.0023	0.0186	0.0511	0.0629
40	0.0000	0.0014	0.0139	0.0447	0.0629
41	0.0000	0.0009	0.0102	0.0382	0.0614
42	0.0000	0.0005	0.0073	0.0318	0.0585
43	0.0000	0.0003	0.0051	0.0259	0.0544
44	0.0000	0.0002	0.0035	0.0206	0.0495
45	0.0000	0.0001	0.0023	0.0160	0.0440
46	0.0000	0.0001	0.0015	0.0122	0.0382
47	0.0000	0.0000	0.0010	0.0091	0.0325
48	0.0000	0.0000	0.0006	0.0066	0.0271
49	0.0000	0.0000	0.0004	0.0047	0.0221
50	0.0000	0.0000	0.0002	0.0033	0.0177

APPENDIX B RANDOM OVERFLOW DELAY CALCULATIONS

The following is the prediction of the average random overflow delay using the principle of expected value of theory of probability.

As an example, using 20 pcu/cycle arrivals and capacity, and assumming that there is a V/C = 1.2 arrivals in the first cycle of overflow, there are four overflowing vehicles arriving. Also assuming the following arrival and discharge headways:

average arrival headway = 4 sec

saturation flow headway = 2 sec

vehicle #	probability of arrival	excepted delay	column 1 X column 2 (sec)
21	0.0846	RED + 4hv + 0hs	4.7
22	0.0769	RED + 3hv + 1hs	4.1
23	0.0669	RED + 2hv + 2hs	3.3
24	0.0557	RED + 1hv + 3hs	2.6

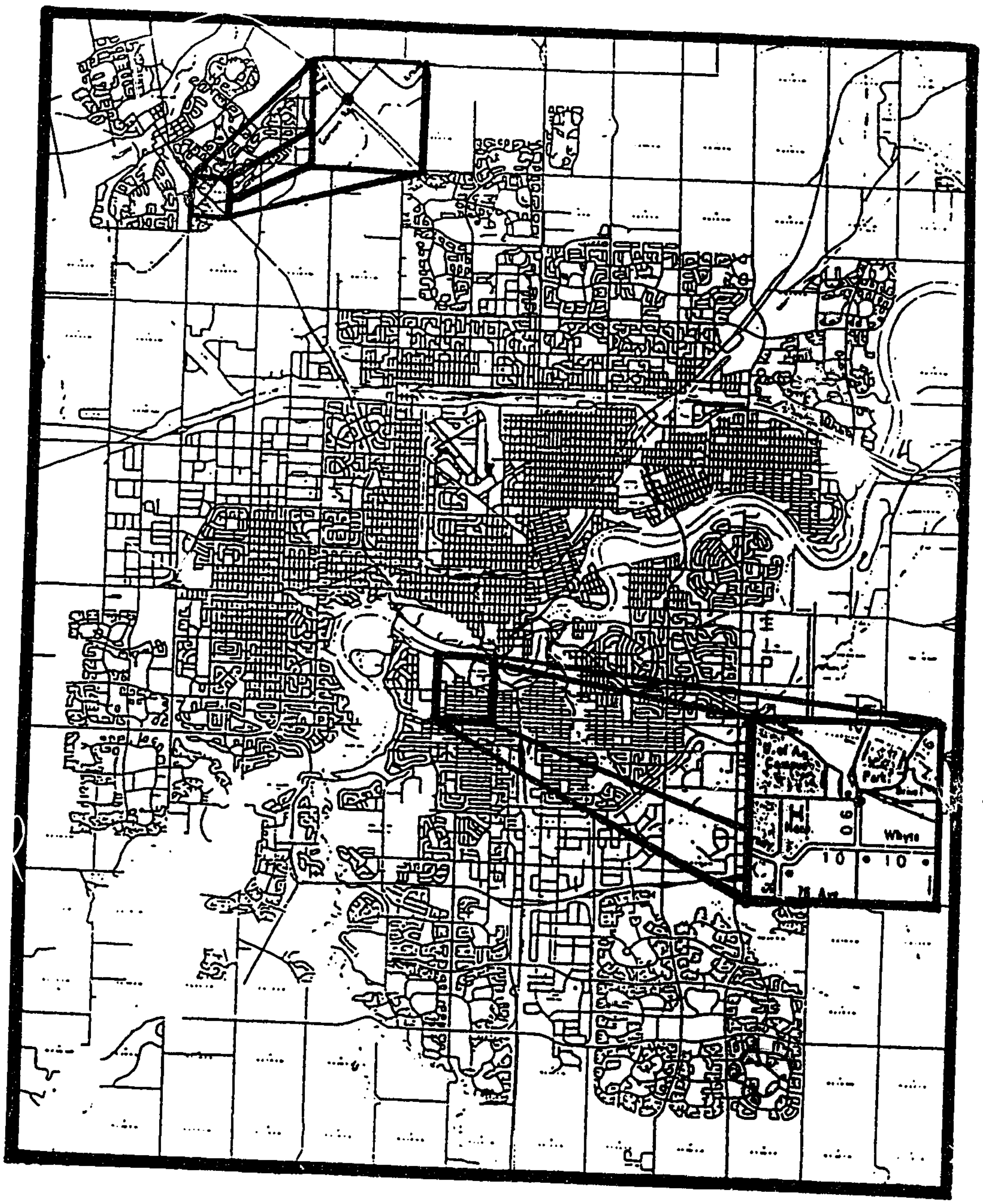
APPENDIX C

LOCATIONS

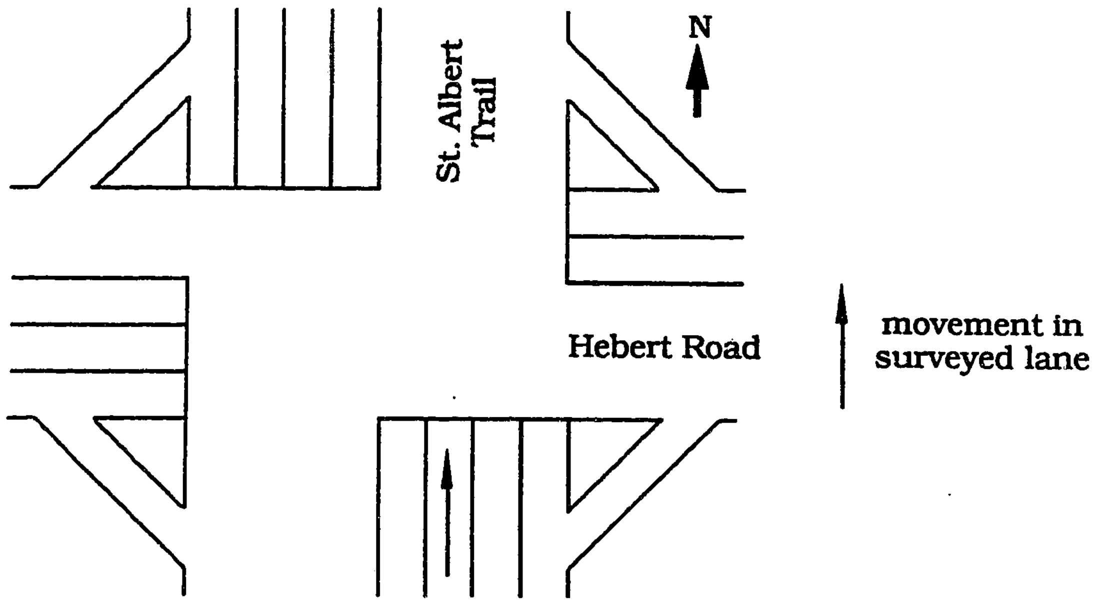
AND

GEOMETRIC LAYOUTS

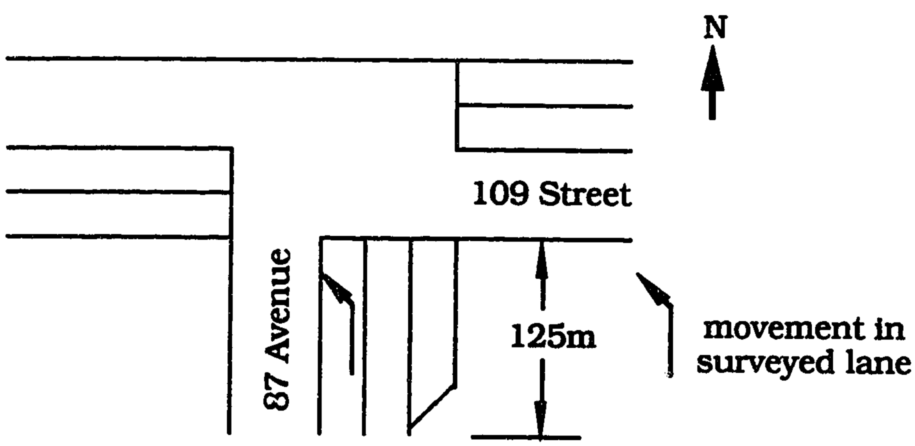
OF SURVEY SITES



City of Edmonton and City of St. Albert



Case 1 : Hebert Road at St. Albert Trail



Case 2 : 87 Avenue at 109 Street

Geometric layouts of surveyed locations

APPENDIX D SURVEY RESULTS

Case 1: Vehicle trajectory reconstruction delay study

Case 2: Stopped delay delay survey

Summary

	Case 1	Case 2
Arrival flow (pcu/h)	760	445
Saturation flow (pcu/h)	1700	1350
Cycle time (sec)	105	75
Green interval (sec)	45	25
Red interval (sec)	60	50
Capacity (pcu/h)	730	450
Evaluation time (min)	24	42
Weather :	sunny	sunny
Pavement	good	good
Traffic	peak hour	peak hour

APPENDIX E FORTRAN PROGRAM LISTING

```
C   PROGRAM SIMULATION
    INTEGER TDIS,TA,TP,T,TIME,TCHK,TSIM,TPH,TSIMH,SEED
    CHARACTER*1 ANS1
     OPEN(6,FILE='OUT',STATUS='OLD')
C
5005 WRITE(*,6001)
6001 FORMAT(2X,'INPUT THE FOLLOWINGS:'/,
   +      5X,'1. INITIAL ARRIVAL FLOW, V1, WHERE V1 > C, (PCU/H)'/,
   +      5X,'2. FINAL ARRIVAL FLOW, V2, WHERE V2 < C, (PCU/H)'/,
   +      5X,'3. CAPACITY (PCU/H)'/,
   +      5X,'4. CYCLE TIME (SEC)'/,
   +      5X,'5. GREEN TIME (SEC)'/,
   +      5X,'6. RED TIME (SEC)'/,
   +      5X,'7. LENGTH OF TIME V1 PERSISTS (H)'/,
   +      5X,'8. SIMULATION TIME (MIN)'/,
   +      5X,'9. SEED NUMBER (ODD 3 DIGIT INTEGER)')
C
    READ(*,*) V1,V2,CAP,CTIME,GTIME,RTIME,TPH,TSIMH,SEED
C
    SUM = 0
    TA = 0
    N = 0
    T = RTIME - 1
    H = 3600 / V1
    HP = 3600 / V2
    HS = 3600 / (CAP*CTIME/GTIME)
    IHS = NINT(HS)
C    THS = HS*2 + 1
C    IHS = INT(THS)
C    VHS = FLOAT(IHS)
C    WHS = VHS / 2.0
C    IHS = NINT(HS)
    TP = TPH * 3600
    TSIM = TSIMH * 60
    MTPL = 25173
    MDLS = 65536
    INCR = 13849
    SCALE = 0.70
C
9001 N = N + 1
    IF(T.GE.TP) THEN
      SCALE = 0.95
      H = HP
    ENDIF
    IPROD = (25173*SEED) + 13849
    SEED = MOD(IPROD,MDLS)
    RSEED = FLOAT(SEED)
    RMDLS = FLOAT(MDLS)
    RNDN = (RSEED/RMDLS)*SCALE
```

```
C
    HV = 2 - (H * LOG(1 - RNDN))
    IHV = NINT(HV)
    TA = TA + IHV
C
1001 T = T + 1
    IF(T.GT.TSIM) GOTO 9999
    IF(T.LT.TA) T = TA
    CYCLE = T/CTIME
    NCYCLE = INT(CYCLE)
    TIME = T - (NCYCLE * CTIME)
    TCHK = TIME - RTIME
    IF(TCHK.LT.0) GOTO 1001
      DO 4001 J=2,25
        TDIS = J * IHS
        IF(TDIS.GT.GTIME) GOTO 1001
        IF(TCHK.EQ.TDIS) THEN
          DELAY = T - TA
          IF(DELAY.LT.0) DELAY = 0
          GOTO 9002
        ENDIF
4001    CONTINUE
      GOTO 1001
9002 SUM = SUM + DELAY
    GOTO 9001
C
9999 ADELAY = SUM / (N - 1)
C
    WRITE(*,6002)ADELAY
    WRITE(6,6002)ADELAY
6002 FORMAT(5X,'AVERAGE DELAY = ',F8.2,'SEC')
C
    WRITE(*,6003)
6003 FORMAT(2X,'RERUN THE PROGRAM?')
    READ(*,2001)ANS1
2001 FORMAT(A1)
    IF(ANS1.EQ.'Y'.OR.ANS1.EQ.'y') GOTO 5005
    STOP
    END
```