UNIVERSITY OF ALBERTA

USE OF DELAY IN ACCIDENT PREDICTION AT INTERSECTIONS WITHOUT TRAFFIC SIGNALS

BY

PAUL LAP PO CHAN

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

DEPARTMENT OF CIVIL ENGINEERING

Edmonton, Alberta
SPRING OF 1994

ISBN 0-612-11173-3

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: PAUL LAP PO CHAN

TITLE OF THESIS: USE OF DELAY IN ACCIDENT PREDICTION AT INTERSECTIONS WITHOUT TRAFFIC SIGNALS

DEGREE: MASTER OF SCIENCE

YEAR THIS DEGREE GRANTED: SPRING OF 1994

(SIGNED) Paul Chan

PERMANENT ADDRESS: *6604 - 187 Street*
Edmonton, Alberta
T5T 2N2

DATED April 15 1994

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled ***Use of Delay in Accident Prediction at Intersections Without Traffic Signals*** submitted by Paul Lap Po Chan in partial fulfilment of the requirements for the degree of MASTER OF SCIENCE in CIVIL ENGINEERING

....
Dr. S. Teply, Supervisor

....
Dr. L.R. Rilett, Chairman

....
Dr. S.M. AbouRizk

....
Dr. J.D. Whittaker

Date April 15, 1994

ABSTRACT

Priority rules and traffic signs regulate the competition of vehicles for space and time at intersections without traffic signals. Potential vehicle conflicts may lead to accidents. In many municipalities, intersections are ranked by the number of accidents that occur annually, or by accident rates, which is usually defined as the number of accidents over million-entering-vehicles. The objective of these ranking methods is to identify intersections with safety problems in order to determine priorities regarding safety analysis and improvement measures.

This thesis summarizes the results of a research project that examined traditional methods of intersection safety ranking and investigates the potential of using delay as the underlying variable in accident prediction models. As delay is experienced directly from the driver's standpoint, it was postulated that delay could be a better parameter in reflecting the risk-taking behaviour of drivers and therefore might result in better accident prediction modelling results.

More than 21,000 vehicle arrivals were recorded and analyzed at 26 priority-ruled T-intersections in Edmonton. The analysis results were also compared with four years of accident data. Additional delay information was generated by a simulation program. The complexity of accident analysis required non-standard analytical techniques.

It was demonstrated that the most common method of ranking intersections by accident ratios, based on the sum of entering volumes, had no theoretical justification and little practical value. The method which related accident frequencies to the product of conflicting traffic streams gave a good indication of the degree of safety. The best model was derived on the basis of total delay as a surrogate measure of the risk to which the drivers were exposed.

ACKNOWLEDGMENTS

I wish to acknowledge the contribution of all individuals and organizations who made the completion of this thesis a reality.

The research was financially supported through an operational research grant from the Natural Sciences and Engineering Research Council of Canada.

I would also like to express my gratitude to Professor W. Brilon and Dr. M. Grossmann of the Ruhr University in Bochum for providing the KNOSIMO intersection program and advice; Ms. K. Bustin for assisting with the translation of the KNOSIMO manual, Dr. T. Fung of the University of Calgary for help with the GLIM computer package, and the corporation of Mr. R. Strynadka and Mr. B. Shacker of the City of Edmonton with regard to accident and other data.

My most heartfelt thanks are expressed to Dr. Stan Teply whose advice, guidance and encouragement contributed to the successful completion of this thesis.

Finally, my deepest gratitude to my family, for without their unbounded patience and understanding none of this would have been possible.

TABLE OF CONTENTS

APPENDIX A	Simulation of Multilane Stop-controlled T-intersections by KNOSIMO in Canada - Paper presented in the Second International Conference on Intersections Without Traffic Signals, Bochum, Germany. Published in Intersection Without Traffic Signals II, 1991 p. 308-319.
APPENDIX B	Goodness-of-fit Plots and Summary of Parameter Estimates for All Models. GLIM Program Listings and Outputs for Selected Models
APPENDIX C	Correlation Plots of Parameters in the Best Models
APPENDIX D	Traffic Safety at Intersection Without Traffic Signals - Paper presented at the 1993 Annual Conference of the Institute of Transportation Engineers District 7 - Canada in Edmonton Compendium of Technical Papers. p.169 - 186

LIST OF TABLES

LIST OF FIGURES

Figures **Page**

1. INTRODUCTION

1.1 Problem Statement

Considerable effort has been directed towards the monitoring of roadway safety Accident records are often used to determine priorities for safety improvements. The success of a safety improvement program relies heavily on the techniques used in identifying locations with high accident potential.

The parameters most commonly used to quantify non-safety are accident frequency and accident rate. They reflect two different aspects of the situation of non-safety. Nevertheless, safety improvement priority lists developed on the basis of these two parameters often do not agree totally with each other.

An alternate approach is to use accident prediction models to reflect the actual safety condition. These accident prediction models require some knowledge about the study locations, such as the amount of traffic or the geometric features of the study locations. Most of the accident prediction models developed in the past, use some combinations of the sums or products of traffic volumes as model parameters.

Accidents are events with a high degree of randomness with numerous factors involved. Factors identified as possible contributors to accidents can be classified into three main groups: driver factors, vehicle factors and environment factors. Driver factor is generally considered the most significant of all contributing issues but also the least predictable. A number of in-depth accident studies (Sabey & Staughten, 1975 and Sabey, 1983) undertaken by "on-the-spot" multi-disciplinary investigation teams concluded that the driver element was dominant in 95 % of roadway accidents.

As suggested by the in-depth accident studies, the driver factor appears to be the major contributor to traffic accidents. It is therefore reasonable to assume that an accident prediction model would have good accident prediction abilities if it contains model parameters that can reflect driver behaviour.

A parameter that is relatively easily to measure and can reflect driver behaviour well is traffic delay experienced by the drivers. Delay is often used in other applications to reflect driver behaviour. For instance, it is a primary parameter used in transportation planning models to simulate the route selection characteristics of drivers. It is also used in a number of intersection capacity analysis methods to reflect the perception of the drivers on the traffic operations at the intersections. It is expected that delay can reflect the effect of human elements in accidents more successfully than other typically used accident prediction modelling parameters, such as traffic volumes.

1.2 Research Objectives

Delay, as a good measure of driver behaviour, has the potential to be a more effective model parameter than the traditionally used traffic volume in accident prediction. Therefore, the primary objective of this research is to investigate the potential of including delay as one of the underlying variables in accident prediction models.

The secondary objective of the research is to examine the traditional methods of intersection safety ranking. A ranking list was developed based on the accident prediction modelling results from the current study. The list was then compared to the ranking lists developed by using the traditional intersection safety ranking methods.

1.3 Scope

To reduce the number of influencing variables in the accident prediction model, the scope of the research was limited to T-intersections, the simplest type of intersection. Study locations were chosen along 4-lane urban arterials at intersections with similar design features to minimize the degree of variability within the data set. The analysis period for the study was limited to weekdays and took into account only daytime accidents. Traffic data were collected at 26 locations on weekdays between 11 a.m. and 4 p.m., which was a period of relatively stable traffic conditions and accident rates.

Field traffic volume, delay, and headway data were collected in January and February, 1990 at 26 selected study locations. Additional delay data were obtained through computer simulation. A total of 3,795 accident records from 1985 to 1988 for 429 priority-ruled intersections were also obtained and analyzed.

1.4 Organization of the Thesis

Chapter 2 contains an in-depth summary of previous accident prediction modelling efforts. The focus is on the accident prediction methodology used and the various safety and model variables selected. A general discussion of accidents is included. The chapter also includes a discussion of various techniques used in the research project for data encoding, data simulation and statistical modelling.

Chapter 3 examines the general accident trends in the City of Edmonton between 1985 and 1988. Results of this preliminary analysis are used to determine the analysis period for the study and to formulate the framework of the modelling process for the study. Elements of the research framework, such as the data collection procedures, the criteria set for the selection of study locations, analysis periods, and the different basic model parameters, are discussed in detail.

Chapter 4 documents the development of various types of accident prediction models. Explanations are given on the approach adopted in developing the model structures.

Chapter 5 compares the results of the various models developed. Different approaches in ranking high accident locations are compared.

Chapter 6 summarizes the research findings. Conclusions and practical implications of the study are provided.

Appendix A includes a paper by the researcher and his supervisor, Dr. Stan Teply, on the application of KNOSIMO. KNOSIMO is a traffic simulation program that was used to simulate traffic delays at the study locations.

Appendix B provides additional information on GLIM. GLIM is the statistical analysis program used in the modelling process. The goodness-of-fit plots for all the models are provided. A summary of parameter estimates for all the delay-based accident prediction models is also provided.

Appendix C contains the correlation plots of selected parameters of the most suitable models.

Appendix D contains a paper co-written by the researcher and his supervisor on the findings of the research project.

2. LITERATURE REVIEW

This chapter consists of four main sections. The first section contains the findings of a literature search on previous accident prediction studies. The focus is on the approaches used to formulate accident prediction models.

The second section examines the concept of "Risk" and "Exposure". This concept was used to formulate the accident prediction models in this research project.

The literature research also revealed that occurrences of traffic accidents were affected by a multitude of factors, of which driver factors played a dominant role. Delay was identified as the parameter that can be used to estimate the risk-taking behaviour of a driver. A discussion on the various factors involved in accidents and the role of drivers in accidents is included in the third section.

The last chapter concludes with a summary on three computer programs used in the research. The Traffic Data Input Program (TDIP), (Kyte & Boesen, 1989), was used to encode traffic data from video recordings. KNOSIMO (Grossmann, 1988), a traffic simulation model, was used to simulate additional delay data for the study locations. The Generalized Linearized Interactive Modelling (GLIM) program (Numerical Algorithms Group, 1987), was used to determine the optimum accident prediction model structures.

2.1 Previous Accident Prediction Approaches

An in-depth literature research was conducted on several previous accident prediction modelling studies. Several accident prediction models have been developed in past research projects to estimate traffic safety at roadway intersections. The most commonly used predicting parameter in these models was traffic volume.

This section compares the different approaches and techniques used in eight selected research projects. The first three models discussed are simple models. The remaining five models are more elaborate comprehensive accident prediction models developed for different types of roadways.

2.1.1 Sum-of-entering-volume Models

Raff (1953) related accidents at intersections to the sum of the flows entering the intersections. He found that the number of accidents per vehicle decreased as the sum-of-entering-volume increased. A functional relationship was not provided in his study. However, he developed a measure to provide some means of comparison of traffic safety between different types of accidents.

Grossman (1954) also defined exposure of accidents as the sum of flows at intersection crossing points. Intersection crossing points were defined as locations where two conflicting flows crossed each other.

2.1.2 Product-of-entering-volume Models

Tanner (1953), in a study of 390 accidents at 232 rural T-intersections in Great Britain, found that the number of accidents was related to the square root of the product of flows on the major and minor roads. The traffic flow parameter used was average daily (16 hours) traffic volumes.

McDonald (1953) studied 1,500 accidents at 150 intersections and defined a relationship between annual number of accidents and the daily number of vehicles on major and minor roads. The power of the volume parameter in the model was determined to be approximately 0.5, which was similar to the findings in Tanner's study.

2.1.3 Product-of-conflicting-volume Models

Hakkert & Mahalel (1978), studied accidents which occurred at 202 urban intersections and 40 "interurban" intersections in Israel. Sixteen hour traffic counts were collected on one weekday, generally between 6:00 a.m. and 10:00 p.m. Two years of accident records from 1971 to 1972 were used in the study. In Israel, only accidents with casualties were reported to the police. The study therefore reflected only those accidents

They concluded that vehicle exposure could be used as the basis for accident prediction. They concluded that the exposure at an intersection can be represented by the sum of the products of flow at all conflict points where vehicle paths crossed or merged

2.1.4 Opportunity-based Regression Models for Signalized Intersections

Plass & Berg (1987), based on 50 case studies in Florida, U.S.A, proposed the use of opportunity-based accident rate expressions to estimate safety at 3-legged and 4-legged signalized intersections. One-year accident records were used for the study. Based on the argument that different vehicles had different probabilities of getting involved in traffic accidents, the researchers claimed that it was not appropriate to use total-entering-vehicle-volume as an exposure measure.

Instead, they correlated accident types to vehicle movements. Expressions were derived for the following types of accident opportunities:

- single-vehicle accident opportunities
- rear-end accident opportunities
- head-on accident opportunities
- angle accident opportunities
- sideswipe accident opportunities

Each accident opportunity type represented a particular type of accident with its own specific combination of vehicle movements. This approach reflected the cause-and-effect relationship of the events in an accident. The total number of accident opportunities for an intersection was calculated by adding the individual opportunity types for all approaches at the intersection.

The opportunities were then used to calculate the opportunity-based accident rates. Three different levels of aggregation were applied, using hourly volumes, peak/off-peak volumes and average daily volumes. It was found that the level of aggregation of the model could significantly affect the resulting accident rates. The researchers concluded that "hourly traffic volumes may be necessary for reliable estimate of opportunity-based exposure levels."

2.1.5 Analytical Impedance Models for Rural 2-Lane Intersections

A study (Stanley Associates Engineering Ltd., 1983) on rural 2-lane 4-legged intersections in Alberta, Canada, attempted to evaluate analytically, the quality of service of 4-legged intersections. The objective of the study was to formulate a simple mathematical analytical model that utilized intersection turning conflict information to provide a measure of quality of service for rural intersections without traffic signals.

Assuming random arrivals of vehicles under a rural roadway setting, an impedance index was developed to account for the interactions of vehicle arrivals. The technique categorized conflicting traffic streams by the related non-priority movement types. Different probabilities of impedance functions were developed using the applicable gap acceptance criteria.

Using these probability functions, the volumes of traffic being impeded were calculated for each non-priority movement type. The sum of the impeded traffic volumes for all non-priority movements represented the total impedance index for the intersection. This calculated level of impedance reflected the quality of service of the intersection.

2.1.6 Disaggregate Product-of-flow Regression Models for Signalized Intersections

Hauer, Ng and Lovell (1989), used traffic volumes and 3 years of accident records collected from 145 locations in Toronto, Canada, to estimate safety at 4-legged signalized intersections. Based primarily on the prior-to-collision movements of the vehicles involved in accidents, 15 accident types were identified for each approach of an intersection.

Accident prediction models were statistically developed using GLIM, which is a generalized linear regression model (Numerical Algorithms Group, 1987). The structures of the models generally consisted of the products of traffic volumes of colliding movements.

Different models were developed to represent the a.m. peak, p.m. peak and off-peak conditions. Most of the 15 model types were found statistically insignificant for all three periods, except where accident frequencies were high. In the case where accident frequencies were low, average daily traffic volumes were used in place of hourly volumes in the model.

2.1.7 Product-of-flow Regression Models for Rural T-intersections

Pickering, Hall and Grimmer (1986), proposed models to predict traffic accidents at 300 rural T-intersections in 40 English counties. Their objective was to investigate the relationship among accidents, traffic flow, geometric layout and other features such as traffic speed and gradient.

Study locations were rural single carriageway roads with speed limits higher than 50 miles per hour. Four hour traffic turning counts were taken at each site to estimate the average daily traffic volume. Adjustments were made to account for road-type, annual traffic growth and monthly variation. Five years of accident records were used in their study.

Although geometric characteristics were included to help explain part of the variation in the accident data, traffic volumes were the primary explanatory variables in the accident prediction models. Accidents were classified by types of manoeuvre of the vehicles before the collision. In total, 25 accident types were identified. For each accident type, an equation was developed statistically using the GLIM regression modelling tool (Numerical Algorithms Group, 1987) to relate accident frequencies to the related traffic volumes. Geometric factors were subsequently added to the equations developed to account for between-sites variations.

2.1.8 Regression-based Models for Roundabouts

Maycock & Hall (1984), using data from 222 roundabouts in England, developed accident prediction models relating frequencies of various types of accidents at roundabouts to a range of explanatory variables. The objective of the study was to relate accident frequency to a range of explanatory variables for accident prediction at roundabouts.

Five years of accident records were used in the study. Sixteen hours turning counts were taken during weekdays from April to June and from September to October. Hourly counts and quarter-hour counts were taken in off-peak hours and peak hours, respectively. The volumes collected were adjusted for traffic growth, vehicle type, day-of-week, and monthly and yearly variation. Site attributes collected at each site included junction type, speed limit and the geometry of the roundabout.

The models were developed based on pre-crash vehicle movements, using GLIM (Numerical Algorithms Group, 1987) as the modelling tool. The following accident types were categorized:

- entering circulating
- approaching
- single vehicle
- others
- pedestrian

The basic models were initially formulated to relate accident frequencies to the corresponding traffic volumes. Geometric variables were subsequently added in the model structures. Models were developed for both the peak and off-peak hours.

2.1.9 Overview of Different Approaches to Accident Prediction Modelling

Several main observations were noted in the review of previous accident prediction studies.

First, most of the models attempted to model cause-and-consequence relationships in the accident events by relating pre-crash vehicle manoeuvres to the corresponding type of accidents. This approach was logical and made intuitive sense. In most of these studies, the approaches used in formulating the accident prediction models were based on the risk and exposure concept. The risk and exposure concept is discussed in more detail in Section 2.2 of this chapter.

Second, the level of aggregation appeared to be critical in the formulation of an accident prediction model. In some of the studies, specific models were developed for peak hour periods. Even quarter-hour traffic volumes were used by Maycock & Hall (1984) to formulate accident prediction models for roundabouts. It appears that different accident prediction models may be needed to model the safety conditions for different times of day or different days of the week.

Third, most models used the frequency of a particular type of accident as the measure of non-safety.

Fourth, the GLIM regression modelling tool (Numerical Algorithms Group, 1987) was used in a number of past studies. Further review on GLIM was carried out to examine its applicability in the current research.

Fifth, all the previous studies reviewed used traffic volumes as the primary model parameter in predicting accidents. It was decided that a general review on the nature of accident occurrences be carried out and the possibility of using an alternative model parameter be considered.

Sixth, the studies reviewed all have extensive data collection programs. Most of the studies used three to five years of accident records accompanied by sizeable traffic counting programs.

2.2 Risk and Exposure Concept

Most of the accident prediction models examined in the literature research were, in one way or another, based on the concept of risk and exposure. Hauer (1982) made a clear explanation of the concept by relating it to the philosophy of chance:

> A unit of exposure corresponds to a trial. The result of such a trial is the occurrence or non-occurrence of an accident (by type, severity, etc). The chance set up is the transportation system (physical facilities, users, and environment) which is being examined and risk is the probability (chance) of accident occurrences in a trial.

The following relationship represents the fundamental structure of a typical risk and exposure accident prediction model:

NON-SAFETY	=	RISK	x	EXPOSURE, or
S	=	R	x	E,

where

S = Measure of intersection non-safety during a period of time

R = Measure of the risk presented to drivers during the same period

E = Measure of exposure, i.e., the number of drivers that are presented with the accident risk during the same period

Under the risk and exposure accident prediction modelling approach, risk and exposure are the independent variables used to predict intersection non-safety, the dependent variable. The non-safety parameter is typically represented by the number of accidents. The risk and exposure parameters are usually represented by some surrogates, which are measured at the study location and used as inputs to the model in estimating the actual number of accidents (i.e., non-safety) at the intersection.

Take as an example the throwing of a die. The chance of throwing a six in a number of throws is equal to the chance of getting the number six in each throw multiplied by the number of throws. Risk can be viewed as the chance of getting a six in one throw. Exposure can be viewed as the number of throws performed. Non-safety, which is the outcome of a combination of risk and exposure, can be viewed as the chance of getting at least a six once after a given number of throws. In other words, based on the risk and exposure concept, the non-safety at an intersection is the product of the risk presented to the vehicles entering the intersection and the exposure of the traffic with the risk.

Details of the model structure as well as the forms and interactions of the various parameters can be determined through the use of statistical modelling tools.

2.3 Driver Factors and Accidents

As suggested by Teply (1988), the complexity of the traffic system can be illustrated by the system concept of Klebelsberg (1982). Under the system concept, the roadway traffic is considered an overall system which consists of sub-systems such as drivers, vehicles, roadways, traffic controls and rules. These sub-systems exhibit various system functions including driver behaviour and roadway conditions. The sub-systems and their functions interact to form system states at different points in time, resulting in various traffic operating conditions or traffic interactions, which may result in traffic conflicts and accidents.

2.3.1 Driver, Vehicle and Environment Factors

The traffic systems is a complex system in which several factors exist and interact with each other. This multitude of factors can be categorized into three major categories: driver, vehicle and environment. Figure 2.1 illustrates the complexity of the interaction among the three factors. It illustrates the role of the driver in the driving process.

Examples of driver factors are age, gender, driving experience, stress level and fatigue. Vehicle factors include factors such as spray from tires, headlights, brakes, viewfield, mirror and vehicle stability. Typical environment factors are geometry of roads, pavement marking, road surface conditions, weather conditions, surface texture, roadway lighting and traffic conditions.

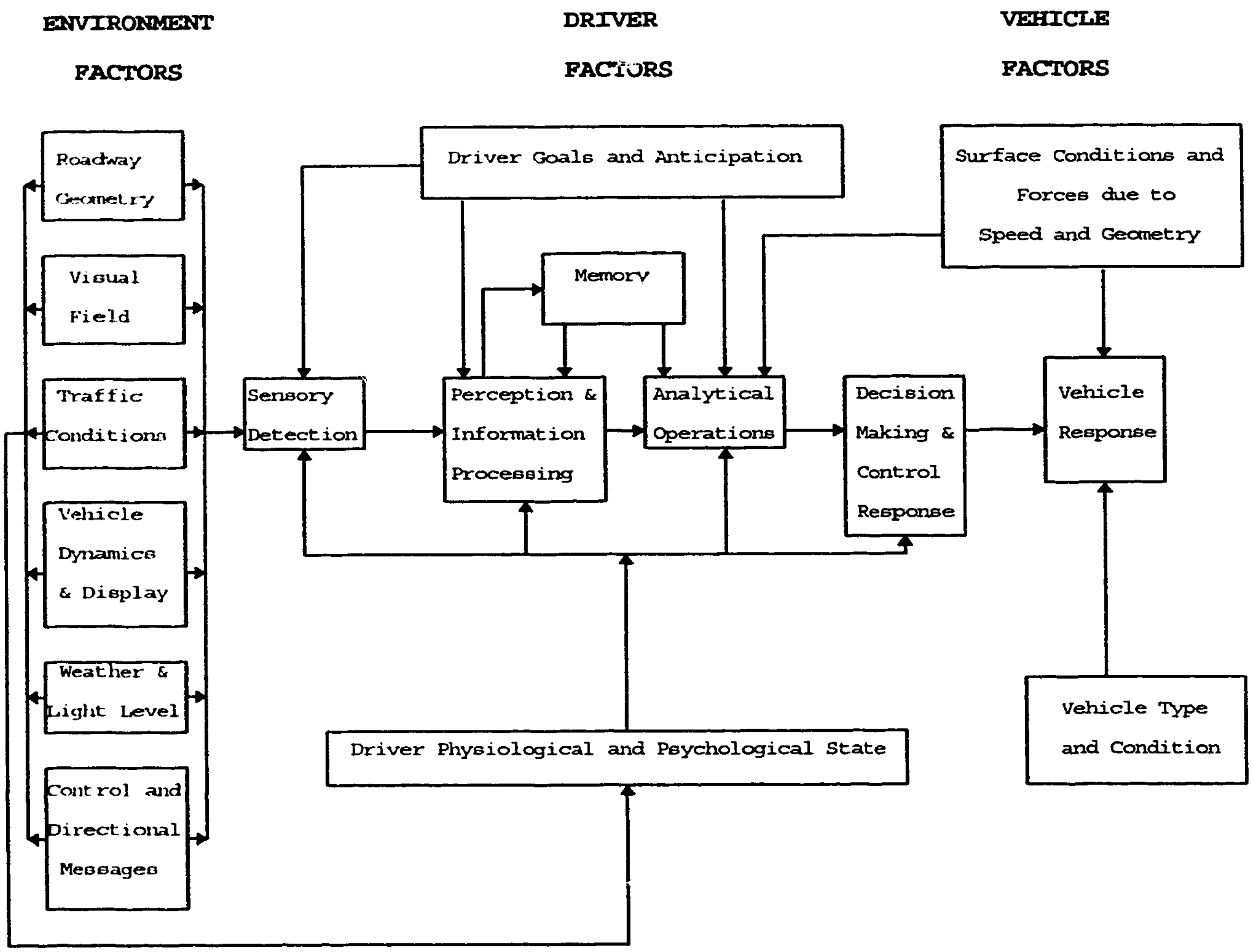


Figure 2.1 Interactions among Driver, Vehicle and Environment Factors

2.3.2 Role of Drivers in Accident Occurrences

Several factors exist and interact to each other in a traffic system. Vehicle drivers, in perceiving and responding to these factors while driving, play a particularly important role within the system.

In a number of in-depth accident investigation studies (Sabey & Staughton, 1975 and Sabey, 1983), driver factors were found present in a majority of the accidents. Figure 2.2 (Sabey, 1983) illustrates that human factors were involved in approximately 95% of the accidents, and were identified as the sole cause in 65% of them.

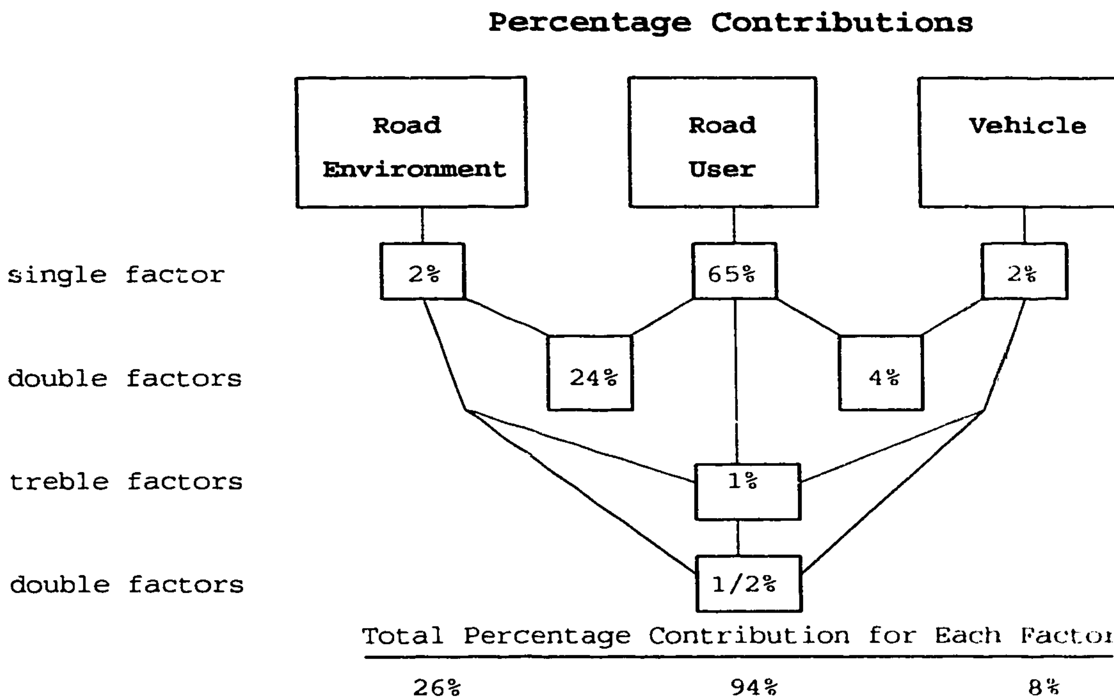


Figure 2.2 Roles of Driver, Vehicle and Environment Factors in Accident Occurrence

Comparatively, environment and vehicle factors contributed respectively to 28% and 9% of the accidents. Environment and vehicle factors were found to be present in approximately 30% of all human factor related accidents .

The in-depth accident investigation demonstrated that the human element was dominant in traffic accidents.

2.3.3 Risk-taking Behaviour Of Drivers

The utility maximization theory (Blomquist, 1986) assumes that individuals pursue multiple goals with limited resources. Each rational individual has goals that are accompanied by some utilities, such as safety, economy, travel time, comfort, etc. An individual, in attempting either consciously or subconsciously to maximize the total benefit or utility to him, compares these goals against each other.

Under this theory, the behaviour of a driver can be explained by the concept of risk compensation (Blomquist, 1986). For instance, the driver may compensate some of his safety utilities for a gain in personal convenience or other personal goals that will bring benefits or pleasure to himself.

Klebelsberg (1982) suggested that driver behaviour in a risk-taking situation is the end-product of a process in which the driver evaluated and acted in response to the subjective and objective risk. Subjective risk is the risk perceived by the driver, whereas objective risk is the "actual risk" that is presented to the driver under a particular combination of driver, vehicle and environment conditions.

In a situation where the subjective "perceived" risk to a particular driver is lower than the objective "actual" risk, the driver expects the situation to be safer than it actually is and may drive in an undesirably aggressive manner. As a result, more accidents than anticipated are likely to occur.

2.3.4 Gap Acceptance

The risk-taking behaviour of drivers can be reflected by the gap-acceptance characteristics of drivers at intersections without traffic signals. Gap acceptance is the choice of a driver, merging from a minor street to a major street or crossing another traffic stream, to accept or reject the gaps available within the major street traffic stream. Critical gap is defined as the gap, in seconds, which 50% of all drivers will accept.

The actual mechanism behind the drivers' gap acceptance behaviour is complicated and can be affected by several driver, vehicle and environmental factors. Abou-Henaidy (1993) concluded that driver's gap acceptance behaviour, when making a major road left-turn manoeuvre, was influenced by the following factors:

- gender
- presence of passengers in the turning vehicle
- queue delay
- front delay
- gap size
- type of opposing vehicle
- opposing traffic flow
- width of cross street
- location of intersection
- number of rejected gaps

Abou-Henaidy (1993) concluded that the probability of a driver to accept a gap decreased when the number of gaps previous rejected by the driver were fewer than 14; whereas, the probability to accept a gap increased when the number of gaps previously rejected exceeded 14. Abou-Henaidy also concluded that a driver was less likely to accept a gap when the front delay was less than 30 seconds, but more likely when front delay exceeded 30 seconds. It was also found that critical gap decreased with increasing traffic volumes in the priority traffic stream, as well as with increasing traffic on the non-priority traffic stream (Brilon, 1988).

The Alberta Traffic Signal Control Display Standardization and Guidelines (Alberta Transportation, 1985) suggested that motorists in more populated communities are likely more inclined to accept shorter gaps.

Gap acceptance behaviour is therefore not necessarily constant. Instead, it often reflects the drivers' risk-taking behaviour as a result of the prevailing traffic conditions and "traffic pressure".

2.3.5 Delay

As delay is a traffic parameter that can be directly experienced by the drivers, it can make a better representation of the actual traffic conditions than traffic volumes. This characteristic makes it possible for delay to be used in transportation planning models to estimate the route choices of drivers within a roadway network.

At the operational level, delay is also an important parameter. Delay is used in both the Canadian Capacity Guide for Signalized Intersection (Institute of Transportation Engineers, 1984) and the Highway Capacity Manual (Transportation Research Board, 1985) as one of the intersection operation evaluation criteria. The amount of delay experienced by drivers at an intersection determines the quality of service provided by the facilities at that location.

The effect of delay on drivers' gap acceptance behaviour reflects that drivers probably evaluate the delay they experienced and compare that to the risk they have to take in accepting gaps in the major traffic stream. This is reflected by the findings in Abou-Henaidy's research (1993), which concluded that the gap acceptance behaviour of drivers changed depending on the "traffic pressure" and the "patience level" of the drivers.

The findings, as summarized in Section 2.3.4, suggested that drivers' gap acceptance behaviour was more cautious when delay was moderate. Drivers were less likely to accept a gap when the front delay was less than 30 seconds or when the number of gaps rejected were fewer than 14.

However, as delay and inconvenience to the driver increases, the gap acceptance behaviour of the driver became more aggressive. Drivers were found more likely to accept a gap when the front delay exceeded 30 seconds, or when the number of gaps rejected exceeded 14.

Evidently delay affects the risk-taking behaviour of drivers. It is therefore postulated that delay may be a better accident prediction parameter than traffic volumes.

2.4 Computer Programs

Three computer programs were used to undertake various analysis tasks in this research project. This section provides a brief discussion on the three programs.

2.4.1 Statistical Analysis Program

The Generalized Linear Interactive Modelling (GLIM) package (Numerical Algorithm Group, 1987), as used by Maycock & Hall (1984), Pickering et al (1986) and Hauer et al (1989) in their studies, was used in the current study. GLIM was required for this study because of the unique distribution of the accident data.

2.4.1.1 Nature of Accident Counts

The object of statistical modelling is to present a simplified representation of the data population. In a typical statistical model, the variation in the data is represented by the systematic components, whereas the unexplained part of the data is treated as the random component. The systematic components can be described by model parameters derived by regression, while the random component can be represented by a probability distribution.

Literature review indicated that the within-site distribution of accidents at a particular location follows the Poisson process (Abbes, Jarret & Wright, 1987 and Maycock & Hall, 1984):

$$P(y \mid x) = \frac{x^y e^{-x}}{y!}, \qquad y = 0, 1, 2 \ldots$$

where

x = mean accident frequency at the site

y = number of accidents in a specific period of time

The mean accident frequency is itself a random variable that varies from site to site. The between-site variation may be described by a gamma distribution (Abbes et al, 1987 and Maycock & Hall, 1984):

$$P(x) = \frac{s^s x^{s-1} e^{-(s/\mu)x}}{\mu^s \Gamma(s)}$$

where

μ = the mean of the distribution

s = a parameter of the gamma distribution

The resulting sampling distribution over all sites will be (Abbes et al, 1987 and Maycock & Hall, 1984):

$$P(y) = P(x)\, P(y \mid x)\, dx$$

which gives

$$P(y) = \frac{\Gamma(s+y)}{\Gamma(s)y!} \left(\frac{s}{\mu+s}\right)^s \left(\frac{\mu}{\mu+s}\right)^y$$

which is the negative binomial distribution.

2.4.1.2 Classical Linear Models

In the classical least-square regression approach, the general form of a multiple linear regression model is as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \varepsilon$$

where β_i is assumed independent and normally distributed with mean zero and variance δ^2, i.e.

$$\varepsilon \sim N(0, \delta^2)$$

However, the negative binomial distribution of accident sampling data does not comply with the normality assumption of the classical linear regression approach.

Further, Aitkin, Anderson, Francis and Hinde (1989), stated that for probability distributions other than normal and extreme value, the representation of the model with an additive error term ε is not satisfactory, because ε does not have any simple distribution.

Based on the above reasons, the classical linear regression approach is not considered appropriate for accident data analysis.

2.4.1.3 Generalized Linear Models (GLM)

Generalized linear model (GLM), as proposed by Nelder & Wedderburn (1972), is a generalized version of multiple linear regression. Instead of requiring a restrictive normality assumption, GLM can describe various error distributions in the exponential family. The exponential families have a general format which includes distributions such as Normal, gamma, Poisson, Chi-square and binomial.

The GLMs have the following structure (Numerical Algorithms Group, 1987):

$$y_i = \mu_i + \varepsilon_i, \quad i = 1, 2, \ldots, n$$

where

y_i = set of independent variables

= observed value or data

μ_i = systematic component

= fitted values or theoretical values

ε_i = random component

A GLM can be defined by three components (Aitkin et al, 1989):

(1) Probability distribution of y_i , which represents the random component.

(2) Linear predictor n_i , which represents the systematic components of the model by describing the linear regression function:

$$n_i = \sum_j x_{ij} \beta_j$$

where

x_{ij} = model parameters

$ß_j$ = model coefficient estimates

(3) Link function $g(\mu)$ that relates the linear predictor n to the mean μ:

$$n_i = g(\mu_i)$$

The classical linear model is a special case of the GLM with Normal probability distribution, additive linear predictor and identity link function.

In application, model fitting depends on the form of the probability distribution. The method of maximum likelihood is used in modelling data with various data distribution. "The likelihood of a model is the probability with which it will occur calculated for a particular set of parameter values." (Aitkin et al, 1989). The likelihood function is:

$$L(\mu_1,\ldots,\mu_n,\phi) = \pi\ f(y_i \mid \mu_i,\phi)$$

where $f(y_i \mid \mu_i, \phi)$ is the individual probability of obtaining the observation y_i

The maximum likelihood estimate $\tilde{\Theta}$ of θ is defined as the value of θ for which $L(\tilde{\Theta}) \geq L(\theta)$ for all θ's. It is obtained by equating the partial differentiates of the likelihood function to zero. It can also be obtained by the method of iterative least squares (McCullagh & Nelder, 1987).

2.4.1.4 Generalized Linear Interactive Modelling (GLIM)

In this research, Release 3.77 of the GLIM program developed by the Royal Statistical Society was used in modelling GLMs. GLIM is specially designed to facilitate the fitting of GLMs. GLIM is one of the few programs that can work with different kinds of probability distributions. Users can specify their own "user-defined models." As a result, models of considerable generality can be described. Examples are negative binomial, censored exponential, Weibull, extreme value and logistic distributions (Gilchrist, Francis & Whittaker, 1985).

In the GLIM modelling process, the data y_i's are matched by a set of theoretical values μ_i's (Numerical Algorithms Group, 1987). This is accomplished by defining a specific GLM structure to specify the appropriate probability distribution, link predictor and link function that are adequate for the set of data to be modelled.

Within GLIM, the method of maximum likelihood is used to select parameter estimates that minimize the "deviance". Deviance is a measure of the goodness-of-fit in GLIM. Its form depends on the distribution assumed.

2.4.1.5 Significance Testing, Goodness-of-fit and Parameter Estimates

To test for significance of GLM, it is necessary to take a closer look at the probability density function. As previously mentioned, the exponential family distributions have a general format of:

$$p(y_i) = \exp\{[y_i\theta_i - b(\theta_i)] / a_i(\o) + c(y_i,\o)\}$$

where

$$\o = \text{scale parameter}$$

The mean and variance of the probability density function can be expressed in terms of θ_i and by the first and second derivative of the function:

$$E(y_i) = b'(\theta_i)$$

$$VAR(y_i) = b''(\theta_i) \cdot a_i(\o)$$

$a_i(\o)$ is usually of the form $\o / w_i$, where w_i is called the prior weight. Therefore, the variance function becomes:

$$VAR(y_i) = b''(\theta_i) \cdot \o / w_i$$

During the process of model fitting, it is necessary to know the significance of extra parameters for a model. The acceptability of the current model as compared to the full model can be determined by comparing the likelihood of the current model (l_c) to the likelihood of the full model (l_f) with the given data. This measure is called the "scaled deviance", S(c,f):

$$S(c,f) = -2 \log (l_c/l_f)$$

$$= -2 [\log(l_c) - \log(l_f)]$$

Substituting for l_c and l_f the probability density function previously established with the maximum likelihood estimates (MLE) of the parameters, the following is obtained (Numerical Algorithms Group, 1987):

$$S(c,f) = 2 \sum\{ y_i(\theta_i - \tilde{\theta}_i) + b(\theta_i) - b(\tilde{\theta}_i)\} / a_i(\varnothing)$$

where θ_i and $\tilde{\theta}_i$ are the MLE of i under the current and full models respectively. The relation can be written as:

$$S(c,f) = D(c,f) / \varnothing$$

where

$D(c,f)$ = deviance of the current model relative to the full model

$\varnothing$ = scale parameter

For negative binomial distribution, the scale deviance is:

$$\text{Scale Deviance} = 2 \sum y \log \frac{y}{\mu} - (y+s) \log \frac{(y+s)}{\mu+s}$$

In the negative binomial scaled deviance, all the parameters are known except s. To calculate the scale deviance, it is necessary to assume different values for s before fitting and find the s that gives the best scale deviance value. The scale deviance is distributed as λ^2 with $t_1 - t_2$ degree of freedom, where t_i is the number of parameters estimated under model i (Numerical Algorithms Group, 1987).

For parameter estimates within a model, the t distribution test can be used as an approximation for the comparison between the parameter estimates and their standard errors, as the t test is only an approximation for distributions other than Normal. An estimate that is more than three times its standard error is usually significant (Numerical Algorithms Group, 1987).

This section serves as a brief introduction to GLM and the GLIM program. More details about the GLIM statistical modelling approach can be found in the GLIM Manual - Release 3.77 (Numerical Algorithms Group, 1987).

2.4.2 Video Traffic Survey Data Encoding Program

The Traffic Data Input Program (TDIP) (Kyte & Boesen, 1989), was used in this research project to encode the data collected in the video survey. Through TDIP, traffic volumes and delay can be encoded directly from a video recording. The operator can place a time stamp on each data entry by pressing a designated number key on the personal computer keyboard to encode the data.

In encoding traffic volumes, the number keys on the personal computer keyboard can be assigned for different traffic streams. The program allows the operator to record up to 4 traffic streams simultaneously. However, data recording for a maximum of 2 streams at one time is recommended.

In encoding delay, two number keys are required to record the arrival and departure time of the vehicles. The time difference between the arrival and departure time stamped is the delay experienced by the drivers. Some judgement is required to estimate the actual time incurred to the drivers as it is necessary to take into consideration the time needed for acceleration and deceleration when the vehicle is leaving and joining the queue. It is recommended that the same operator be used in encoding delay data to maintain consistency in data encoding.

2.4.3 Traffic Delay Simulation Program

The field collected delay data exhibited a high degree of variability. Because of the time and budgetary constraints of this pilot research project, it was not practical to have an extensive data collection program. To supplement the field data collected, a traffic simulation model, KNOSIMO (Grossmann, 1988), was used to estimate the delay at the study locations.

KNOSIMO is a simulation computer program for intersections without traffic signals. It was developed at the Ruhr University in West Germany for the West Germany Federal Minister of Transport. The program is applicable for 3 or 4-legged single lane roadways.

KNOSIMO is an event-oriented simulation program which requires a relatively short time per simulation. The program is interactive and can be run on various types of personal computers under the PC-DOS environment. The model consists of individual vehicles taking certain actions at certain events. Vehicles are generated in the program using the hyper-erlang probability distributions, which can represent partial constraint of single lane traffic streams realistically.

The operations of non-priority stream manoeuvres are determined by using the principle of gap acceptance. Different critical gaps, t_c, and move-up times, t_f, are used for different types of roadway and traffic conditions. Once a particular set of t_c and t_f are determined, they will be applied to each gap accepting operation according to an erlang probability distribution to reflect the actual distribution of the two values among drivers. During the simulation process, total delay and queue are constantly stored for each simulated interval. Reports are available at the end of the simulation period.

A copy of the KNOSIMO program was obtained from Ruhr University. An English translation of the User Manual (in German) was also prepared as part of this research.

Use of simulation models at low traffic volume locations can greatly reduce data collection time. As most intersections without traffic signals have relatively low traffic volumes, a simulation model can simulate a sufficiently long data collection period while keeping the various travel conditions stable to minimize variability in delay data.

Because of the ease of its application, the KNOSIMO program was chosen to estimate traffic delays. However, the model at the present stage was only developed for

2-lane roadways and was not readily applicable to multi-lane situations. To simulate traffic conditions for multi-lane intersections, it was necessary to address the differences in the conditions considered in the program and at the studied locations. A paper written by Chan & Teply (1992), described the modifications needed to apply the KNOSIMO program to an urban 4-lane situation, and is included in Appendix A.

3. DATA COLLECTION

This chapter summarizes the selection of model parameters for each of the non-safety, risk and exposure variables. The accident and traffic data provided by the City of Edmonton was used to determine the framework of the research project. This included the study scope, analysis period and study locations. A detailed description of the data collection method is provided in this chapter.

3.1 Selection of Measures of Non-safety

3.1.1 Problems with Historical Accident Records

Most accident prediction models used historical accident records as the measure of non-safety at the study locations. A number of problems in using accident records were investigated. The biggest problem of using accident records is the low frequency of accident occurrences, which often translates into significant fluctuations in accident trends. Accident records for a number of locations over a relatively short study period will be a set of data of small numbers. Data with small numbers is difficult to analyze statistically. To compensate for this, longer study periods are often used to create a more workable set of data.

However, the use of a longer study period also leads to other difficulties. Traffic and roadway conditions often do not remain stable over an extended period of time, resulting in additional variability in the accident data and making analysis more difficult.

The reporting criteria of accidents can also affect the validity of accident records. An accident that does not involve injuries or fatalities is only reported when the minimum cost of the property damage exceeds the regulated reporting criteria. As vehicle repair costs increase with inflation, more traffic accidents will have damages above the reporting criteria and become reportable. As Hauer et al (1989) stated, the "inflation that eats away at the value of the dollar causes an inflation of reportable accidents".

There are limited alternatives to accident records. The common alternatives are traffic conflicts and critical incidents. These alternatives also have some disadvantages and are discussed in the following sections.

3.1.2 Traffic Conflict

There are many definitions for traffic conflict. The definition by Amundsen and Hyden (1977) seemed to be most appropriate:

> a traffic conflict is an observable situation in which two or more road users approach each other in space and time to such an extent that there is a risk of collision if their movements remain unchanged

However, many traffic conflict studies were based on different definitions of traffic conflicts. For this reason, it was difficult to compare the results of one study to another.

The use of traffic conflict as a model parameter requires a large amount of data. The method often relies on subjective interpretations by observers during the data collection and encoding stages. Therefore, a high level of judgement by trained observers is required for data collection and data encoding.

The biggest shortcoming is that the validity of relating traffic conflicts to accidents is still unproven. The traffic conflict technique attempts to estimate the number of accidents based on the assumption that a particular type of conflict will lead to a certain number of related accidents. Williams (1981) concluded that the findings from various traffic conflict studies on the ratio were contradicting.

3.1.3 Critical Incident

Critical incident is another alternative to accident records. Teply (1987), in a "before" and "after" study on pavement markings, employed a critical incident technique to assess driver behaviour. Teply defined critical incidents as "driver actions which are inconsistent with roadway design or traffic situations".

Data collection for the critical incident approach is also labour-intensive, as compared to typical methods of traffic volume collection. Subjective judgement and interpretation by the observer are needed to identify each critical incident type.

3.1.4 Selection of Non-safety Measures

Accident record, despite its shortcomings, is preferable over traffic conflict and critical incident techniques for the present study. The biggest advantage of accident record is that it is an original source of data. Contrary to the traffic conflict and critical incident techniques, the procedures for accident data collection and documentation are often well-established, thereby minimizing additional data collection requirement. As well, accident record, being an original source of traffic non-safety data, does not require any correlation parameters such as the accident-to-conflict ratio.

Accident frequency was therefore chosen for use as the non-safety parameter in the accident prediction model in this study.

3.2 Selection of Measures of Risk

3.2.1 Traffic Volume and Delay

At intersections without traffic signals, the gap acceptance characteristics of drivers are closely related to the risk-taking behaviour of drivers. Two parameters, traffic volumes and delay, were used as the surrogates representing the risk presented to drivers in the accident prediction models.

Traditionally, traffic volume is the most often used parameter in accident prediction. Main road traffic volume is frequently used to represent the risk to the minor road traffic. An advantage of using traffic volume as a risk parameter is that it can be easily collected.

Delay, on the other hand, may be a better risk surrogate measure because delay can be experienced more directly by the minor street drivers. However, it is realized that delay data is more difficult to collect than volume data. However, the use of a computer simulation program such as KNOSIMO can significantly reduce data collection efforts.

In this study, accident prediction models were developed using the volume and delay data, respectively. Comparison of the models could determine whether delay or volume data would result in a accident prediction.

3.2.2 Stopped Delay

The delay measure used in this study is "stopped delay". Stopped delay does not consider the time lost in accelerating and decelerating the vehicle as part of the delay. At a stop-controlled intersection, the delay experienced by a minor road vehicle driver is the delay that he faces after the vehicle has come to a mandatory stop in front of the stopline. Stopped delay which discounted the expected time lost in acceleration and deceleration can reflect a more realistic delay situation to the drivers.

3.3 Selection of Measures of Exposure

Traffic volumes were used to reflect the exposure of the traffic that were faced with the accident risk. Depending on the model structure, different combinations of traffic volumes were used.

3.4 Relevant Accident Trends in Edmonton

The City of Edmonton's accident records were analyzed in establishing the research framework, the scope of study as well as the analysis period. The City of Edmonton Transportation Department maintains an accident inventory computer file that contained, at the time of the study, all reportable on-street motor vehicle accidents in Edmonton from 1985 to 1988. A reportable accident is defined as "an accident involving property damage in excess of $500 and/or resulting in an injury, and not occurring on private property or off-highway" (City of Edmonton Accident Decoding Manual, 1989).

A total of 3,795 reportable accident records from 1985 to 1988 for 429 intersections without traffic signals were extracted from the computerized database. The 429 intersections without traffic signals were selected from a Edmonton area map on the basis that the intersections were not located at a traffic circle or a curve. Of the 429 intersections selected, 129 were 3-legged and 300 were 4-legged. Summaries of accident data for both types of intersections are in Table 3.1.

Table 3.1 Accidents at 429 Selected Intersections Without Traffic Signals (1985-1988)

Number of Accidents (1985-1988)	3-legged	4-legged	Total
Number of intersections	129	300	429
Total Number of accidents	733	3062	3795
Average number of accidents in 4 years	5.7	10.2	8.8

In addition to the accident records, the following data were obtained:

1. Annual Meteorological Summary for the Edmonton Municipal Airport, 1985-1988. (Source: Environment Canada)

2. Arterial Roadway Construction Schedule, 1985-1988. (Source: City of Edmonton Transportation Department, Design and Construction Branch)

3. Key and Regular Station Traffic Count Reports, 1985-1988, various count station locations. (Source: City of Edmonton Transportation Department, Transportation Planning Branch, Monitoring Unit)

3.5 Effects of Travel Conditions and Weather

3.5.1 Travel Conditions

The general accident trends for the 429 intersections are illustrated in Figures 3.1a-f. Trends for accidents at 3-legged intersections are shown by a line profile on the same graphs.

Figure 3.1a indicates that accident trends remained relatively steady between 1985 and 1988. Very different trends were observed between daytime/night-time accidents and weekday/weekend accidents. Figure 3.1c illustrates that there were more accidents on Fridays as compared to the other days of the week, especially Saturdays and Sundays. Figure 3.1d illustrates that accidents peaked during the peak hours. The proportion of accidents in the daytime was significantly greater than that in night-time.

To examine in more detail the underlying trends between these time periods, the accident records were categorized into four arbitrary modules to represent weekday/weekend and daytime/night-time conditions, as summarized in Table 3.2.

Table 3.2 Accidents on Weekday/Weekend and in Day/Night

Number of Accidents (1985-1988)	Weekday MON-FRI	Weekend SAT and SUN	Weekly
Daytime (6 am to 9 pm)	2814 (74%)	532 (14%)	3346 (88%)
Night-time (9 pm to 6 am)	272 (7%)	177 (5%)	449 (12%)
Daily	3086 (81%)	709 (19%)	3795 (100%)

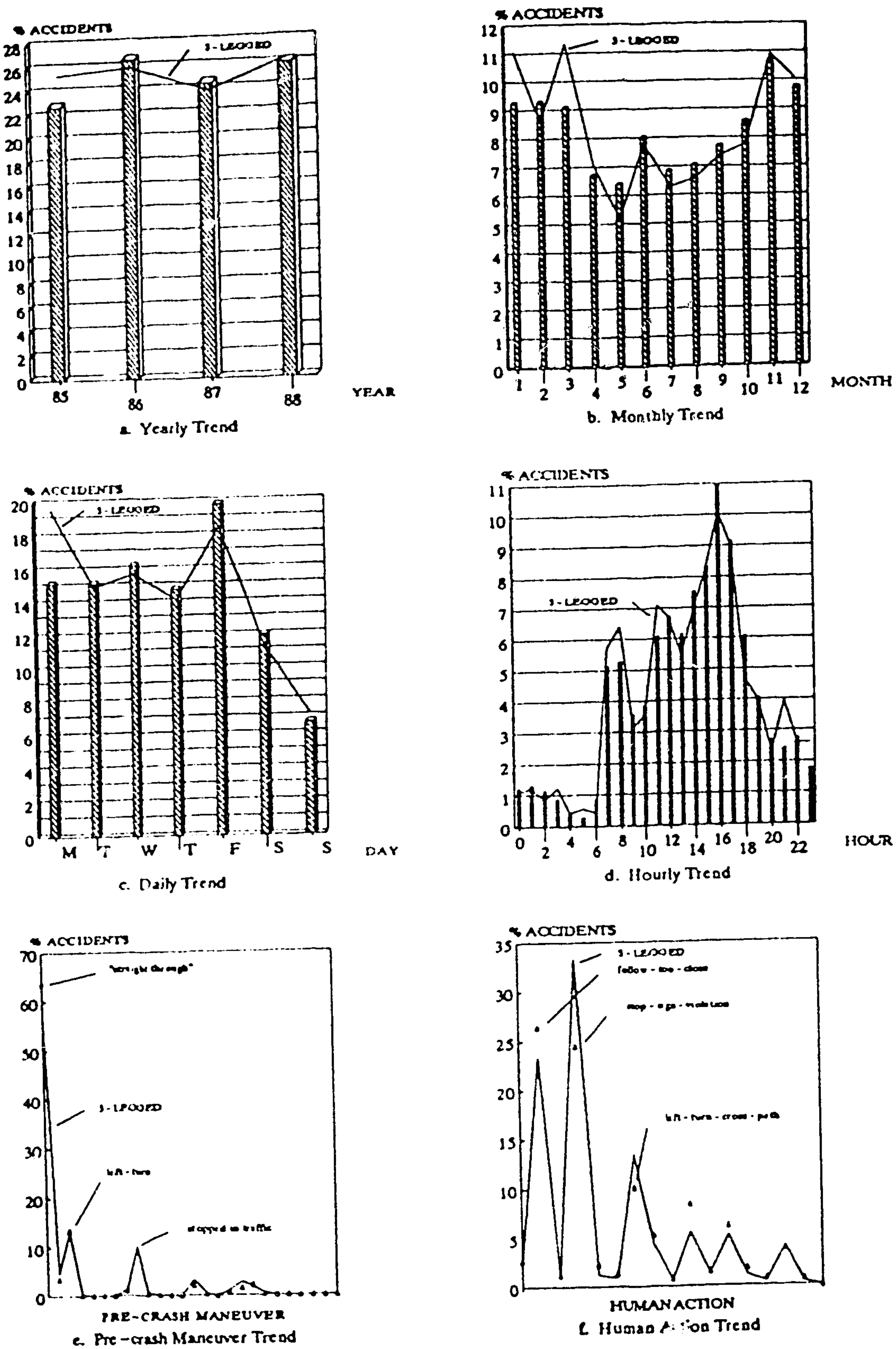


Figure 3.1 Accident Trends At 429 Intersections Without Traffic Signals In Edmonton

Differences in trends between weekday and weekend accidents, as well as between daytime and night-time accidents are illustrated in Figures 3.2a-c and Figures 3.3a-c.

Figures 3.2a-c illustrates that approximately 12% of all accidents occurred in night-time. However, when examining weekday accidents separately from weekend accidents, night-time accidents were found to occur much more frequently during weekends with 26% being night-time accidents and only 9% of the weekday accidents being night-time accidents.

Observations made from a slightly different perspective revealed another aspect of the accident trend. Figure 3.3a illustrates that, on the average, 19% of all accidents occurred on the weekends. Further examination using Figures 3.3b-c indicated that only 16% of all daytime accidents actually occurred on weekends, whereas approximately half (42%) of the night-time accidents occurred on weekends.

Based on the observed accident trends, it is apparent that there are different accident patterns between weekday and weekend and between daytime and night-time. The difference in accident trends among the four patterns may be explained by the fact that the underlying conditions within each time period may be significantly different from the others. These time periods can be considered as distinctive "travel condition modules" in that each has its own individual characteristics in traffic, driver and environment factors and/or conditions. These factors may interact with each other within each module and create different traffic operation and safety conditions. For example, during the morning peak hours, travellers are mostly comprised of commuters who have similar home-to-work or home-to-school trip purposes and similar trip destination locations. Traffic and ambient conditions are also similar.

Based on this argument, different accident prediction models should be developed for each travel condition module because of the difference in accident experiences and underlying travelling conditions.

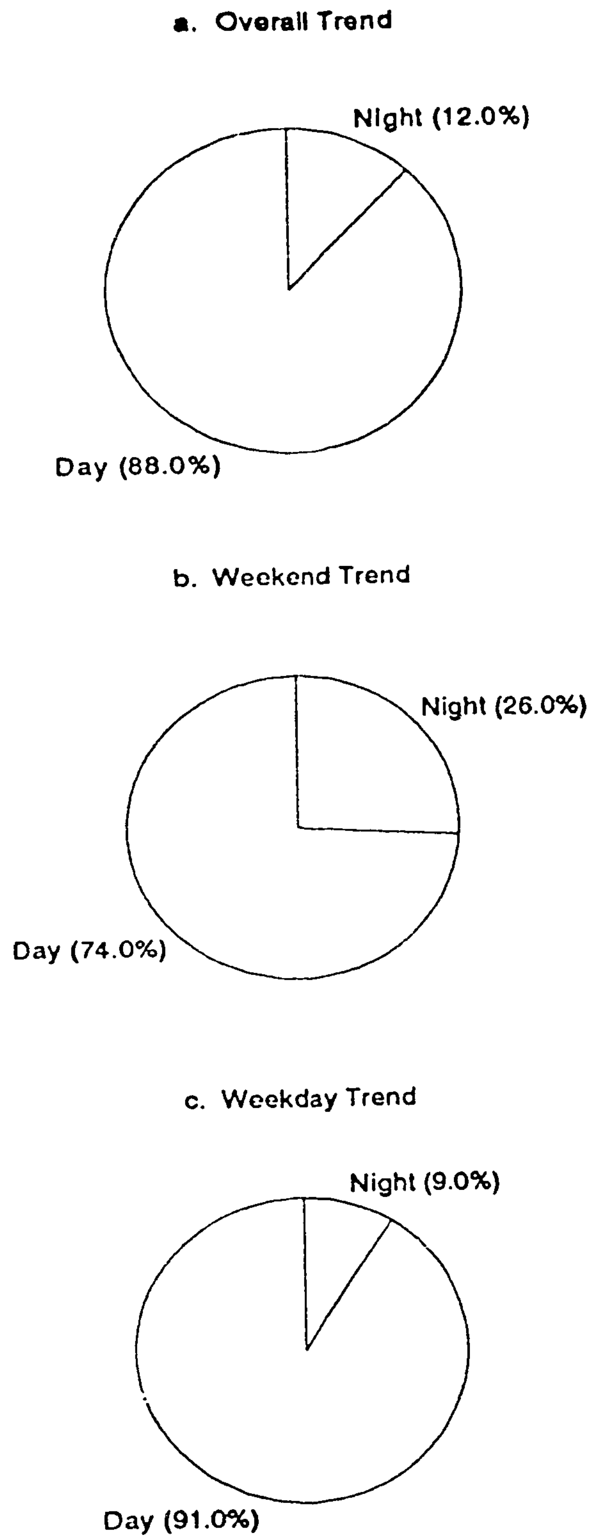


Figure 3.2 Weekday and Weekend Accident Trends

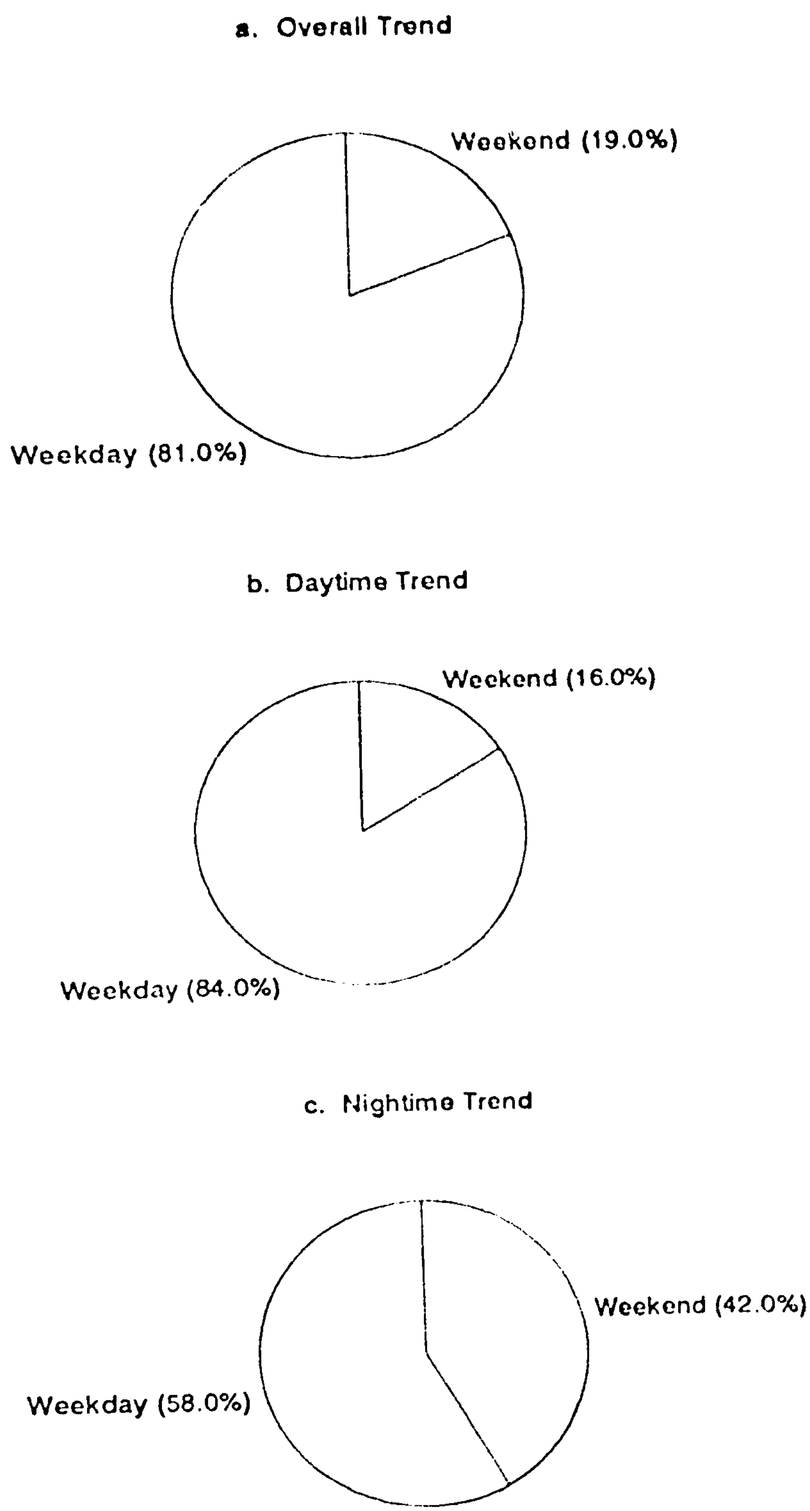


Figure 3.3 Daytime and Nightime Accident Trends

3.5.2 Weather Conditions

The winter and summer climate in Edmonton are very different. In the winter, the treacherous roads and the cold temperatures have a significant effect on vehicle performance. Accident occurrences peak in months with a considerable amount of snowfall. With in excess of 4 months of winter each year, the drivers in Edmonton are relatively accustomed to the winter driving conditions. Major roadways within the city are routinely snow-ploughed and sanded. It was decided that the accident statistics for the entire year, including those that occurred in the winter, would be included in the accident database for modelling.

Cumulated frequency curves as shown in Figures 3.4a to 3.4c were used to determine graphically the 95th percentile daily accident levels for the 429 intersections without traffic signals. The figures indicate that 95% of the intersections had 9 or fewer accidents per day during the winter months from 1985 to 1988. The 95th percentile figure was approximately 6 accidents per day during the summer months.

The 3,795 accident records in the accident database were evaluated using the two criterion developed above. The two 95th percentile criterion were exceeded 17 times between 1985 and 1988. Of the 17 days with unusually high accident frequencies, the 9 days with the highest daily accident frequencies occurred between November and February. Almost all 9 days with the worst accident records had unusually heavy snowfalls and very slippery road conditions. As a result they had a considerably higher number of accidents than an average winter day. These accidents were considered to occur under "extreme winter conditions". Under "extreme winter conditions", road environments, vehicle performances and fumes from the vehicle exhaust result in very poor driving situations. The weather on the 9 days with the worst accident records between 1985 and 1988 are summarized in Table 3.3.

Although only a few of these "extreme weather conditions" day were present in the database, they could result in bias in the database because of their high frequencies. It is necessary to ignore these accidents as they occurred as a result of "extreme winter

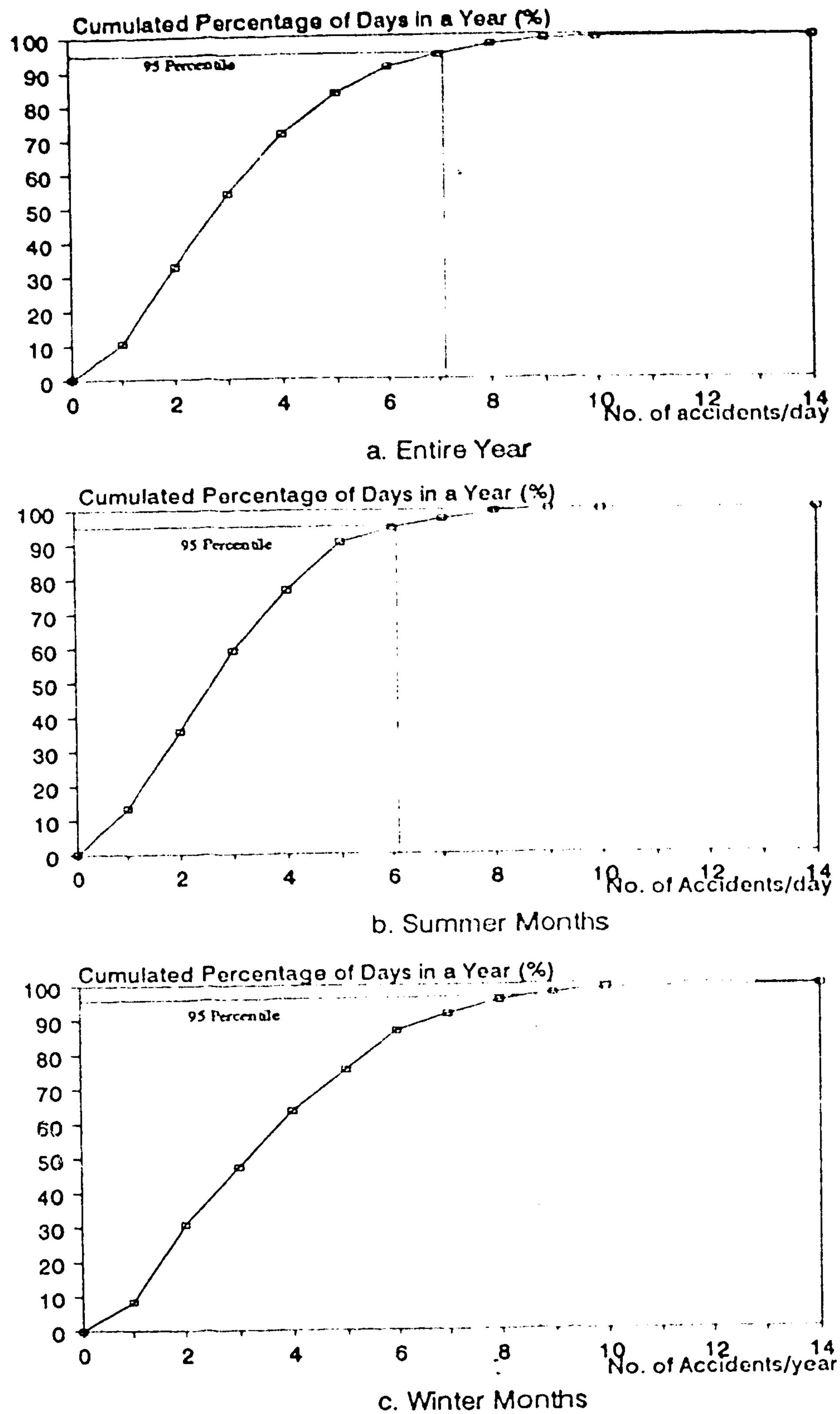


Figure 3.4 Percentile Plots for Accidents in the Summer, Winter and Entire Year

conditions". It is not the intention of this research project to model the effects of these special weather conditions.

Table 3.3 Weather Conditions for the 9 Days with the Worst Accident Records (1985-1988)

	Date	# of accid	Snow (cm)	Rain	Temp (°C)	Day of Week	Comment
1	86/03/24	48	9.9		-5.1	MON	First day of snow in 5 days, heavy snow
2	87/12/29	19	5.2	tr	-9.6	TUE	First day of snow in 5 days, heavy snow, freezing rain
3	88/11/12	15	3.2		-7.2	SAT	Heavy snow,slippery,accidents mostly between 1:30 pm & 4:30 pm
4	88/12/09	14	tr		-7.1	FRI	Slippery, accidents mostly between 1:30 pm & 5:00 pm
5	85/11/16	12	2.0		-8.0	SAT	Very slippery, quite heavy snow
6	86/01/17	11	2.4	tr	-3.9	FRI	Quite heavy snow,freezing rain,thaw & freeze on pavement
7	86/02/21	10	0.3		-18.6	FRI	
8	87/12/02	10			-1.5	WED	Thaw and freeze on pavement
9	87/12/18	10	1.4		-4.4	FRI	Thaw and freeze on pavement

Adjustment factors were therefore applied to accidents that occurred during the 9 days with the worst accident records, to reduce the accident frequency to that comparable to a 95th percentile probability accident frequency level of an average winter day. If the accident frequency was higher on a particular day, a bigger adjustment factor was applied. For example, for a particular day with 21 accidents, the adjustment factor would be 0.43 (9/21=0.43). Any accident that occurred on that day would be multiplied by 0.43.

3.6 Selection of Analysis Period

General accident prediction models comparing traffic safety over a wide variety of conditions often have poor predictability. The poor predictability is due to severe variability in the different underlying driver, vehicle and environment conditions at the time of the accident.

However, there are some time periods when the factors affecting accident occurrences are relatively stable. For modelling purposes, it is desirable to minimize the variability of underlying conditions by focusing on a representative analysis period in which the underlying conditions were relatively homogeneous. By selecting an analysis period with relatively homogeneous travel conditions and characteristics, variability in underlying conditions can be reduced.

The analysis period selected for this study was the weekday daytime period, which contained 74% of the accidents. The accident prediction models developed in the study therefore were applicable to this weekday daytime analysis period only.

3.7 Selection of Data Collection Period

Considering the pilot study nature of this research project, an extensive data collection and modelling program were not only impractical, but also deviated from the original objective of the study. To minimize the efforts required in data collection, the following requirements were considered in selecting the data collection period:

1. It is in a time period when traffic conditions and accident trends are relatively stable.

2. It should be representative of the conditions of a period when most accidents occur.

3. It should allow convenient data collection.

In traffic capacity modelling, the analysis period represents the most critical hour of system operation. In accident analysis, however, the criteria for analysis period selection is quite different as traffic accidents happen at all times of the day. Sharp fluctuations in traffic conditions during peak hours may not be the most appropriate analysis period.

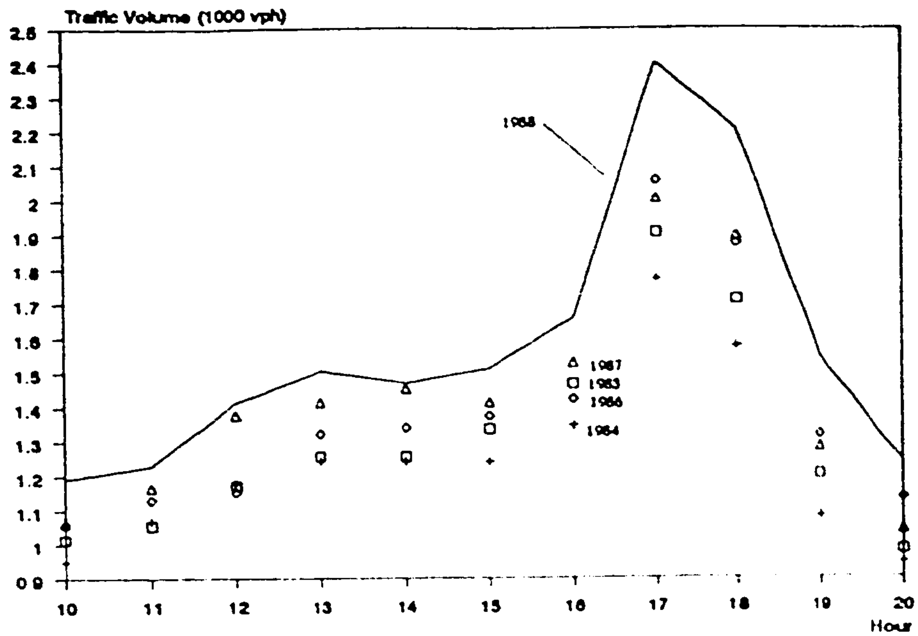
Machine counts from two typical arterial roads are plotted in Figures 3.5a and 3.5b. The solid lines in both figures represent the hourly volume from 1988 and 1989. The symbols represent counts from the years prior to 1988 and 1989. Although the two locations exhibit different degrees of changes in volume over the years, traffic volumes seem to be most stable between 11 a.m. and 4 p.m., with differences averaging to a maximum of approximately 15%.

The daily and hourly traffic variations on various arterial roads are indicated in Figures 3.6a and 3.6b. Traffic volumes on arterials were found to be very stable from 11 a.m. to slightly before 4 p.m.. In Figure 3.6b, the cluster of 4 lines at the bottom of the figure represents the daily fluctuation of traffic from 11 a.m. to 4 p.m.. The 2 lines at the top of the figures are for the 5 and 6 p.m. traffic. It shows that the daily variation of hourly traffic volume is small in the early afternoon and the traffic is more unstable near the p.m. peak.

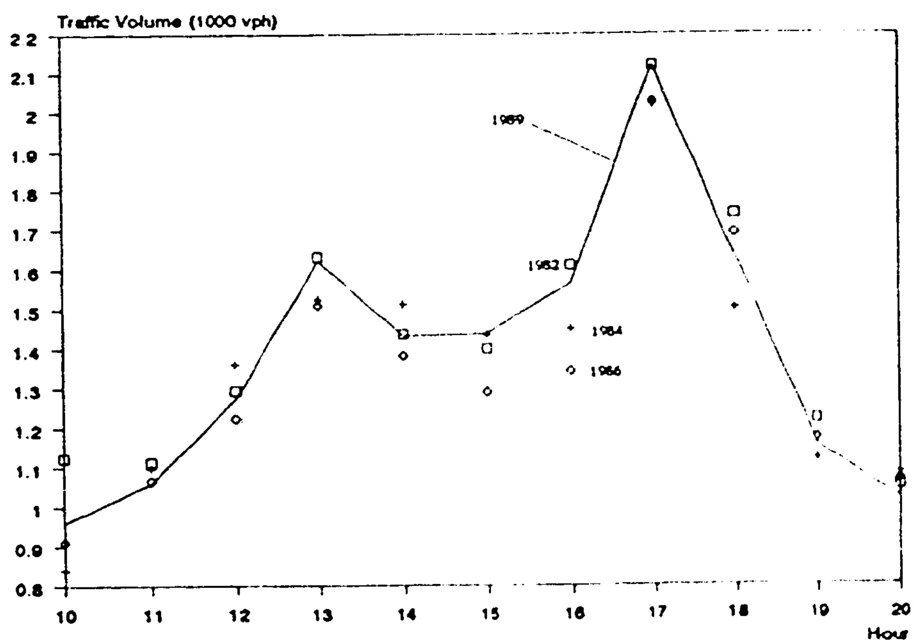
The observation suggests that traffic conditions were relatively stable during the early weekday afternoon between 11 a.m. and 4 p.m.. 38 % of total accidents also occurred during this time period. Therefore, the time period from Monday to Thursday between 11 a.m. and 4 p.m. was considered most appropriate for data collection because of its long duration, stable traffic and safety conditions.

3.8 Selection of Study Locations

As indicated in Section 3.4, of the 429 intersections without traffic signals, 129 were 3-legged intersections and 300 were 4-legged intersections. It was demonstrated in Figure 3.1 that the annual, monthly, daily and hourly accident trends between the 3-legged and 4-legged intersections were similar. The pre-crash manoeuvres and the driver's behaviour leading to the accidents were also similar between the two types of intersections. It was concluded that

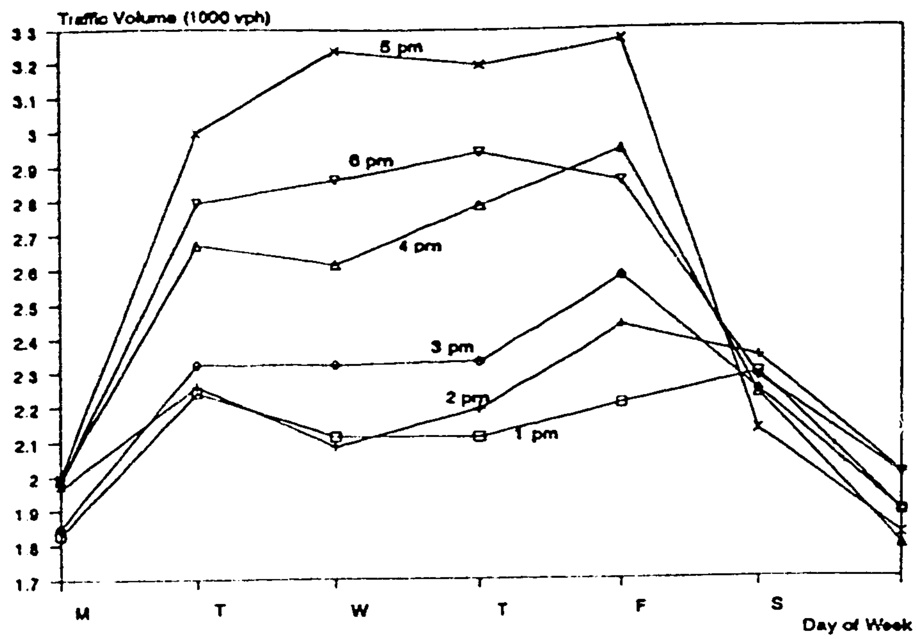


a. Yearly Traffic Variation at 149 Street South of 100 Avenue

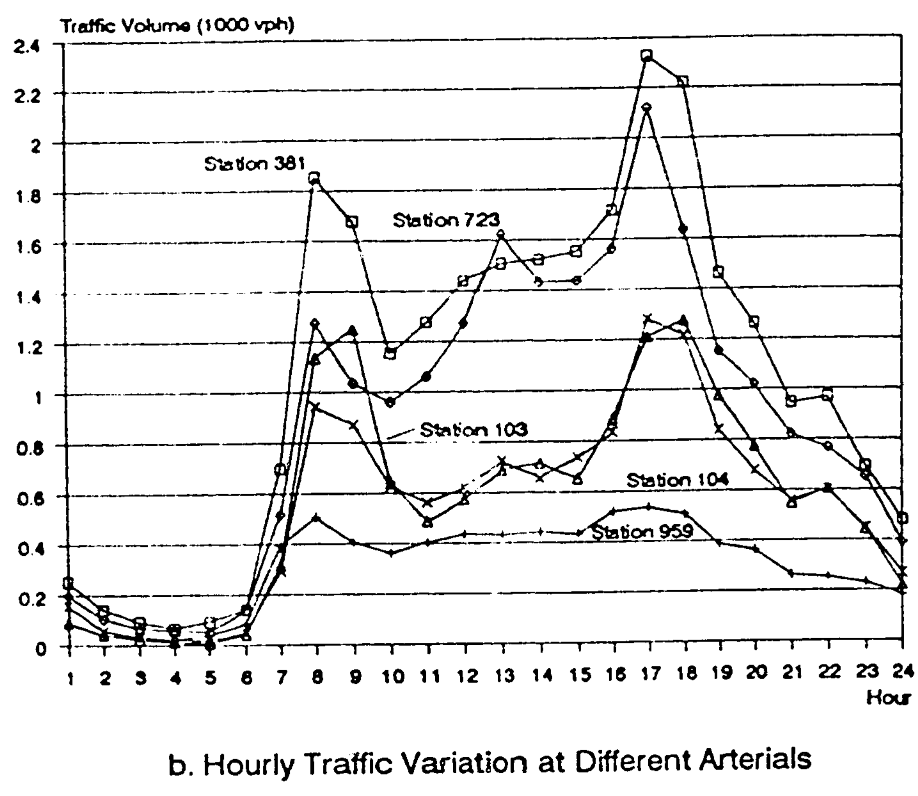


b. Yearly Traffic Variation at 50 Street South of 101 Avenue

Figure 3.5 Hourly Traffic Volume Profiles on Two Arterial Roads



a. Daily variation of Afternoon Traffic at 75 Street North of 82 Avenue



b. Hourly Traffic Variation at Different Arterials

Figure 3.6 Daily and Hourly Traffic Variations on Various Arterial Roads a study concentrating solely on 3-legged intersections (T-intersections) would reasonably represent the conditions at intersections without traffic signals. Because of low accident frequencies at the majority of the intersections, a random selection of study locations would result in severe under-representation of high accident locations. Accordingly, study locations were selected to provide good representation of sites with varying accident frequency records. The following criteria were used as a guideline in selecting the study locations:

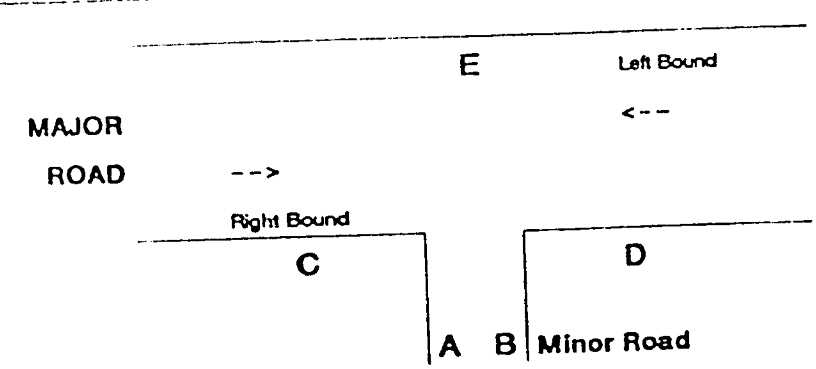
1. Intersections with unusual geometric features were not used.
2. Only stop-controlled T-intersections were considered.
3. Only locations with 4 lane main roads were considered.
4. Locations with parking on the main road were not considered.
5. Intersections with oblique angles were not considered.
6. Offset T-intersections were not considered.
7. Locations with one-way main road traffic were not considered.
8. Intersections on slopes were not considered.
9. Locations with left turn restrictions were not considered.
10. Locations in a curve were not considered.
11. Locations with major construction between 1986 and 1988 were not considered (records for 1985 construction locations were not available).

The above criteria were selected to minimize the variability within the data set so that the analysis could be concentrated on the main objective of the study, which was to compare the relative ability of volume and delay in estimating non-safety. Using the above criteria, 26 T-intersections without traffic signals were chosen. The locations of the selected sites are listed in Table 3.4.

Collision diagrams were constructed for the 26 selected locations. Unusual accident types were examined and, if necessary, discarded from the sample. For example, run-off-road accidents might mean improper sanding or no snow ploughing at the minor road and should

Table 3.4 Intersection Attributes at the 26 Study Locations

Site No.	Locations	No. of Lanes --> Right Bound	No. of Lanes <-- Left Bound	Left-Turn Lane	Median	Parking on Minor Street	Sight Line	Access (see plan below)	Comments
1	41 Ave / 99 St	2	2	-	-	-	-	-	Industrial
2	47 Ave / 99 St	2	2	YES	-	PARK	POOR	E	Strip mall
3	51 Ave / 104 St	2	3	YES	YES	PARK	-	A,E	Restaurant,pizza
4	51 Ave / 105 St	3	3	-	YES	-	-	E	Gas bar,busy mall
5	62 Ave / 122 St	3	2	YES	YES	-	-	-	Residential
6	66 Ave / 99 St	2	2	YES	YES	PARK	POOR	A,B	Pizza,tire store
7	91 Ave / 50 St	2	2	YES	YES	-	-	B	Motel
8	92 Ave / 149 St	2	2	-	-	PARK	-	-	Residential
9	93 Ave / 50 St	3	2	YES	YES	-	-	A,B	Bulb,service road
10	93 Ave / 178 St	2	2	YES	YES	-	-	-	Residential
11	94 Ave / 149 St	2	2	-	-	PARK	-	B	Service road
12	95 Ave / 175 St	2	2	YES	YES	-	-	-	Residential
13	102 Ave / 149 St	2	2	-	-	PARK	POOR	-	Adjacent to signals
14	103 Ave / 149 St	2	2	-	-	PARK	-	-	Adjacent to signals
15	115 Ave / 149 St	2	2	-	-	-	-	D	Small strip mall
16	116 Ave / 142 St	2	2	-	-	-	POOR	-	Bus stop @ E
17	118 Ave / 130 St	2	2	-	-	PARK	-	B,D	Gas bar,apartment
18	127 Ave / 78 St	2	2	-	-	PARK	-	-	No lane marking
19	127 Ave / 90 St	2	2	-	-	PARK	-	A	Service road
20	127 Ave / 119 St	2	2	-	-	PARK	POOR	-	Residential
21	127 Ave / 122 St	2	2	-	-	PARK	-	-	Residential
22	128 Ave / 66 St	2	2	-	YES	PARK	-	C	Strip mall, car dealer
23	134 Ave / 127 St	2	2	YES	-	PARK	-	-	Residential
24	137 Ave / 108 St	2	2	YES	YES	-	-	A,B	Resid/service road
25	137 Ave / 111 St	2	2	YES	YES	-	-	A,B	Resid/service road
26	137 Ave / 135 St	2	2	YES	YES	-	-	A,B	Residential



not be considered a candidate for study locations. After the collision diagrams were constructed, an additional check was made to locate any unusual accident data that might warrant further examination. Accident data for each intersection were plotted on a graph in a cumulated accident frequency plot to reveal inconsistent accident history and unstable accident trends.

To ensure that the 26 sites chosen were representative, the accident trends for the 26 sites were compared to the accident trends for all 129 T-intersections in the database. Figures 3.7a to 3.7e illustrate the results of the comparison. The two sets of data were very similar in almost all aspects, indicating that the selected sites were representative of the other T-intersections.

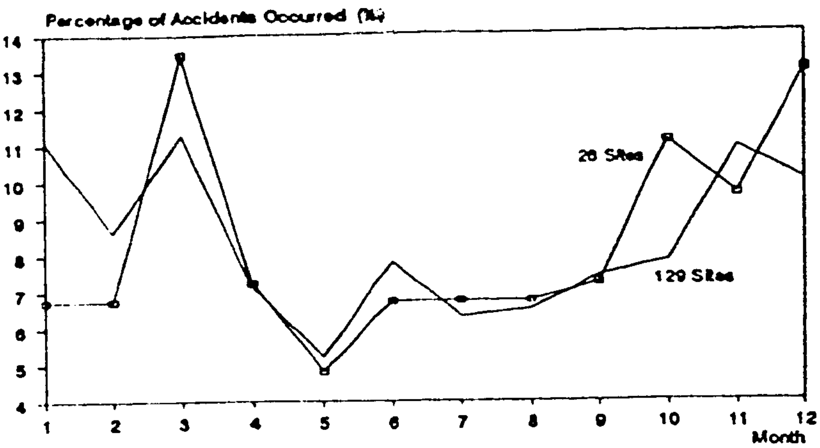
3.9 Data Collection Methods

3.9.1 Intersection Data

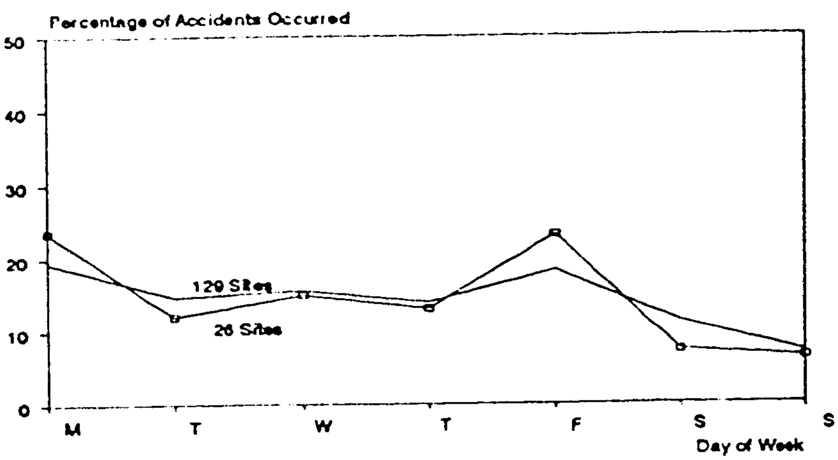
Site attributes recorded at each study location included: sightline, presence of adjacent commercial access, adjacent parking, left-turn bay and raised concrete median. The site attributes recorded are summarized in Table 3.4.

3.9.2 Accident Data

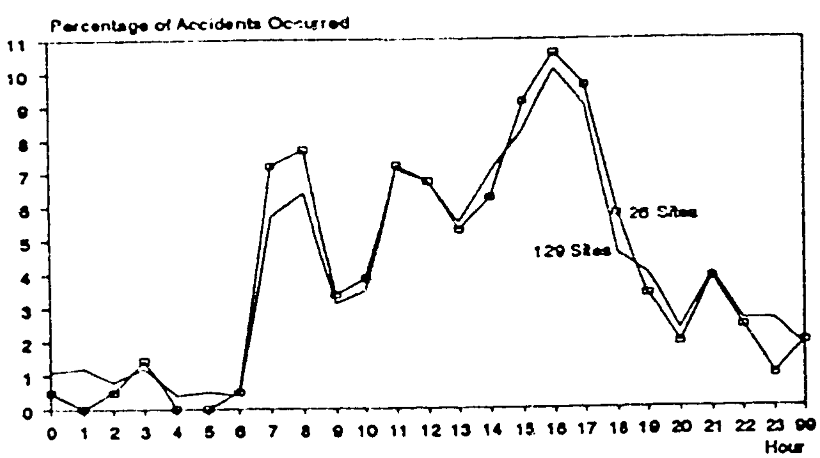
Accident data used for analysis in this study were derived from the accident inventory file of the City of Edmonton. The inventory file contained all reportable on-street motor vehicle accidents in the City from 1985 to the present. Accident data between 1985 and 1988 for the 26 study locations were sorted from the accident records. Only accidents occurring during the weekday daytime analysis period, i.e., from Monday to Friday between 6 a.m. and 9 p.m., were selected.



a. Monthly Accident Distribution

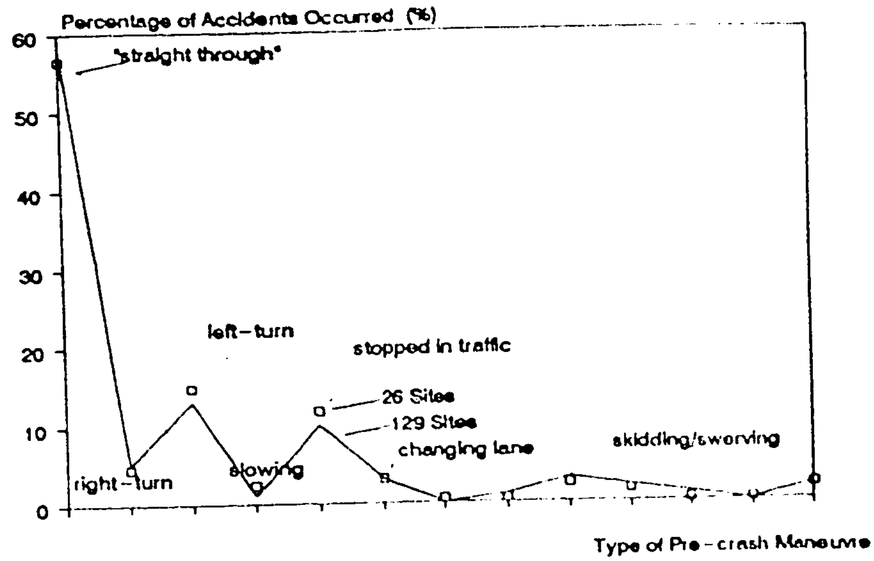


b. Daily Accident Distribution

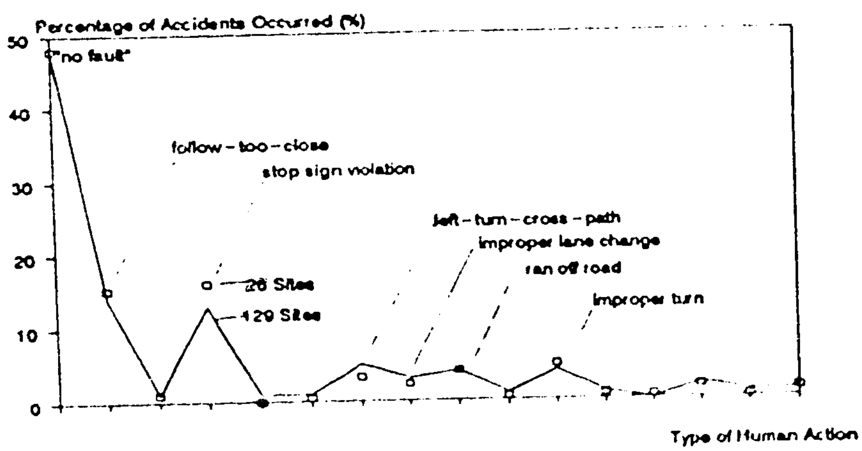


c. Hourly Accident Distribution

Figures 3.7 a–c Accident Trends at the 26 Study Locations



d. Types of Pre-crash Maneuvre



e. Types of Human Action Before Crash

Figures 3.7 d-e Accident Trends at the 26 Study Locations

Accident trends were used instead of actual annual accident numbers as the safety parameter. The accident trends were obtained by determining the average number of accidents over a 4 year period.

The accidents were further categorized into 4 non-priority movement types as well as into pre-crash manoeuvre types for use in disaggregate accident prediction modelling, following the methodology discussed in Chapter 4.

The average 4 year accident frequencies for all accidents and for each individual non-priority manoeuvre accidents and rear-end / right-angle accidents are summarized in Table 3.5.

3.9.3 Volume Data

Intersections without traffic signals often have much lower priority over signalized intersections in the traffic counting program. Availability of volume data for the 26 study locations were therefore limited. Moreover, counts were often collected during peak hours only and counts were often collected over different years.

For modelling purposes, it is desirable to have comparable and current volume and delay data collected simultaneously during the data collection period. Data collection was carried out in January and February of 1990 at the 26 chosen locations in Edmonton. In total, over 15 hours of real time events with 21,069 vehicle arrivals were analyzed and transferred to a computer data base with the assistance of the Traffic Data Input Program (TDIP) (Kyte & Boesen, 1989).

All data were collected on weekdays from Monday to Thursday, from 1 p.m. to 4 p.m., within the chosen data collection period. Data collection was carried out only on days with normal winter conditions.

A video camera was used to collect traffic data required for the analysis. Other than providing an excellent permanent record, the videotaping approach was also economical as only one person was required to operate the camera. After the camera was set up for

Table 3.5 Accident, Volume and Delay Data for the Study Locations

Location	Expected or Actual Number of Accidents (1985–1988)							Volume (vph)						Measured Delay (sec)			Simulated Delay (sec)		
	S	S3	S4	S6	S7	Sre	Sra	V2	V3	V4	V6	V7	V8	D4	D6	D7	D4	D6	D7
1 41 Ave/99 St	3.3	0.80	1.20	0.80	0.80	3.00	1.00	467	13.5	5.6	28.1	27.0	504	5.6	7.5	3.7	7	3	3
2 47 Ave/99 St	5.8	0.05	3.00	1.40	1.40	3.34	5.17	838	43.9	36.6	40.3	69.5	772	38.1	25.8	9.6	44	10	9
3 51 Ave/104 St	8.4	0.05	7.10	1.30	0.05	3.00	1.00	866	50.1	42.2	39.6	76.5	892	26.4	10.1	7	31	4	6
4 51 Ave/105 St	7.7	1.30	0.50	0.80	5.10	5.00	1.00	840	20.8	3.0	29.8	32.8	935	21.4	12.1	11.5	24	4	5
5 62 Ave/122 St	2.0	0.05	1.00	0.05	1.00	0.05	2.00	261	6.2	5.2	17.7	13.5	249	6.6	4.2	2.3	4	2	2
6 66 Ave/99 St	6.6	0.05	2.60	1.30	2.60	2.00	3.00	937	78.5	19.6	41.7	80.9	1111	44.2	9.5	8.3	36	6	9
7 91 Ave/50 St	4.1	0.05	2.90	1.20	0.05	0.05	4.00	796	19.2	16.8	16.8	9.6	940	22	8.9	17.9	26	5	10
8 92 Ave/149 St	4.1	0.05	2.70	0.05	1.40	1.00	2.00	944	17.0	18.3	11.8	10.5	653	15.6	12	3.8	19	10	6
9 93 Ave/50 St	20.1	1.40	10.50	6.90	1.40	0.05	1.00	742	57.5	57.5	109.0	81.8	902	16.9	7.3	3.1	15	3	4
10 93 Ave/178 St	5.7	0.05	1.50	2.30	1.90	1.00	2.00	921	27.2	9.4	26.1	32.4	1005	17	7.4	4.7	23	5	6
11 94 Ave/149 St	1.0	0.05	1.00	0.05	0.05	0.05	1.00	777	3.5	7.0	3.5	5.2	615	21.5	3.9	3.8	22	6	4
12 95 Ave/175 St	1.0	0.05	0.70	0.30	0.05	0.05	1.00	277	9.1	18.3	21.3	21.3	317	9.4	2.7	1.5	5	2	2
13 102 Ave/149 St	10.1	0.05	3.40	5.70	1.00	5.00	4.00	829	51.1	84.1	92.0	33.9	692	27.8	14.9	8.3	28	4	5
14 103 Ave/149 St	4.0	1.30	1.30	0.05	1.30	2.00	1.00	783	17.5	5.8	7.6	13.6	851	45.9	3.2	4.7	7	3	3
15 115 Ave/149 St	4.9	1.60	0.05	0.05	3.30	3.00	0.05	689	37.7	43.3	67.1	28.0	707	21.8	6.1	7	23	5	7
16 116 Ave/142 St	11.5	0.05	5.30	4.70	1.40	1.80	5.97	567	68.4	39.9	91.1	120.0	459	36.7	6.6	7.4	23	4	6
17 118 Ave/130 St	7.6	0.05	1.20	1.80	4.60	3.00	2.00	801	13.1	2.9	13.1	32.1	801	45.6	9.4	7	24	4	7
18 127 Ave/78 St	5.8	1.00	1.40	0.60	2.90	5.00	1.00	183	9.9	7.4	24.7	17.3	222	3.9	3.2	2.5	2	2	2
19 127 Ave/90 St	10.5	1.00	3.90	2.30	3.30	0.84	9.17	288	25.8	46.8	42.1	103.0	358	15	3.9	4.6	11	2	3
20 127 Ave/119 St	1.0	0.05	1.00	0.05	0.05	0.05	1.00	365	4.0	4.0	4.0	5.3	351	7.1	3.2	2.5	8	2	1
21 127 Ave/122 St	2.0	0.05	0.40	0.60	1.00	0.05	2.00	285	4.6	1.5	7.6	4.6	314	1.7	2.9	3.2	3	1	1
22 128 Ave/66 St	7.8	0.05	3.00	2.50	2.30	1.00	4.00	548	11.2	22.5	18.0	33.7	517	9.7	4.8	3.9	10	3	3
23 134 Ave/127 St	2.0	0.05	2.00	0.05	0.05	0.05	1.00	523	10.4	0.1	5.2	22.4	933	0.05	2.8	1.8	0.05	4	3
24 137 Ave/108 St	7.5	3.80	0.05	0.05	3.80	3.00	2.00	952	5.0	8.8	22.6	15.1	839	21.8	11.8	7.3	18	6	6
25 137 Ave/111 St	0.1	0.05	0.05	0.05	0.05	0.05	0.05	887	4.4	8.8	13.3	13.3	962	26	2.4	6.5	22	4	7
26 137 Ave/135 St	7.0	2.80	1.70	2.50	0.05	3.00	2.00	722	17.6	20.2	37.8	37.8	779	24.3	13.4	9.5	27	7	7

Legend:

S Total number of accidents
S3 Major road right-turn accidents
S4 Minor road left-turn accidents
S6 Minor road right turn accidents
S7 Major road left-turn accidents
Sre Rear-end collisions
Sra Right-angle collisions

V2 Major road rightbound through volume
V3 Major road right-turn volume
V4 Minor road left-turn volume
V6 Minor road right-turn volume
V7 Major road left-turn volume
V8 Major road leftbound through volume

D4 Average delay for minor road left-turn traffic
D6 Average delay for minor road right-turn traffic
D7 Average delay for major road left-turn traffic

recording, the camera operator would be free to record the attributes of the study area as well as other observations.

The video camera was placed at a location where all the non-priority movements could be recorded. In most cases, putting the camera at the minor road about 50 to 100m from the intersection provided the best view of the T-intersection. Figure 3.8 illustrates the location of the camera relative to the intersection. Because of the sub-zero outdoor temperature in Edmonton in January and February, the camera was placed inside a car. Placing the camera inside the car also allowed the videotaping to be carried out inconspicuously. The length of recording at each location was approximately 45 minutes to 1 hour.

Data encoded from the video recording were traffic volumes, delays, vehicle categories and headway distribution. TDIP was used to encode the data from the video traffic survey.

The encoded volume data were then converted to an equivalent one-hour "design hour" volume for comparison and modelling purposes. The adjusted traffic volumes for all manoeuvres are summarized in Table 3.5.

3.9.4 Delay Data

As discussed in the previous section, delay data were also encoded from the video recordings. The camera was placed at a location so that the entire queue length on the minor road could be covered by the view field of the camera.

In encoding the delay data, some judgement was required to estimate the actual delay time incurred to a driver. For instance, the time needed for acceleration and deceleration when leaving and joining the queue was taken into consideration. As some degree of judgement was needed in encoding delays, all delay data were encoded by the same person to provide better consistency. The encoded delay data for each non-priority manoeuvre are summarized in Table 3.5.

3.9.5 Simulated Delay

The delay data encoded from the videotape at each location displayed a high degree of variability. Additional data were necessary to supplement the field delay data collected. A number of calculation methods and simulation models were considered to estimate delays based on the measured traffic volumes. It was found that the application of calculation methods in realistic situations would not be satisfactory because of the limitations as a result of oversimplified assumptions.

Alternatively, simulation models, being microscopic in model nature, have the ability to handle more complex traffic situations. They can greatly reduce data collection time especially for low traffic volume situations. At most intersections without traffic signals, traffic volume is relatively low. The use of a simulation model can simulate a sufficiently long data collection period while keeping the various travel conditions stable.

Because of its ease of application and ability to simulate complex traffic interactions, the KNOSIMO program (Grossmann, 1989) was chosen to estimate traffic delays. However, the model was only developed for 2-lane roadways and was not readily applicable to multi-lane situations. To simulate traffic conditions for multi-lane intersections, it was necessary to address the difference in the conditions considered in the program and at the studied locations. A paper written by Chan and Teply (1991), described the modifications needed to apply the KNOSIMO program to an urban 4-lane situation, and is included in Appendix A. The simulated delay data are included in Table 3.5.

3.10 Database

The non-safety, traffic volume and delay data are summarized in a database, as shown in Table 3.5. The database at this stage is ready for use in further analysis and accident prediction modelling.

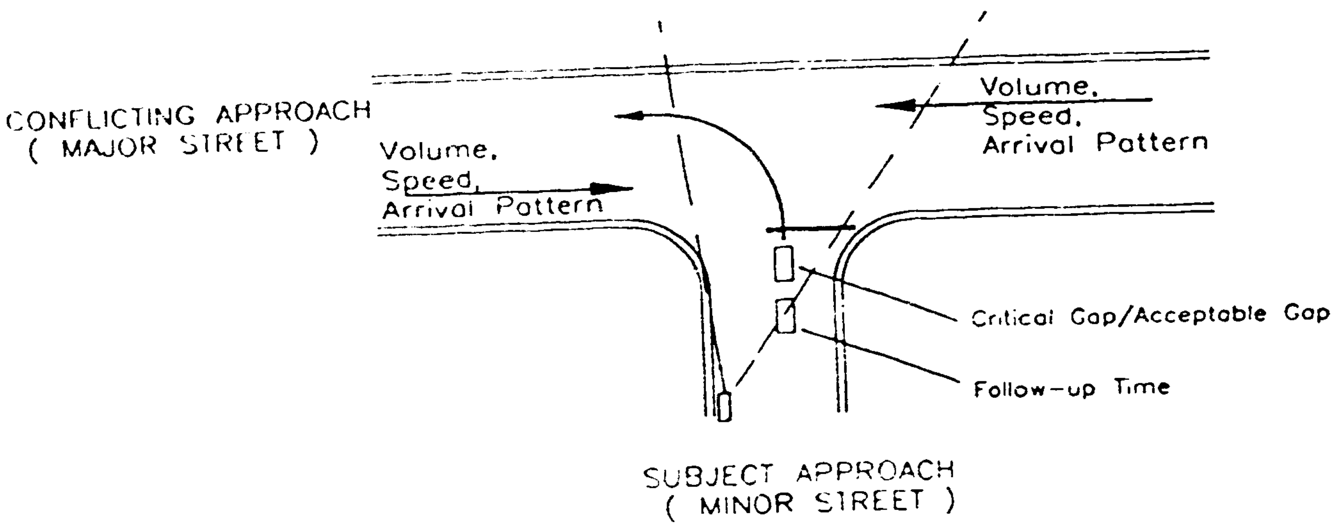


Figure 3.8 Camera Location and View Angle

4. ACCIDENT PREDICTION MODELS

At this stage of the research project, the framework of the research project has been determined. In establishing the research framework, a preliminary analysis on available accident and traffic records was carried out, as was discussed in detail in Chapter 3. The major elements of the research framework are:

1. analysis period
2. study locations
3. parameters to collect
4. data collection method

Traffic volume and delay data were collected at 26 selected intersections. The collected data were encoded with the use of the Traffic Data Input Program, TDIP (Kyte and Boesen, 1989). Additional delay data was also simulated through the use of the KNOSIMO simulation program (Grossmann, 1988).

With all the data in place, the next stage of the research was to determine the model structures and formulate the actual accident prediction models by using the various volume and delay data collected, as well as the 4 years of accident records.

4.1 Application of the Risk and Exposure Concept

The risk and exposure concept were used in formulating the model structures. The concept states that the non-safety at an intersection is the product of the risk presented to the traffic entering the intersection and the exposure of the traffic with the risk.

Table 4.1 Risk and Exposure Concept

Safety	= Risk x	Exposure
Sum of Flow Model	1	Sum of all entering flows
Product of Flow Model	Major road flows	Minor road flows
Delay-based Models	Average delay experienced by non-priority flows	Non-priority flows

As illustrated in Table 4.1, in the sum-of-flow model, safety is the number of accidents which have occurred at an intersection. It is estimated by the total number of vehicles entering the intersection. The total number of vehicles entering the intersection can be regarded as an exposure parameter. Since the risk factor is not dealt with, the sum-of-flow model should be considered as an exposure model instead of a risk and exposure model.

In the other two accident prediction models, the risk and exposure parameters are often represented by the major and minor road flows, or some combination of the two flows. In a simple product-of-flow model, the major road flow is used as a surrogate of risk to the merging minor road traffic. The minor road flow is the measure of exposure for the traffic being presented with the risk of merging. In that sense, the safety parameter in a simple product-of-flow model is restricted to represent the non-safety situation for the minor road traffic only. Based on the same concept, the simple product-of-flow model can be transformed into a simple delay-based accident prediction model by using non-priority flow as the exposure measure, and the average delay experienced by the non-priority flow as the risk measure

The details of the model structure as well as the forms and interactions of the various parameters was then determined through the use of statistical modelling tools. The various non-safety, risk and exposure model parameters available for modelling are summarized in Table 4.2.

4.2 Statistical Model Structure Options

Literature research indicated that a non-standard statistical model would be needed for accident prediction modelling because of the negative binomial distribution of accident data samples. The GLIM program (Numerical Algorithm Group, 1987) is capable of handling unique data distributions and has been successfully used in numerous accident prediction modelling research projects. For this reason, the GLIM program was chosen for this study.

The three components of a GLM, as described in Chapter 2, are the probability distribution function of the data; a linear predictor that describes the linear regression function; and a link function that relates the linear predictor to the mean. It is possible to specify different models with either an additive or a multiplicative link function.

4.2.1 Additive Model Structures

A model with an additive model structure can be formulated by using an identity link function. The resulting model structure will be as follows:

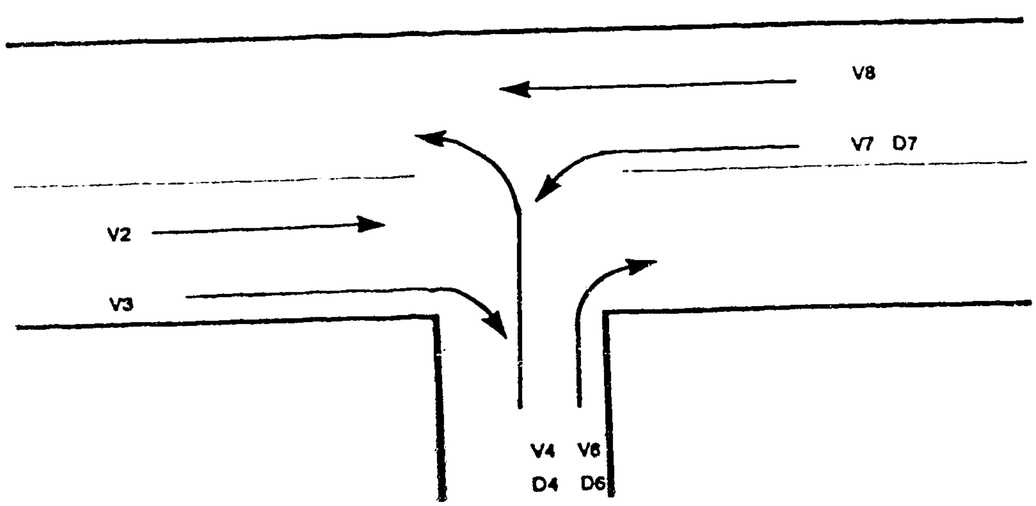
$$S = a_0 + a_1 R_1E_1 + a_2 R_2E_2 + a_3 R_3E_3 + \ldots$$

where

S = measure of non-safety, the dependent variable

Table 4.2 Accident Prediction Model Parameters

Parameter	Description
Measures of non-safety	
S	Total number of accidents in 4 years
Sre	No. of rear-end collisions over 4 years
Sra	No. of right-angle collisions over 4 years
S3	No. of accidents related to main road right-turn movements over 4 years
S4	No. of accidents related to minor road left-turn movements over 4 years
S6	No. of accidents related to minor road right-turn movements over 4 years
S7	No. of accidents related to main road left-turn movements over 4 years
Volume parameters	
V2	Main road rightbound through volume
V3	Main road right-turn volume
V4	Minor road left-turn volume
V6	Minor road right-turn volume
V7	Main road left-turn volume
V8	Main road leftbound through volume
Delay parameters	
D4	Average stopped delay for minor road left-turn flows
D6	Average stopped delay for minor road right-turn flows
D7	Average stopped delay for main road left-turn flows



R_i, E_i = model parameters, which can be the risk and exposure parameters

a_i = coefficients to be modelled by GLIM

4.2.2 Multiplicative Model Structures

A model with a multiplicative model structure can be formulated by using a log-linear link function. The model structure below is a typical multiplicative model:

$$S = a_1 R_1^{b1} E_1^{c1} * a_2 R_2^{b2} E_2^{c2} * a_3 R_3^{b3} E_3^{c3} + \ldots$$

where

S = measure of non-safety, the dependent variable

R_i, E_i = model parameters, which can be the risk and exposure parameters

a_i = model coefficients to be modelled by GLIM

$^{bi}, ^{ci}$ = exponential coefficients to be modelled by GLIM

To specify a multiplicative model structure in GLIM, a log-transformation is required to alter the non-linear model structure to linear. The log-linear link function can then relate the transformed linear predictor to the transformed observed values. The transform model structure is as follows:

$$\log(S) = \log(a_1) + b_1 * \log(R_1) + c_1 * \log(E_1) + \log(a_2) + b_2 * \log(R_2) + c_2 * \log(E_2) + \ldots$$

4.3 Level of Aggregation and Potential Model Parameters

In an accident prediction model, the main objective of the modelling process is to provide satisfactory model prediction results that will match the observed values in the field. In predicting accidents at a particular intersection, certain attributes of the intersection, as required by the model structure, will be used. The resulting output of the model will be a predicted non-safety value. If the predicted non-safety value matches well with the actual observed non-safety value, the model is considered accurate.

The non-safety measure chosen has a primary effect on the set-up of the model structure. Once a non-safety measure is chosen, the model parameters should be selected accordingly so that they are relevant to the type of accident being considered.

The literature research indicated that different levels of aggregation could be used in accident prediction modelling. Some of the more aggregate models predicted the number of daily accidents by using average daily traffic. Other less aggregate models predicted accidents during different peak hour periods by using traffic volumes of specific conflicting movements at the intersection. Attempts were made in this research to try to use different levels of aggregation for the non-safety measure.

4.4 Disaggregate Accident Prediction Model For Each Basic Accident Type

The modelling approach similar to that adopted by Pickering et al (1986) and Hauer et al (1989) was used to examine the safety situation for each basic accident type. Various types of accidents can be identified based on the pre-crash manoeuvres of the vehicles involved in the collision. As shown in Figure 4.1, it is possible to have 13 basic accidents types at a T-intersection. Table 4.3 illustrates the frequencies of 13 basic accident types at the 26 study locations.

The problem that became immediately obvious was the low accident frequencies for most of the basic accident types. For example, of the 129 accidents that occurred at the 26 T-intersections, only 10 accidents were rightbound rear-end collisions. Moreover, the right-bound rear-end collisions were absent in all but 4 of

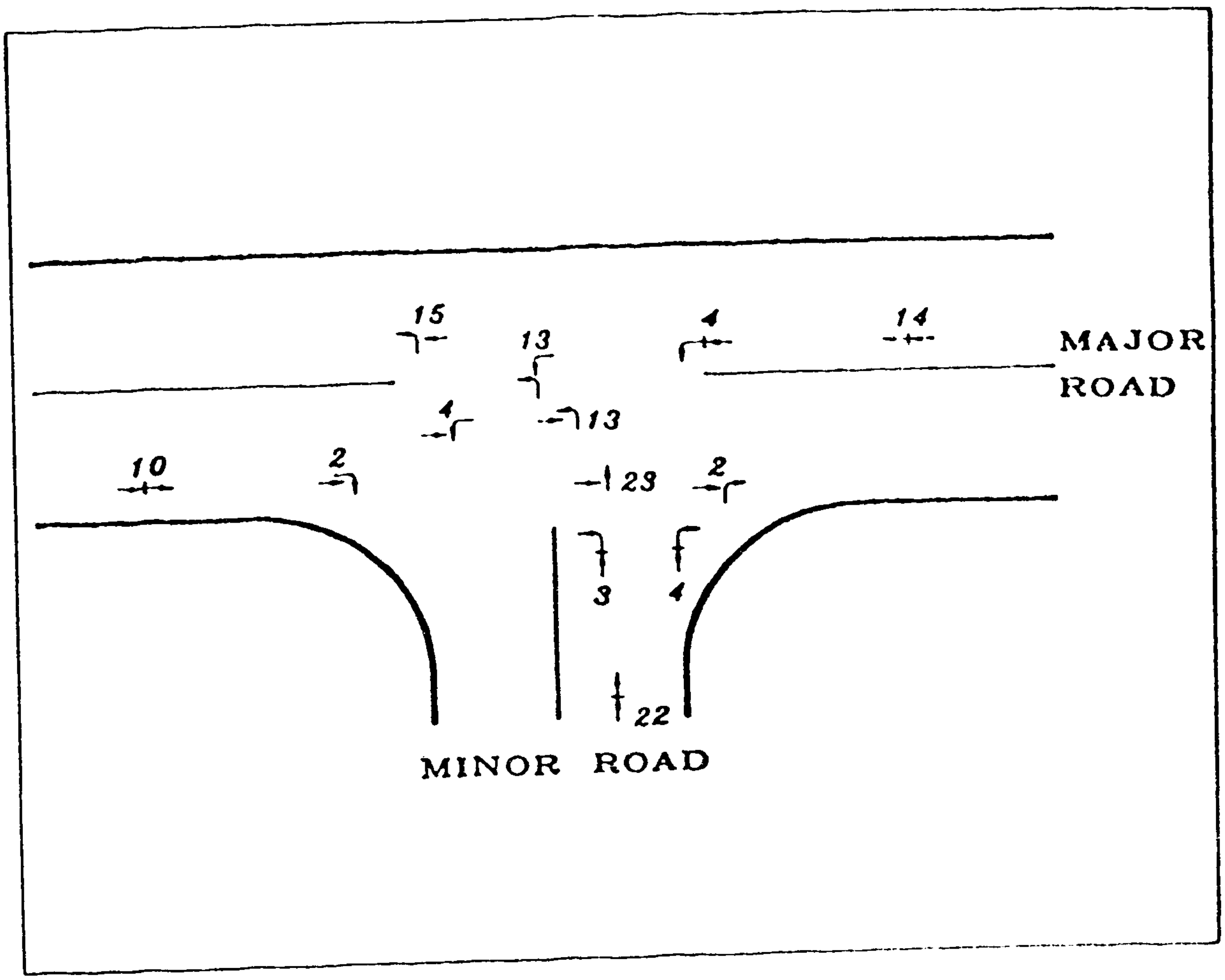


Figure 4.1 13 Basic Accident Types and Frequencies at 26 Study Locations

Table 4.3 13 Basic Accident Types at the Study Locations

Site No.	Locations	13 Basic Accident Types 1	2	3	4	5	6	7	8	9	10	11	12	13	Sum
1	41 Ave/99 St	*				*			*					*	4
2	47 Ave/99 St					***		**	****			*		*	11
3	51 Ave/104 St				**	*			*						4
4	51 Ave/105 St	*							*			*	*	**	6
5	62 Ave/122 St					*						*			2
6	66 Ave/99 St			**							*	*		*	5
7	91 Ave/50 St			**				**							4
8	92 Ave/149 St			*		*							*		3
9	93 Ave/50 St			*	*	***		**	*****		*	*			15
10	93 Ave/178 St							*	*			*			3
11	94 Ave/149 St					*									1
12	95 Ave/175 St					*		*							2
13	102 Ave/149 St			*				**	****	*	*	*			10
14	103 Ave/149 St	*		*										*	3
15	115 Ave/149 St	*												**	3
16	116 Ave/142 St			***				***	*		*	*			9
17	118 Ave/130 St							**						***	5
18	127 Ave/78 St	*				*			*				**	*	6
19	127 Ave/90 St		*			*	***	****	**			**			13
20	127 Ave/119 St			*											1
21	127 Ave/122 St							*				*			2
22	128 Ave/68 St			*			*	*		*				*	5
23	134 Ave/127 St					*									1
24	137 Ave/108 St	**	*									**		*	6
25	137 Ave/111 St														0
26	137 Ave/135 St	**						**	*						5
	Sum	10	2	13	3	15	4	24	22	2	4	13	4	14	129

the 26 study locations. It was concluded from this observation that, because of low accident frequencies, it is unlikely that accident modelling for intersections without traffic signals could be done for each discrete basic accident type. Therefore, no models were developed at the basic accident type level.

4.5 Disaggregate Right-angle and Rear-end Accident Prediction Models

Accident data as recorded in police accident reports often categorize accidents into distinctive descriptive accident types. Common types of accident categories used in police accident reports are rear-end collisions, right-angle collisions, accidents involving left-turning vehicles and head-on collisions.

Collision diagrams were drawn for all the accidents which occurred at the 26 study locations. It was found that most of the accidents could be grouped into two accident types: rear-end collisions and right-angle collisions. Accident frequencies for other accident types were low and were also difficult to classify into any particular accident types. Figure 4.2 illustrates how the different types of accidents can be grouped into the rear-end and right-angle collisions.

4.5.1 Rear-end Accident Prediction Model (Model 512)

A typical rear-end accident model (Model 512) has the following structure:

$$S_{RE} = a + a_0(V_2*V_3) + a_1 V_4*D_4) + a_2(V_6*D_6) + a_3(V_7*D_7)$$

where

S_{RE} = number of rear-end collisions from 1985 to 1988

V_i,D_i = model exposure and risk parameters

a_i = coefficients to be modelled by GLIM

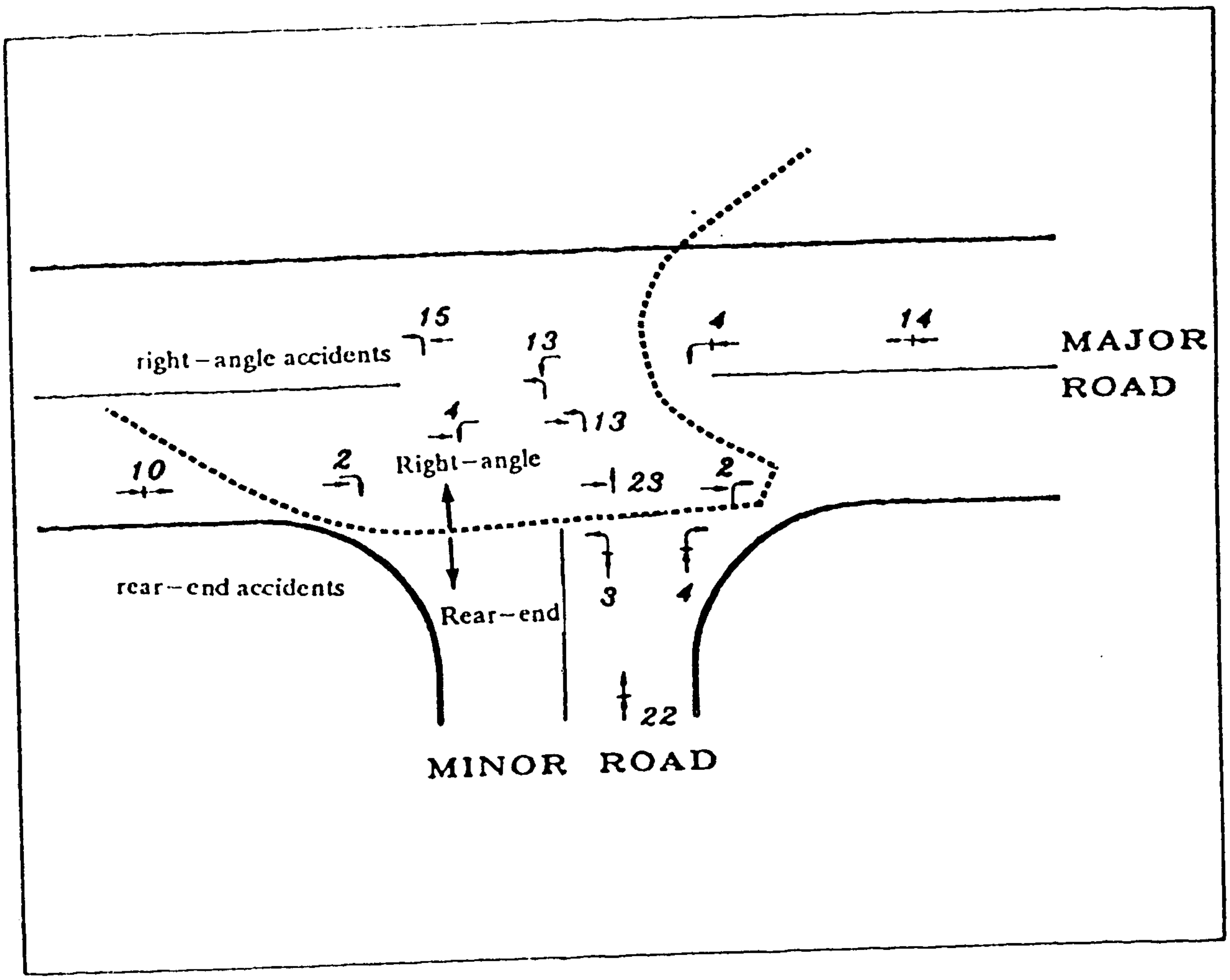


Figure 4.2 Categorization of Rear-end and Right-angle Accidents

The first set of parameters, (V_2*V_3), are products of the rightbound through and right-turn volumes. It is used to predict the number of rear-end collisions which may occur to the rightbound traffic movements. The use of a product format assumes that rear-end collisions for the rightbound traffic are related to the potential conflicts between major road rightbound through traffic and major road right-turn traffic.

Based on similar assumptions, the other three sets of parameters, (V_4*D_4), (V_6*D_6) and (V_7*D7), were developed for the rear-end collisions related to the other non-priority flows. Delay and volume were used to represent risk and exposure for the other three non-priority movements.

4.5.2 Right-angle Accident Prediction Model (Model 612)

A typical right-angle accident model (Model 612) has the following structure:

$$S_{RA} = a + a_0 (V_4*D_4) + a_1 (V_7*D_7) + a_2 (V_6*D_6)$$

where

S_{RE} = number of right-angle collisions from 1985 to 1988

V_i, D_i = model exposure and risk parameters

a_i = coefficients to be modelled by GLIM

The first set of parameters, (V_4*D_4), are products of the minor road left-turn volumes and the delays experienced by that movement. Delay is used as a risk measure for the non-priority movements. Traffic volume for the movement being delayed is used as a exposure measure. Three sets of parameters are included in the model: minor road left-turns (V4), minor road right-turns (V6), and major road left-turns (V7).

4.6 Disaggregate Non-priority Movement Accident Prediction Models (Models 311 to 314)

At intersections without traffic signals, drivers in the non-priority traffic stream have to use judgement to either reject or accept a gap. Focusing on non-priority movements reflects a logical cause-and-consequence relationship and therefore have a potential for satisfactory accident prediction.

Table 4.4 illustrates the type of basic accident types that can be related to the 4 turning movements at a T-intersection. It can be seen from Table 4.4 that most accident types relate predominantly to one turning movement, with the exemption of rear-end collisions and accident types 6 and 7. Some interpretation was needed to group these accidents under these 4 turning movements. As all four turning movements do not have the first priority to carry out the manoeuvre, this approach can be considered as the non-priority movement accident prediction approach.

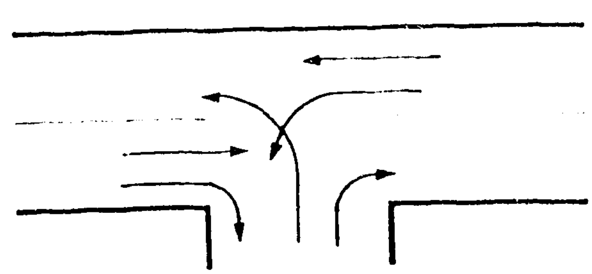
The rightbound right-turn traffic was also considered a non-priority traffic movement. A rightbound right-turn vehicle travelling on the far side will have to yield to the traffic on the near side lane, before it can change lanes to make a right turn manoeuvre.

By grouping the basic accident types into four non-priority movement related groups, the number of accident groups in the data can be reduced from 13 to 4, resulting in higher accident frequencies for the non-priority accident types. However, as seen in Table 4.4, accident frequencies are still low for two of the non-priority movements.

t44

Table 4.4 Accident Categories By Non-Priority Movements

Basic Accident Types	Frequency	Non-priority Movement Categories			
1	10				
2	2				
3	13				
4	3				
5	15				
6	4				
7	24				
8	22				
9	2				
10	4				
11	13				
12	4				
13	14				



4.6.1 Major Road Right-turn Accident Prediction Model (Model 311)

The traffic movements related to the major road right-turn accident prediction model are the major road through and right-turn volumes (refer to Figure 4.3a). This model has the following structure:

$$S_3 = e^{a3} E_3^{b3} R_3^{c3}$$

where

S_3 = number of accidents related to the major road right-turn movement from 1985 to 1988

E_3 = exposure parameter for the major road right-turn traffic

R_3 = risk parameter for the major road right-turn traffic

$a3, b3, c3$ = coefficients to be modelled by GLIM

The major road right-turn movement does not have conflicting movements that are as obvious as the other non-priority movements. It is therefore difficult to formulate a set of risk and exposure parameter to represent a cause-and-effect relationship.

An attempt was made by using the major road right-turn traffic volume as the exposure parameter and the major road leftbound through traffic volume as the risk parameter.

The resulting accident prediction model (Model 311) for major road right turn traffic is as follows:

$$S_3 = e^{a3} V_2^{b3} V_3^{c3}$$

where

S_3 = number of accidents related to the major road right-turn movement from 1985 to 1988

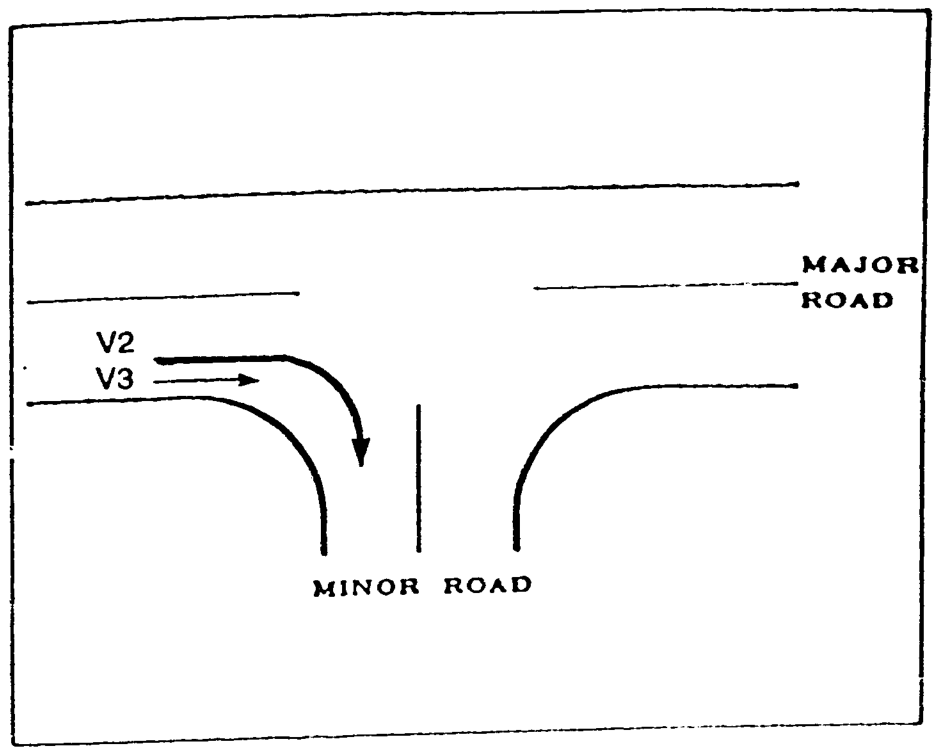


Figure 4.3a Major Road Right–turn Accidents Related Movements

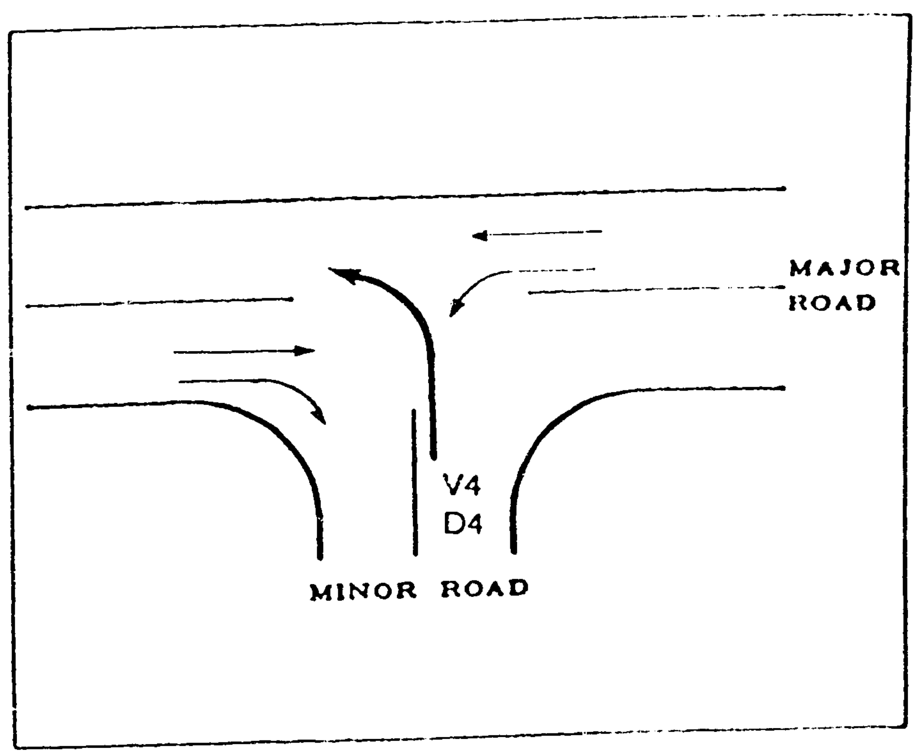


Figure 4.3b Minor Road Left–turn Accidents Related Movements

V_3 = exposure parameter, represented by the volume of right-turn traffic

V_2 = risk parameter, represented by the volume of rightbound through traffic on the major road

$a3,b3,c3$ = coefficients to be modelled by GLIM

4.6.2 Minor Road Left-turn Accident Prediction Model (Model 312)

The traffic movements related to the minor road left-turn accident prediction model are the minor road left-turn traffic and the sum of the major road traffic. Refer to Figure 4.3b, this model has the following structure:

S_4 = $e^{a4} E_4^{\ b4} R_4^{\ c4}$

where

S_4 = number of accidents related to the minor road left-turn movement from 1985 to 1988

E_i = exposure parameter for the minor road left-turn traffic

R_i = risk parameter for the minor road left-turn traffic

$a4,b4,c4$ = coefficients to be modelled by GLIM

The conflicting movements in this case are much more obvious. In this case, the delays experienced by the minor road left-turn traffic are used as the risk parameter. The minor road volume itself represents the exposure. The resulting accident prediction model (Model 312) for minor road traffic is as follows:

S_4 = $e^{a4} V_4^{\ b4} D_4^{\ c4}$

where

S_4 = number of accidents related to the minor road left-turn movement from 1985 to 1988

V_4 = exposure parameter, represented by the minor road left-turn traffic

D_4 = risk parameter, represented by the delay experienced by the minor road left-turn traffic

a_4, b_4, c_4 = coefficients to be modelled by GLIM

4.6.3 Minor Road Right-turn Accident Prediction Model (Model 313)

The minor road right-turn model was developed based on the same assumptions used for the minor road left-turn accident prediction model (see Figure 4.3c). The model (Model 313) has the following structure:

$$S_6 = e^{a6} V_6^{b6} D_5^{c6}$$

where

S_6 = number of accidents related to the minor road right-turn movement from 1985 to 1988

V_6 = exposure parameter, represented by the volume of minor road right-turn traffic

D_6 = risk parameter, represented by the delay experienced by the minor road right-turn traffic

a_6, b_6, c_6 = coefficients to be modelled by GLIM

4.6.4 Major Road Left-turn Accident Prediction Model (Model 314)

The major road left-turn model was also developed based on the same risk and exposure concept using delay as a risk parameter (see Figure 4.3d). The model (Model 314) has the following structure:

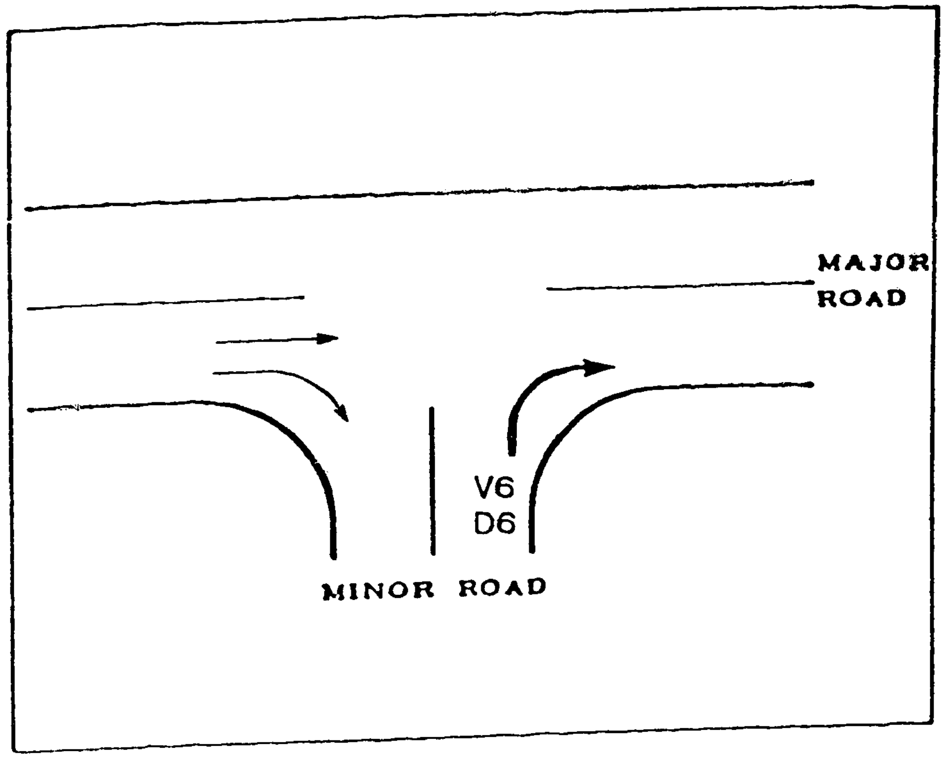


Figure 4.3c Minor Road Right–turn Accidents Related Movements

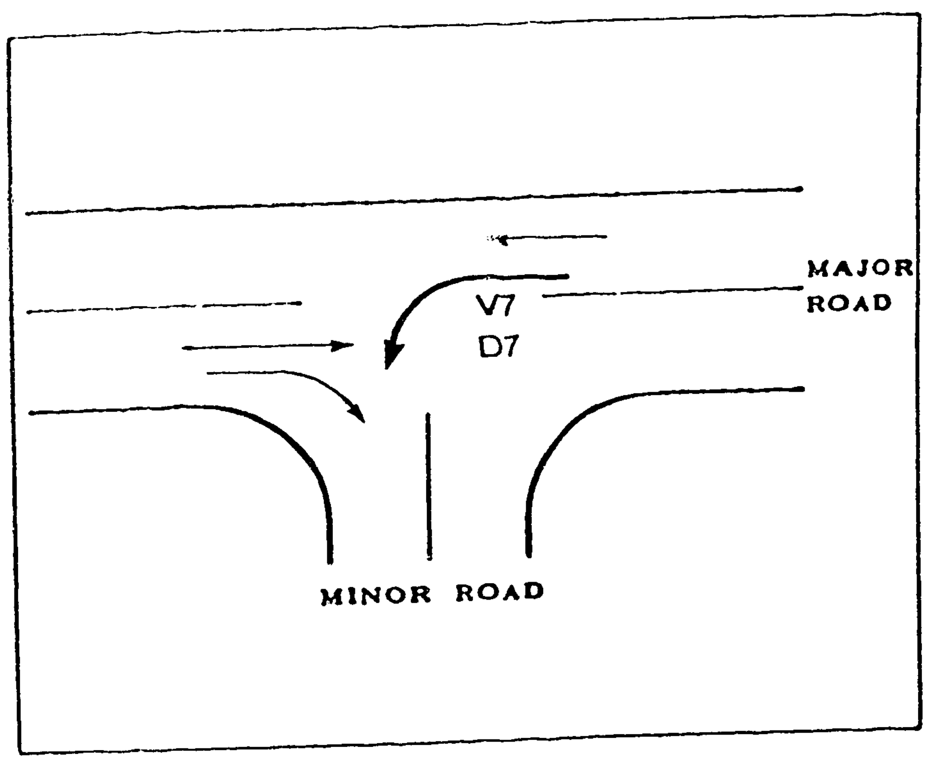


Figure 4.3d Major Road Left–turn Accidents Related Movements

$$S_7 = e^{a7} V_7^{b7} D_7^{c7}$$

where

S_7 = number of accidents related to the major road left-turn movement from 1985 to 1988

V_7 = exposure parameter, represented by the volume of major road left-turn traffic

D_7 = risk parameter, represented by the delay experienced by the major road left-turn traffic

$a7, b7, c7$ = coefficients to be modelled by GLIM

4.7 Aggregate Accident Prediction Models With Volume based Parameters

The models developed up to this point are relatively disaggregate models. The disaggregate models have low accident frequency. For this reason, the aggregate approach will be used for the rest of the models to be developed. The safety parameter in the models will be the number of accidents that occurred at the intersection over the 4 year period of time from 1985 to 1988. The different models developed have varying degrees of complexity in model structures.

4.7.1 Sum-of-flow Accident Prediction Models (Models 111 to 112)

Two sum-of-flow model structures were examined. The first one (Figure 4.4a) is a less complex model structure, with the non-safety parameter related directly to the sum of all traffic movements entering the intersection. The structure of the first sum-of-flow model (Model 111) is:

$$S = a_0 + a_1 [V_2+V_3+V_4+V_6+V_7+V_8]$$

where

S = number of accidents from 1985 to 1988

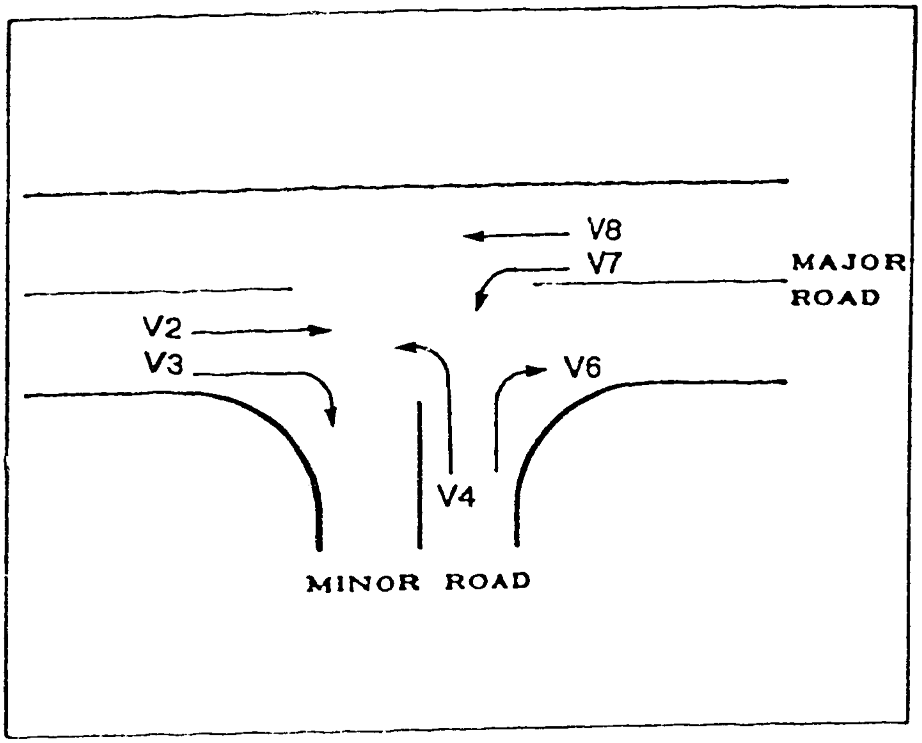


Figure 4.4a Parameters in Sum-of-flow Model (Model 111)

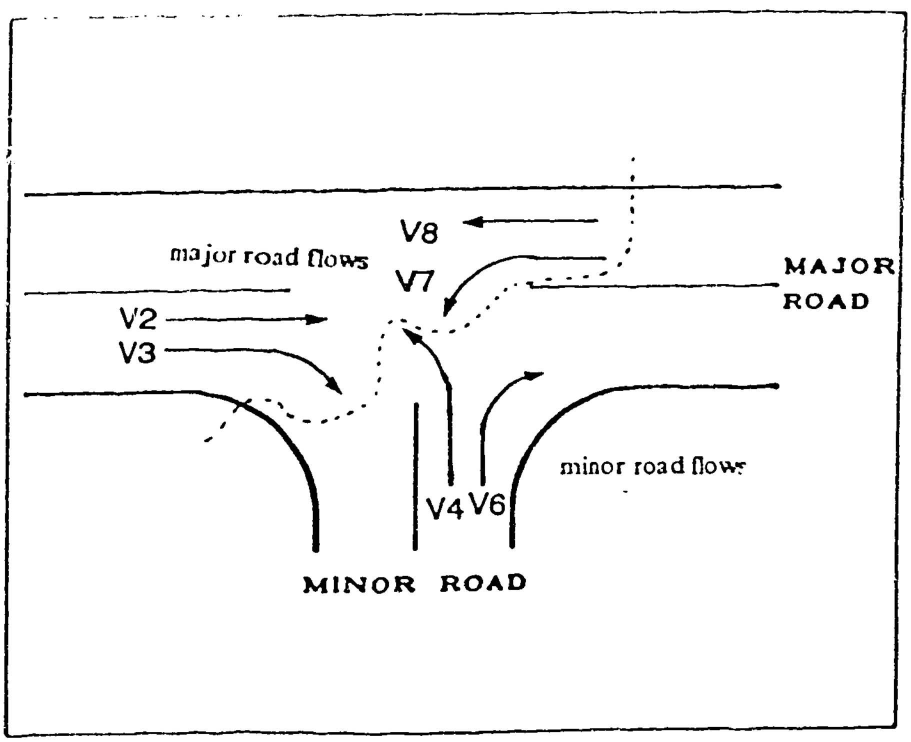


Figure 4.4b Parameters in Product-of-flow Model (Model 122)

V_i = various traffic movement volumes

a_i = coefficient to be modelled by GLIM

In the second model, the sum of conflicting volumes of each accident type is added to form a separate model parameter. Coefficients are developed separately for each sum-of-flow parameter. For example, the first group of parameters, $[v_2+v_3+v_4+v_7+v_8]$, represents the sum of the conflicting volumes involved with the minor road left-turn movement, whereas the third group of parameters, $[v_2+v_3+v_7]$, represents the sum of the conflicting volumes involved with the major road left-turn movement. The structure of this model (Model 112) is:

$$S = a_0 + a_1 [V_2+V_3+V_4+V_7+V_8] + a_2 [V_2+V_6] + a_3 [V_2+V_3+V_7]$$

where

S = number of accidents from 1985 to 1988

V_i = various traffic movement volumes

a_i = coefficients to be modelled by GLIM

4.7.2 Product-of-flow Accident Prediction Models (Models 121 to 124)

Four product-of-flow models with varying degrees of complexity were developed. The models differ in the way the volume parameters were aggregated within the model.

The first product-of-flow model employs three groups of parameters that represent the sum of the product-of-flow for each non-priority movement type. The model structure (Model 121) is shown as follows:

$$S = a_0 + a_1 [V_4*(V_2+V_3+V_7+V_8)] + a_2 [V_6*V_2] + a_3 [V_{\text{[illegible]}} * V_2 \text{ [illegible]}]$$

where

S = number of accidents from 1985 to 1988

V_i = various traffic volumes

a_i = coefficients to be modelled by GLIM

The second model groups all traffic volumes into two categories: minor road traffic and major road traffic. The minor road traffic can be considered as the exposure. The major road traffic can be considered as the risk to the minor road traffic. The model structure (Model 122) is shown as follows as well as in Figure 4.4b:

$$S = a_0 + a_1 [(V_2+V_3+V_7+V_8) * (V_4+V_6)]$$

where

S = number of accidents from 1985 to 1988

V_i = various traffic volumes

a_i = coefficients to be modelled by GLIM

The third model has the same model parameters, except that the model structure is multiplicative / exponent instead of the additive format of the second model. The model structure (Model 123) is as follows:

$$S = e^{a0} * (V_2+V_3+V_7+V_8)^{a1} * (V_4+V_6)^{a2}$$

where

S = number of accidents from 1985 to 1988

V_i = various traffic volumes

a_i = coefficients to be modelled by GLIM

The fourth product-of-flow model (Model 124) has the most complex model structure, indicated as follows:

$$S = e^{a0} (V_2)^{b1} (V_4)^{b2} (V_6)^{b3} (V_7)^{b4} (V_2+V_3)^{b5} (V_2+V_3+V_7+V_8)^{b6}$$

where

S = number of accidents from 1985 to 1988

V_i = various traffic volumes

a_i, b_i = coefficients to be modelled by GLIM

4.8 Aggregate Accident Prediction Models With Delay-based Parameters (Models 211 to 216)

Six delay-based models were developed. In the six models, delay was used to replace traffic volume as the risk parameter. The volume of the non-priority traffic was used as the exposure parameter.

4.8.1 Delay-based Accident Prediction Model 1 (Model 211)

The first delay-based model is a function of the sum of the risk and exposures of the four non-priority movements. Delay is used as a surrogate of risk, whereas volumes are being used as a surrogate for exposure. The model structure of the first delay based model (Model 211) is linear in format and is shown below:

$$S = a_0 + a_1 (V_2{*}V_3) + a_2 (D_4{*}V_4) + a_3 (D_6{*}V_6) + a_4 (D_7{*}V_7)$$

where

S = total number of accidents from 1985 to 1988

Vi = exposure parameter, which is the various traffic volumes

D_i = risk parameter, measured by the delay to drivers

a_i = coefficients to be modelled by GLIM

4.8.2 Delay-based Accident Prediction Model 2 (Model 212)

The second delay-based model simply grouped all parameters together and formed the product of all the parameters. The model structure is also additive, as follows (Model 212):

S	=	$a_0 + a_1 (V_2 * V_3 * D_4 * V_4 * D_6 * V_6 * D_7 * V_7)$
where		
S	=	total number of accidents from 1985 to 1988
V_i	=	exposure parameter, which is the various traffic volumes
D_i	=	risk parameter, measured by the delay to drivers
a_i	=	coefficients to be modelled by GLIM

4.8.3 Delay-based Accident Prediction Model 3 (Model 213)

The third delay-based model has a multiplicative model structure in the form of a power function. The parameters in this model are similar to the first model, except that the model structure is multiplicative instead of additive. The model structure is shown as follows (Model 213):

S	=	$e^{a0} (V_2 * V_3)^{b1} (D_4 * V_4)^{b2} (D_6 * V_6)^{b3} (D_7 * V_7)^{b4}$
where		
S	=	total number of accidents from 1985 to 1988
V_i	=	exposure parameter, which is the various traffic volumes
D_i	=	risk parameter, which is the delay to drivers
a_i, b_i	=	coefficients to be modelled by GLIM

4.8.4 Delay-based Accident Prediction Model 4 (Model 214)

The fourth delay-based model is a variation of the third model. The power function, instead of being applied to both the risk and exposure parameters, was applied to the risk parameters only. The resulting model structure is as follows (Model 214):

$$S = e^{a0} (V_3 * V_4 * V_6 * V_7)^{c1} \; (V_2^{b1} * D_4^{b2} * D_6^{b3} * D_7^{b4})$$

where

S	=	total number of accidents from 1985 to 1988
V_i	=	exposure parameter, which is the various traffic volumes
D_i	=	risk parameter, which is the delay to drivers
$_{ai,bi,ci}$	=	coefficients to be modelled by GLIM

4.8.5 Delay-based Accident Prediction Model 5 (Model 215)

The fifth delay-based model is a variation of the previous model. In this model, the power function is applied to the exposure parameters and not to the risk parameters.

$$S = e^{a0} (V_2 * D_4 * D_6 * D_7)^{b1} \; (V_3^{c1} * V_4^{c2} * V_6^{c3} * V_7^{c4})$$

where

S	=	t [illegible]ber of accidents from 1985 to 1988
V_i	=	exposure parameter, which is the various traffic volumes
D_i	=	risk parameter, which is the delay to drivers
$_{ai,bi,ci}$	=	coefficients to be modelled by GLIM

4.8.6 Delay-based Accident Prediction Model 6 (Model 216)

The sixth delay-based model has a power function for all the parameters. The structure is shown below as well as in Figure 4.5 (Model 216):

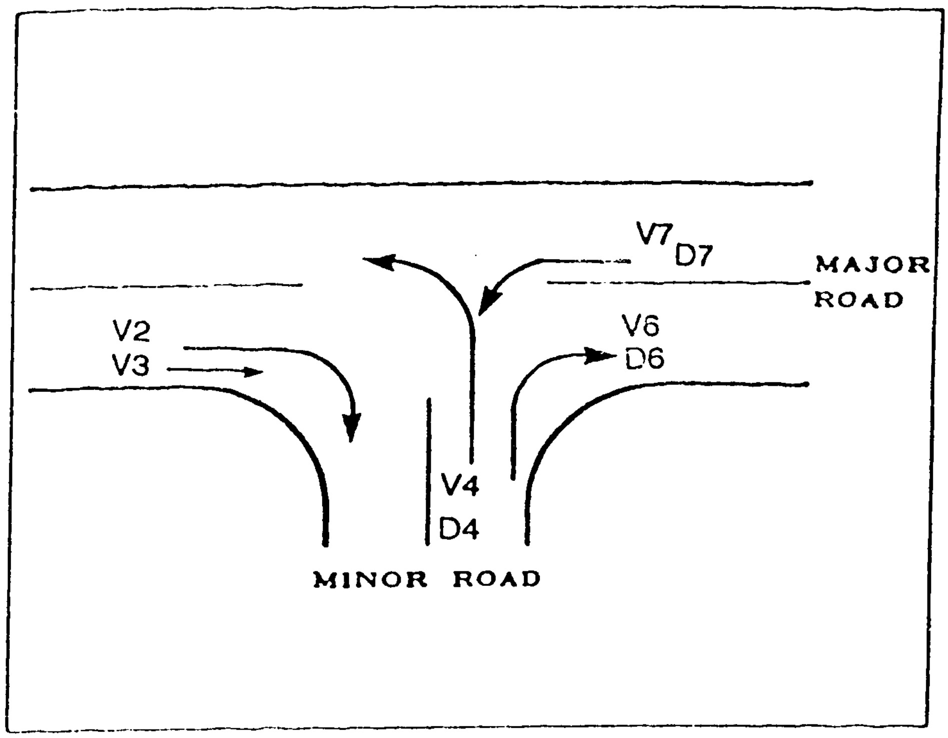


Figure 4.5 Parameters in Delay-based Model (Model 226)

$$S = e^{a0} (V_2^{b1} * V_3^{c1}) (D_4^{b2} * V_4^{c2}) (D_6^{b3} * V_6^{c3}) (D_7^{b4} * V_7^{c4})$$

where S = total number of accidents from 1985 to 1988

V_i = exposure parameter, which is the various traffic volumes

D_i = risk parameter, which is the delay to drivers

b_i, c_i = coefficients to be modelled by GLIM

$$S = e^{a0} (V_2^{b1} * V_3^{c1}) (D_4^{b2} * V_4^{c2}) (D_6^{b3} * V_6^{c3}) (D_7^{b4} * V_7^{c4})$$

where

S = total number of accidents from 1985 to 1988

V_i = exposure parameter, which is the various traffic volumes

D_i = risk parameter, which is the delay to drivers

b_i, c_i = coefficients to be modelled by GLIM

4.9 Aggregate Accident Prediction Models With Simulated-delay-based Parameters (Models 221 to 226)

The model structures of the simulated-delay-based models are identical to that of the delay-based models. The simulated-delay-based models were developed using simulated delay data instead of field measured delay.

4.10 Summary

At this stage, all the accident prediction models have been developed. The models developed cover a wide combination of approaches with different safety measures, degree of aggregation and level of complexity.

5. COMPARISON OF MODEL RESULTS AND PARAMETERS OF BEST MODELS

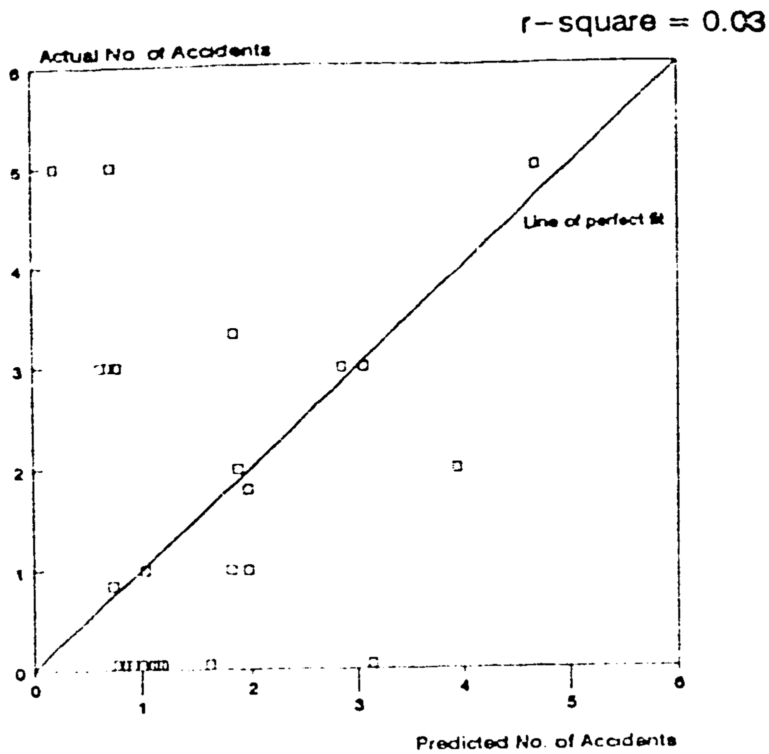
This chapter discusses and compares the results of accident trend prediction of the various models developed in the study by using goodness-of-fit plots. The best models and their model parameters are compared to assess the adequacy of the traditional intersection safety ranking methods.

5.1 Rear-end and Right-angle Accident Prediction Models (Models 512 and 612)

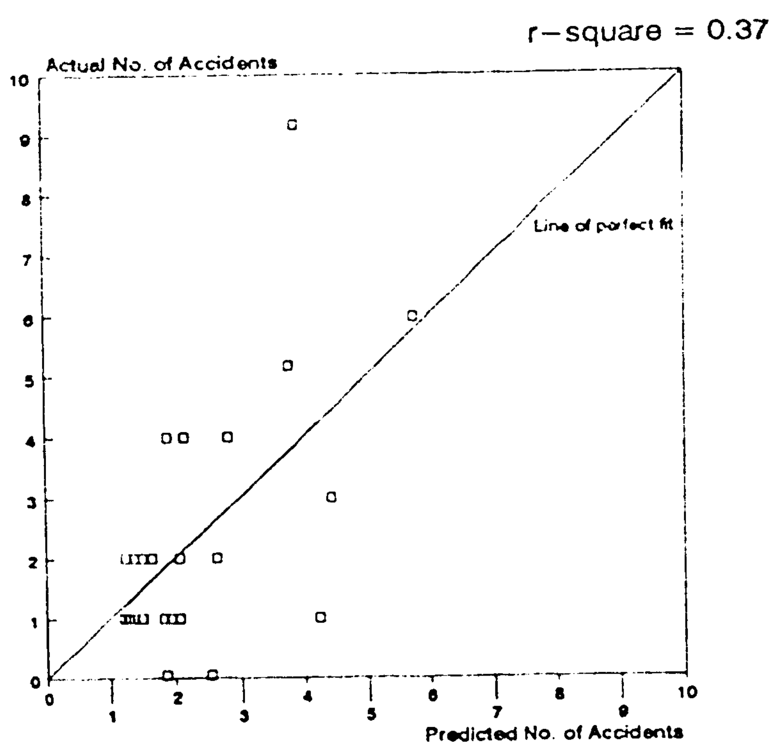
Figures 5.1a and 5.1b illustrate the goodness-of-fit of the rear-end (Model 512) and right-angle (Model 612) accident prediction models. In the goodness-of-fit plot, the y-axis represents the actual number of accidents at the intersection, which is the average number of accidents that occurred during weekday daytime between 1985 and 1988. The x-axis represents the number of accidents predicted by the accident prediction model.

There are 26 data points in each goodness-of-fit plot. Each data point represents the result of accident prediction at a particular intersection. The position of the data point indicates the success of the model in predicting the number of accidents. The straight diagonal line in each graph represents the line of perfect fit. A data point lying on the line of perfect fit means that the predicted number of accidents is the same as the actual number of accidents that occurred. A data point located below the diagonal line indicates that the number of accidents predicted by the model is higher than the actual number of accidents observed.

Accordingly, a model with the 26 data points located closely adjacent to the best-fit line would have better accident trend prediction results than one with data



a. Rear-end Collision Model (Model 512)



b. Right-angle Collision Model (Model 612)

Figures 5.1 a-b Goodness-of-fit Plots for Rear-end and Right-angle Accidents points scattered far away from the best-fit line. Models with higher r-square values with their goodness-of-fit plots were identified as better accident prediction models.

Both Figures 5.1a and 5.1b exhibit a significant degree of scattering of data points. The scattering is more disperse for the rear-end accident prediction model, Model 512. The resulting r-square values for the rear-end and the right-angle accident prediction models are 0.03 and 0.37, respectively. The results of accident prediction is considered poor, especially for the rear-end accident prediction model.

The poor predictability of the rear-end collision model may be due to the inability of the model to reflect the cause-and-effect relationship in rear-end collisions. Grouping accidents as rear-end collisions does not correlate accidents directly to the traffic movements involved in the accidents.

Comparatively, the right-angle accident model has better prediction results than the rear-end accident model. The model groups together parameters that are directly related to the pre-crash maneuvers of the vehicles involved in the collision.

The degree of disaggregation apparently has significant effect on the predictability of the models. As the total number of accidents that occurred at an intersection have to be split between the rear-end and right-angle accident prediction models, the resulting accident frequency in each model is lower. This is most obvious in the rear-end collision case, where there were no rear-end collisions from 1985 to 1988 at 9 of the 26 intersections.

Tables 5.1a and 5.1b summarize the GLIM program output results for the rear-end and right-angle accident prediction models. The results indicate that the parameter estimates for both models have high standard errors. For a negative binomial accident data distribution, a parameter estimate is considered insignificant unless it is three times greater than the standard error. Many parameter estimates in both the rear-end and right-angle accident predication models are insignificant.

Table 5.1a Parameter Estimates for Rear–end Accident Prediction Model (Model 512)

Model 512		$S_{RE} = a + a_0 (V_2 * V_3) + a_1 (V_4 * D_4) + a_2 (V_6 * D_6) + a_3 (V_7 * D_7)$
Deviance	52.42	
R–square	0.03	

#	Coefficients in Model Structure	Parameters in GLIM Output	Parameter Estimates	Standard Errors
1	a	%gm	8.44E–01	0.499
2	a0	q1	5.870E–05	4.070E–05
3	a1	q2	1.830E–03	2.360E–03
4	a2	q3	–1.430E–03	3.410E–03
5	a3	q4	3.330E–03	E2.860E–03

Table 5.1b Parameter Estimates for Right–angle Accident Prediction Model (Model 612)

Model 612		$S_{RA} = a + a_0 (V_4 * D_4) + a_1 (V_7 * D_7) + a_2 (V_6 * D_6)$
Deviance	14.21	
R–square	0.37	

#	Coefficients in Model Structure	Parameters in GLIM Output	Parameter Estimates	Standard Errors
1	a	%gm	1.225E+00	3.620E–01
2	a0	q1	1.520E–03	1.420E–03
3	a1	q2	–2.270E–03	E2.360E–03
4	a2	q3	4.170E–03	E2.330E–03

5.2 Non-priority Movement Accident Prediction Models (Models 311 to 314)

Four models are developed for the four non-priority movements at T-intersections. Traffic volume and delay parameters are used to model the risk and exposure aspects of the accidents related to each of the non-priority movements. The resulting structures of the models provide a logical cause-and-effect relationship for each accident type.

The major problem of the approach is the low accident data density that resulted from this level of disaggregation. The resulting accident densities are very low because the total number of accidents is divided into 4 groups according to the relevant non-priority movements.

In the major road right-turn model (Model 311), a total of 15 accidents occurred at 9 intersections over 4 years. The model has a r-square value of 0.10. There were no major road right-turn accidents at 17 intersections. Figure 5.2a illustrates the goodness-of-fit of the prediction results of Model 311.

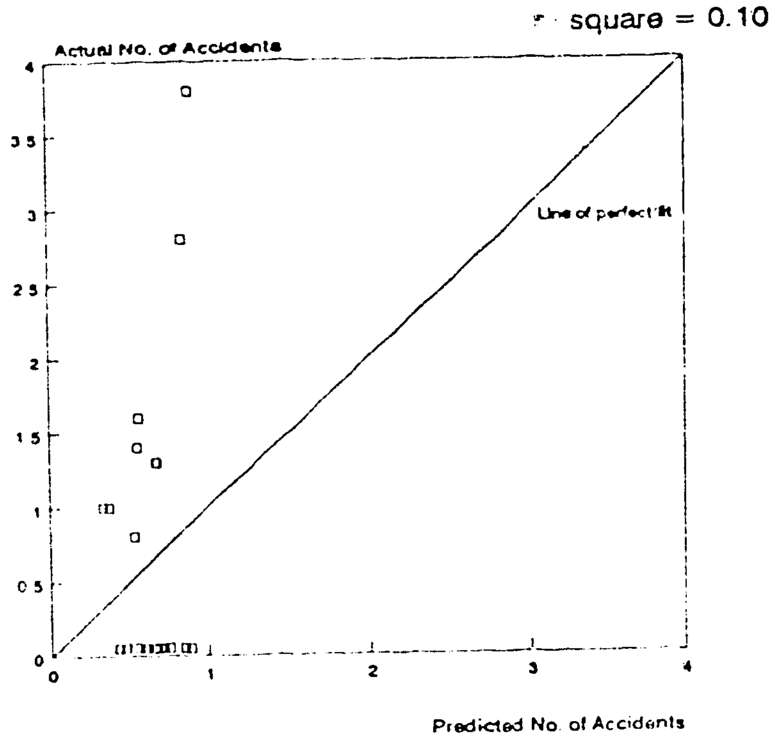
In the minor road left-turn model (Model 312), a total of 49 accidents occurred at 22 intersections over 4 years. The model has a r-square value of 0.40. Figure 5.2b illustrates the goodness-of-fit of the prediction results of Model 312

In the minor road right-turn model (Model 313), there were a total of 37 accidents at 17 intersections over 4 years. The model has a r-square value of 0.69. Figure 5.2c illustrates the goodness-of-fit of Model 313.

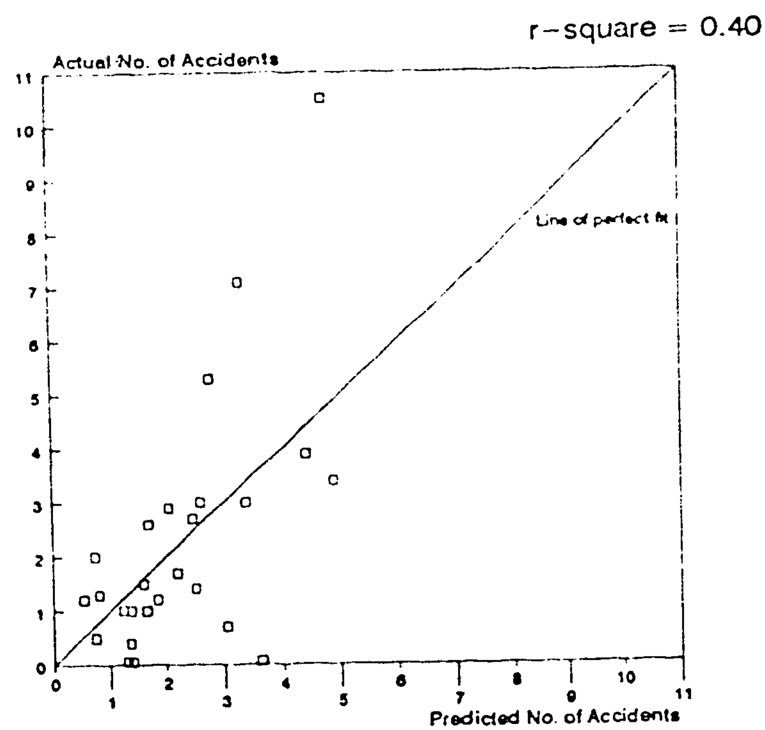
Finally, in the major road left-turn model (Model 314), there was a total of 41 accidents at 18 intersections over 4 years. The model has a r-square value of 0.10. Figure 5.2d illustrates the goodness-of-fit of Model 314.

Parameter estimates were also found to be insignificant in all four models, as indicated in Tables 5.2a to 5.2d.

Among the four non-priority accident prediction models, the models with the best accident trend prediction results are the minor left-turn and minor road

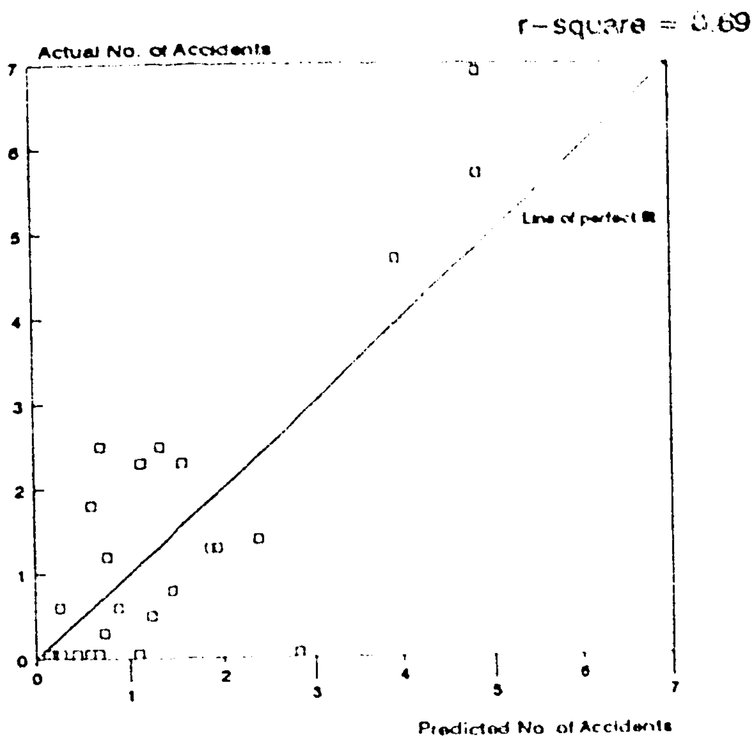


a. Major Road Right–turn Model (Model 311)

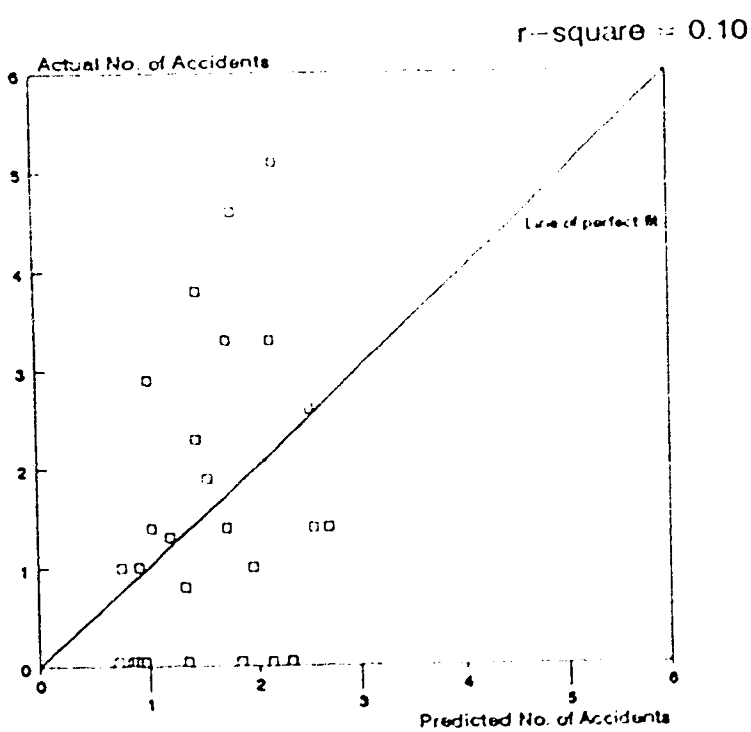


b. Minor Road Left–turn Model (Model 312)

Figures 5.2 a–b Goodness–of–fit Plots for Non–Priority Movement Models



c. Minor Road Right-turn Model (Model 313)



d. Major Road Left-turn Model (Model 314)

Figures 5.2 c-d Goodness-of-fit Plots for Non-Priority Movement Models

Table 5.2a Parameter Estimates for Major Road Right-turn Accident Prediction Model (Model 311)

Model 311		$S_3 = e^{a3} V_2^{b3} V_3^{c3}$
Deviance	21.65	
R-square	0.10	

#	Coefficients in Model Structure	Parameters in GLIM Output	Parameter Estimates	Standard Errors
1	log(a3)	%gm	-3.563E+00	4.157
2	b3	e1	-1.450E-01	3.170E-01
3	c3	r1	5.380E-01	6.710E-01

Table 5.2b Parameter Estimates for Minor Road Left-turn Accient Prediction Model (Model 312)

Model 312		$S_4 = e^{a4} V_4^{b4} D_4^{c4}$
Deviance	14.96	
R-square	0.40	

#	Coefficients in Model Structure	Parameters in GLIM Output	Parameter Estimates	Standard Errors
1	log(a4)	%gm	2.820e-01	4.180e-01
2	b4	e2	5.960e-01	1.820e-01
3	c4	r2	-4.020e-01	2.130e-01

Table 5.2c Parameter Estimates for Minor Road Right-turn Accident Prediction Model (Model 313)

Model 313		$S_6 = e^{a6} V_6^{b6} D_6^{c6}$		
Deviance	13.07			
R-square	0.69			

#	Coefficients in Model Structure	Parameters in GLIM Output	Parameter Estimates	Standard Errors
1	log(a6)	%gm	-3.684E+00	9.370E-01
2	b6	e3	1.019E+00	2.470E-01
3	c6	r3	2.420E-01	3.240E-01

Table 5.2d Parameter Estimates for Minor Road Left-turn Accient Prediction Model (Model 314)

Model 314		$S_7 = e^{a7} V_7^{b7} D_7^{c7}$		
Deviance	20.73			
R-square	0.10			

#	Coefficients in Model Structure	Parameters in GLIM Output	Parameter Estimates	Standard Errors
1	log(a7)	%gm	-1.161E+00	8.990E-01
2	b7	e4	2.870E-01	2.370E-01
3	c7	r4	3.890E-01	3.630E-01

right-turn accident prediction models, with r-square values of 0.40 and 0.69 respectively. The high r-square value for the right-turn accident model is mainly due to the presence of a few data points with high accident frequencies at the upper right hand corner of the figure. The actual data scattering would be extensive if these data points were removed from the model, and would result in much worse r-square values. The minor road left-turn model is therefore considered the best non-priority movement accident prediction model. The model has the highest accident frequencies and the least number of locations with no accident occurrences (empty cells) during the study period.

The predictability of major road non-priority movement accidents is poor, with r-square value of 0.10. The levels of priorities for the major road non-priority movement are more difficult to distinguish than the other movements, which may explain the poor predictability of the model. For example, a major road right-turn movement otherwise is normally not required to yield priority to any other movements. However, when a right-turn vehicle is required to change lanes to a right-turn lane prior to making a right turn maneuver, the right turn maneuver is then considered as a non-priority movement.

In general, non-priority movement accident prediction models provide poor accident trend prediction due to low accident frequencies and the presence of a large number of empty cells in the data base. An attempt to correlate the sum of the various predicted non-priority movement accidents with the actual measured number of accidents at the 26 study locations is unexpectedly successful with a r-square value of 0.63. There are two possible explanations for this situation. First, the over-prediction in one disaggregate model may be compensated by under-prediction. Secondly, the erroneous categorization in one model compensates missed categorization in the other three non-priority movement models.

5.3 Aggregate Accident Prediction Models With Volume-based Parameters

The models developed up to this point are relatively disaggregate models. In these disaggregate models, the non-safety parameter is either divided into rear-end and right-angle accidents, or into groups of accidents by non-priority movements. The advantage of disaggregate models is that it is easier to model the actual interactions in a sub-group of data that is more homogeneous. The categorization and segmentation of the data results in a more homogeneous data base. However, due to low accident frequencies, division of the database into smaller sub-groups has created significant problems in modelling statistical information. This resulted in insignificant model parameters and poor accident trend prediction.

An aggregate modelling approach was used for the remainder of the accident prediction models, using the total number of accidents which occurred during weekday daytime between 1985 and 1988 as the non-safety parameter in the model.

5.3.1 Sum-of-flow Accident Prediction Models (Models 111 and 112)

Two models (Models 111 and 112) were developed based on parameters comprised of the sum of various entering flows. Both models have poor accident predictability with r-square values of 0.12 and 0.21, as well as insignificant model parameter estimates, as indicated in Table 5.3a.

Figure 5.3a illustrates the goodness-of-fit of Model 111, which uses the sum of all entering flows as its only model parameter. It appears that the predicted number of accidents at the 26 study locations does not correlate well to the measured accident numbers, which indicates that the sum-of-flow approach is not appropriate for accident prediction.

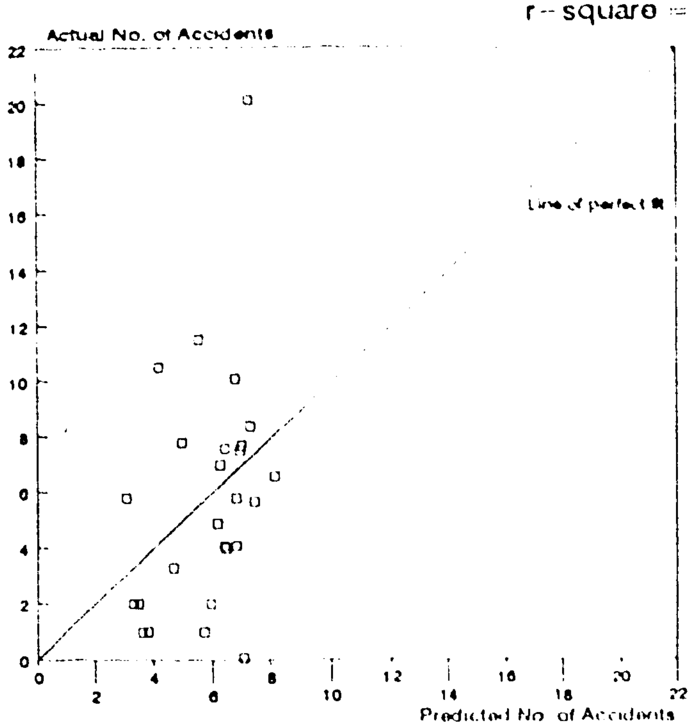
Table 5.3a Parameter Estimates for Sum–of–flow Accident Prediction Model (Model 111)

Model 111			$S = a_0 + a_1 [v_2+v_3+v_4+v_6+v_7+v_8]$	
Deviance	25.53			
R–square	0.12			
#	Coefficients in Model Structure	Parameters in GLIM Output	Parameter Estimates	Standard Errors
1	a0	%gm	1.736E+00	1.837E+00
2	a1	e	2.820E–03	1.400E–03

Table 5.3b Parameter Estimates for Product–of–flow Accient Prediction Model (Model 122)

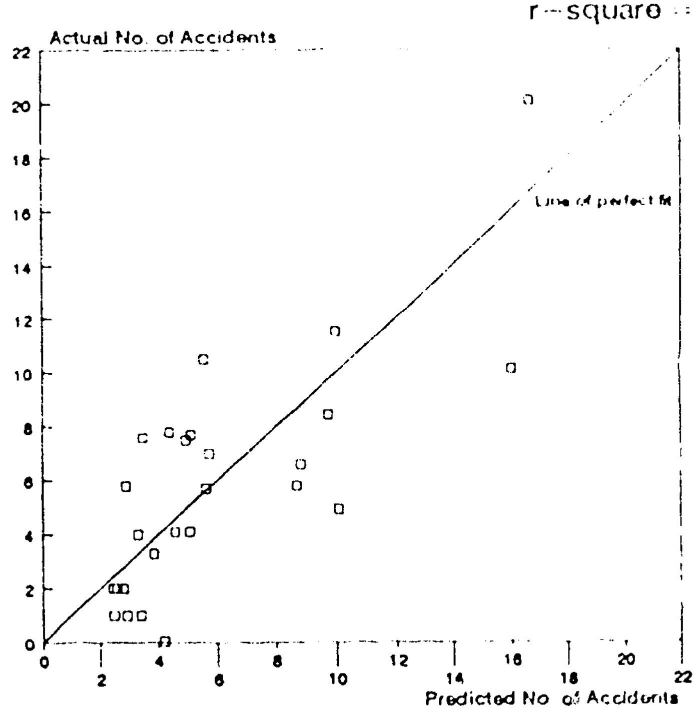
Model 122			$S = a_0 + a_1 [(V_2+V_3+V_7+V_8) * (V_4+V_6)]$	
Deviance	12.34			
R–square	0.59			
#	Coefficients in Model Structure	Parameters in GLIM Output	Parameter Estimates	Standard Errors
1	a0	%gm	2.189E+00	6.690E–01
2	a1	e	4.890E–05	1.310E–05

r-square = 0.12



a. Sum-of-flow Model (Model 111)

r-square = 0.59



b. Product-of-flow Model (Model 122)

Figures 5.3a-b Goodness-of-fit Plots for Volume-based Models

5.3.2 Product-of-flow Accident Prediction Models (Models 121 to 124)

Four models were developed based on parameters comprised of the products of flows entering the intersection (Models 121 to 124). The four models contained various levels of complexity in model structures. Models 121 and 122 have a linear model form, whereas Model 123 and 124 have a multiplicative exponential model form.

Models based on parameters comprised of the products of flows had much better prediction power than those based on the sum of entering flows. All four product-of-flow models exhibit good accident predictability, with r-square values of 0.70 for Model 124 and 0.59, 0.59 and 0.60 for the other three models.

Table 5.3b indicates that the parameter estimates in the models are insignificant with high standard errors, except for the model with the simplest model structure, Model 122. This model is the simple product-of-flow model which correlates the actual number of accidents to the product of the major and minor road traffic volumes. It appears that overly complicated model structures result in model parameters that are statistically insignificant. The goodness-of-fit graph for Model 122 is illustrated in Figure 5.3b.

The trend prediction power of the simple product-of-flow model was only slightly lower than that provided by the more complicated product-of-flow models. However, unlike the more complicated models, all its parameters are statistically significant. It appears that an added level of model complexity did not improve the accident predictability of the model.

5.4 Aggregate Accident Prediction Models With Delay-based Parameters (Models 211 to 216, and 221 to 226)

Twelve delay-based models were developed. The first six models (Models 211 to 216) were developed by using the delay data collected at the study locations. The remaining six models (Models 221 to 226) were developed by using the delay data derived through the use of a simulation program, KNOSIM. The structures of these 12 models are more complex than the models previously developed.

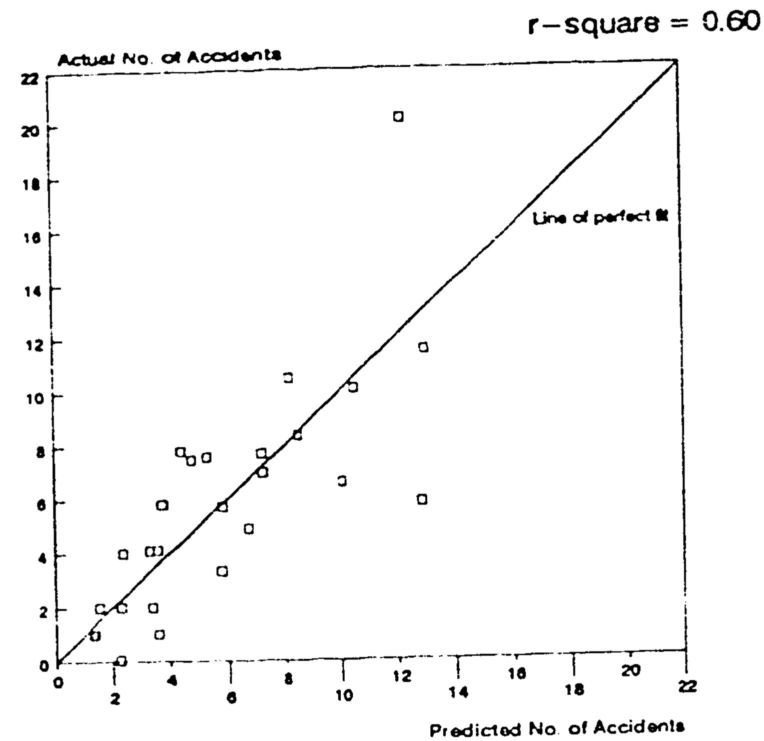
It appears that the simulated delay data produced better accident prediction models than the measured delay data. The average r-square value for 4 accident prediction models developed based on simulated delay data (Models 223 to 226) is 0.63. The average r-square value for 4 models developed based on measured delay data (Models 213 to 216), with the same model structures, is lower at 0.56.

The improved predictability of the model developed by using the simulated delay data may be due to the lower degree of variability within the simulated delay data. Figures 5.4a and 5.4b illustrate the goodness-of-fit of the two complex delay-based models, Models 216 and 226 (r-square values of 0.60 and 0.74). However, as indicated in Tables 5.4a and 5.4b, most parameters of this two complex delay-based models are statistically insignificant with high standard errors.

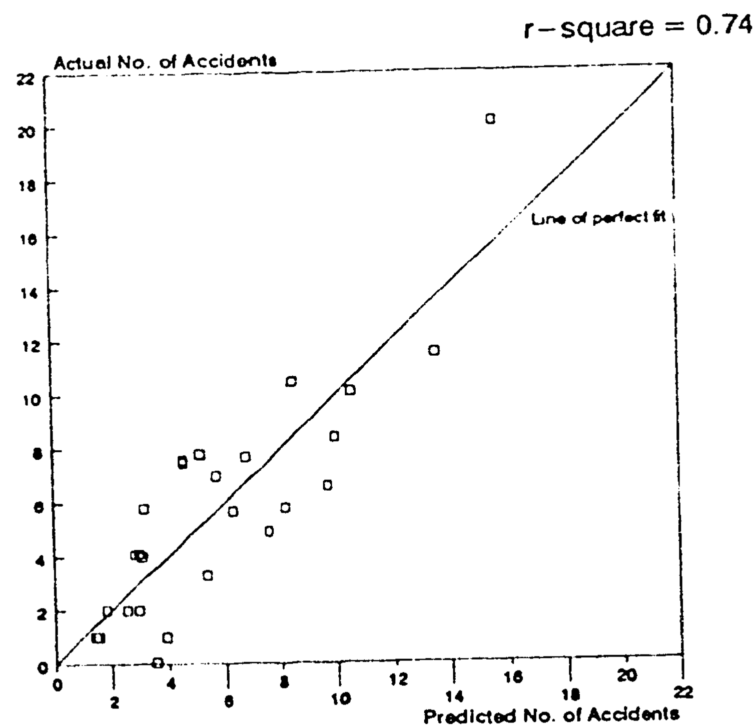
A comparison between the volume-based product-of-flow models and the delay-based models indicates that the two groups of models both produce relatively high levels of accident predictability. However, in general, it appears that models that include delay parameters are slightly better in accident prediction than those of the product-of-flow models which do not include delay parameters.

5.5 Simple Aggregate Accident Prediction Models

Of all the models developed through the statistical modelling process, only one model (Model 122) contains parameters that are all statistically significant.



a. Measured Delay Based Model (Model 216)



b. Simulated Delay Based Model (Model 226)

Figures 5.4 a-b Goodness-of-fit Plots for Delay-based Models

Table 5.4a Parameter Estimates for Delay-based (Measured Delay) Accident Prediction Model (Model 216)

Model 216		$S = e^{a0} (V_2^{b1} * V_3^{c1}) (D_4^{b2} * V_4^{c2}) (D_6^{b1} * V_6^{c3}) (D_7^{b4} * V_7^{c4})$
Deviance	8.46	
R-square	0.66	

#	Coefficients in Model Structure	Parameters in GLIM Output	Parameter Estimates	Standard Errors
1	log(a0)	%gm	-5.840E-01	2.341E+00
2	b1	r1	-8.120E-02	4.070E-01
3	b2	r2	-3.650E-02	1.920E-01
4	b3	r3	3.520E-01	3.080E-01
5	b4	r4	-5.590E-02	3.520E-01
6	c1	e1	3.330E-02	2.950E-01
7	c2	e2	-3.400E-02	1.840E-01
8	c3	e3	3.090E-01	3.210E-01
9	c4	e4	3.280E-01	2.690E-01

Table 5.4b Parameter Estimates for Delay-based (Simulated Delay) Accient Prediction Model (Model 226)

Model 226		$S = e^{a0} (V_2^{b1} * V_3^{c1}) (D_4^{b2} * V_4^{c2}) (D_6^{b3} * V_6^{c3}) (D_7^{b4} * V_7^{c4})$
Deviance	8.86	
R-square	0.74	

#	Coefficients in Model Structure	Parameters in GLIM Output	Parameter Estimates	Standard Errors
1	log(a0)	%gm	-3.120E+00	3.062E+00
2	b1	r1	4.250E-01	5.440E-01
3	b2	r2	1.520E-03	2.140E-01
4	b3	r3	-1.160E-01	5.420E-01
5	b4	r4	-1.350E-01	5.660E-01
6	c1	e1	-5.940E-03	2.970E-01
7	c2	e2	-2.050E-03	1.990E-01
8	c3	e3	3.970E-01	3.260E-01
9	c4	e4	3.490E-01	2.760E-01

Model 122 is a simple product-of-flow model. Parameters in the more complicated models were found to be not significant in similar comparisons, although these models have higher r-square values. Based on this observation, the more complicated models should not be considered more favourable than the simpler simple product-of-flow model, even though they gave marginally better accident prediction results. The low accident frequencies at intersections without traffic signals apparently limited the practical level of aggregation for modelling.

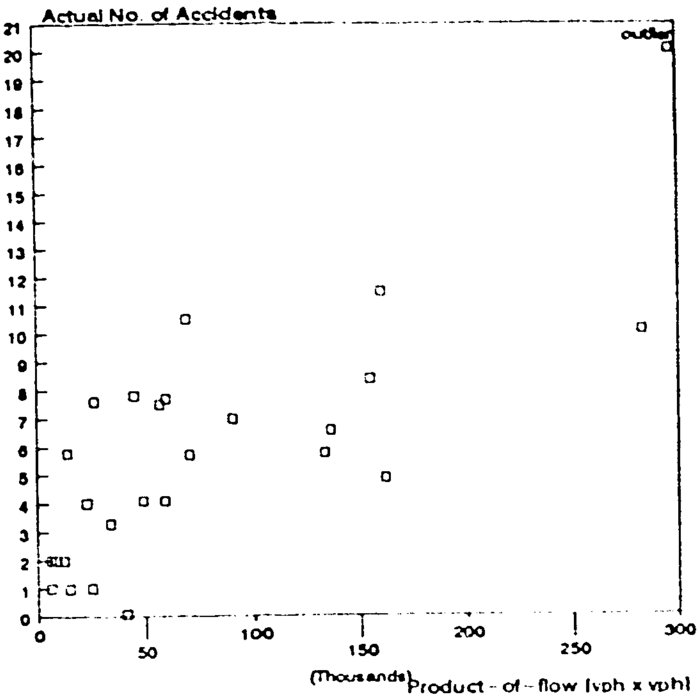
Two models with simple model structures were subsequently examined: the simple product-of-flow model and the total-delay model.

5.5.1 Simple Product-of-flow Models

Two simple product-of-flow models are considered. The two models differ in the way the conflicting flows are chosen. The first model uses the major road flow and the minor road flow as the conflicting flows. The second model uses the priority and non-priority flows as the conflicting flows. The priority flow is the sum of the main road through traffic flow at the intersection. The non-priority flow is the sum of all the turning traffic volumes at the intersection.

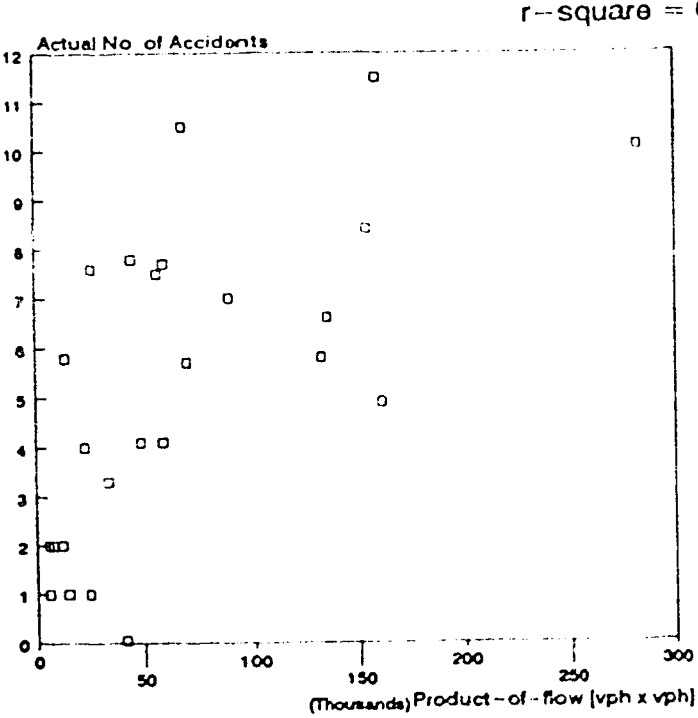
In both product-of-flow models, the major road flow or the priority flow can be considered as the risk presented to the minor road flow or the non-priority flow. The minor road flow or the non-priority flow, on the other hand, represents the exposure parameter in the accident prediction model. The product of the risk and exposure parameters is then related to the non-safety parameter of the intersection, which is the number of accidents.

Figure 5.5a is a plot of the product of major and minor road flows against the measured actual number of accidents. The r-square of the model is 0.59. A closer examination on the scatter plot indicated a potential outlier at the upper right hand corner of the graph. The data point corresponds to the 93 Avenue - 50 Street r-square = 0.59



a. with outlier

r-square = 0.41



b. without outlier

Figure 5.5 Correlation of Accident Frequency with Product-of-flows intersection, which has the highest accident frequency of the 26 intersections. This data point, being so remote from the rest of the data points, creates an artificial linearity effect. This increased the r-square value of the graph significantly. The r-square value of the model is reduced to 0.41 with the outlier removed from the data set (Figure 5.5b).

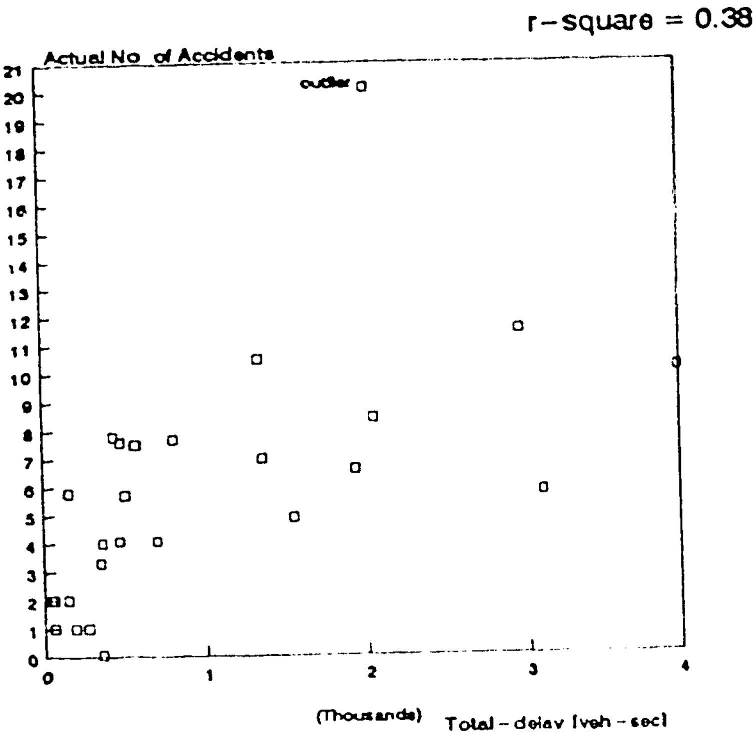
5.5.2 Simple Total-delay Model

Figure 5.6a illustrates the correlation between accidents and the total delay experienced by the non-priority traffic flows. Total delay is a product of the traffic volumes of the non-priority movements and the average delay occurred to the non-priority movements. The r-square of the total-delay model is 0.38, which can increase to 0.45 with the outlier at 93 Avenue - 50 Street removed from the database Figure 5.6b).

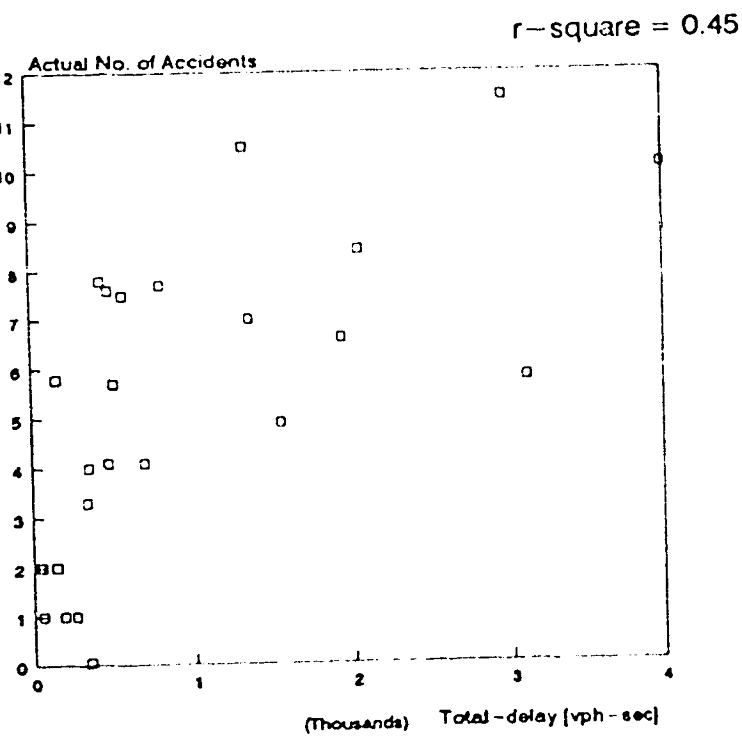
The structures of the two models are nearly identical. Both models use non-priority flow as the exposure parameter. The difference of the two models is in their choice of the risk parameter.

In the total-delay model, the average delay experienced by the non-priority stream traffic is used as the risk parameter, and the volume of the non-priority stream is used as the exposure parameter. The product of non-priority traffic flow and the average delay it experienced is the total delay experienced by the non-priority traffic streams.

In the product-of-flow model, the major road flow is used as the risk parameter, and the minor road flow is used as the exposure parameter. An experiment with the use of different risk and exposure parameters with different priority flow and non-priority flow combinations showed the same level of accident prediction results.



a. with outlier



b. without outlier

Figure 5.6 Correlation of Accident Frequency with Total-delay

The total-delay model, with a r-square value of 0.45, has better accident prediction results than the product-of-flow model, with a r-square value of 0.41.

5.5.3 Square Root of Product-of-flow and Total-delay Models

The square root of both the product-of-flow and the total-delay provided the best prediction results (Figures 5.7 and 5.8). The r-square of the product-of-flow model improved from 0.41 to 0.48 when the square root of the product-of-flow was used instead of product-of-flow. When the square root of total-delay is used in the total delay model, the r-square value improved from 0.45 to 0.54. Therefore, while the total-delay model provides the best accident trend prediction, the model can be further improved by using the square root of delay as the risk measure.

5.6 Ranking Comparison

Traditionally, traffic engineers rank intersections, from the safety point of view, by measures such as accident frequency and accident rate. Accident rate is typically derived by dividing accident frequency by the number of entering vehicles.

The models developed in this study can also be used to rank intersections by their "relative degree of safety." In Figure 5.9, the goodness-of-fit plot of Model 226 is used as an example to show that an intersection can be considered "more dangerous than average", if the actual number of accidents occurring at the intersection is significantly higher than the predicted number of accidents. That is, the accident frequency at that intersection is higher than what should be expected. Ranking approaches using the accident frequency and accident rate ranking criteria are used in selecting "dangerous accident locations".

An 95% significance envelop is placed in the goodness-of-fit graphs. If the data point of a particular intersection is located above the significance envelop, the

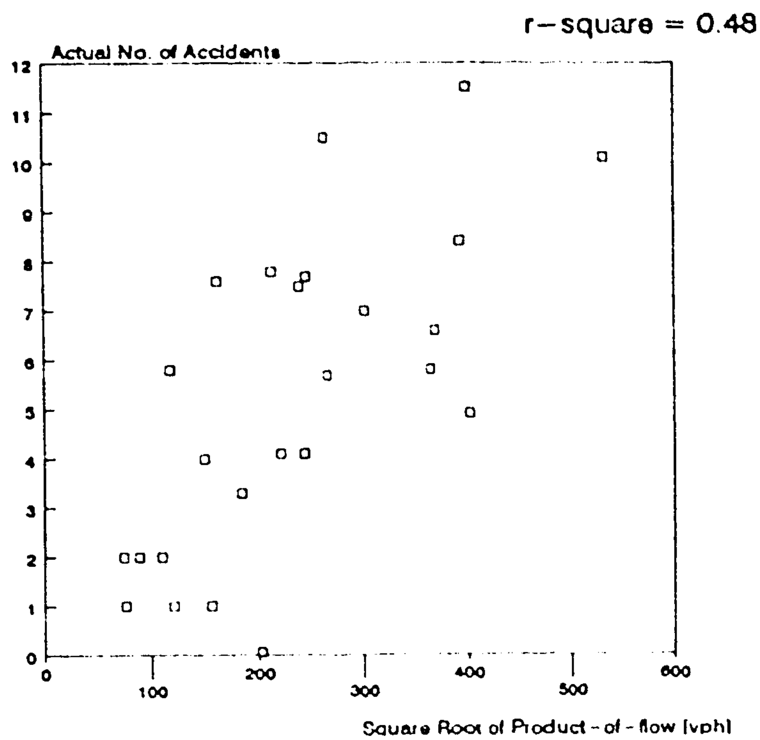


Figure 5.7 Correlation of Accident Frequency with Square Root of Product--of--flow

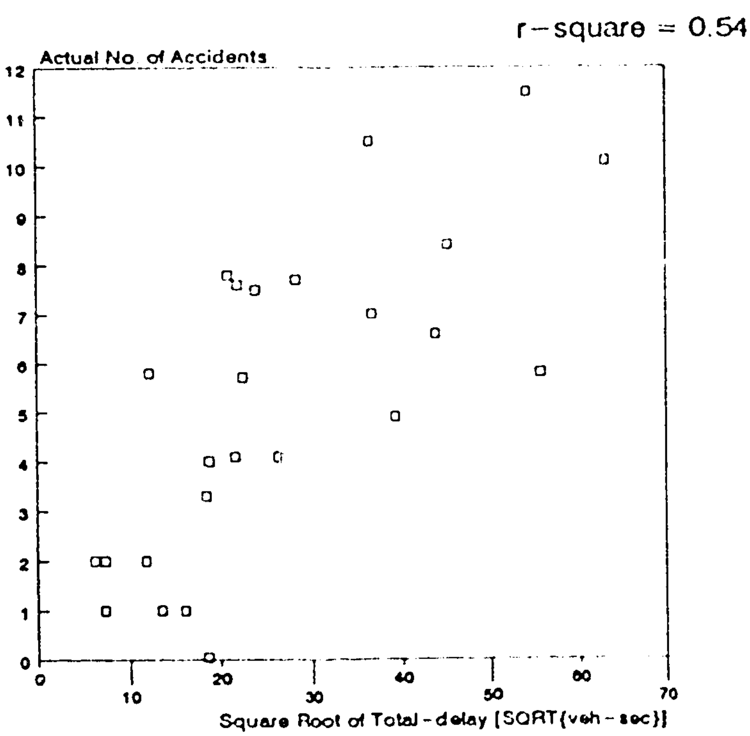


Figure 5.8 Correlation of Accident Frequency with Total-delay

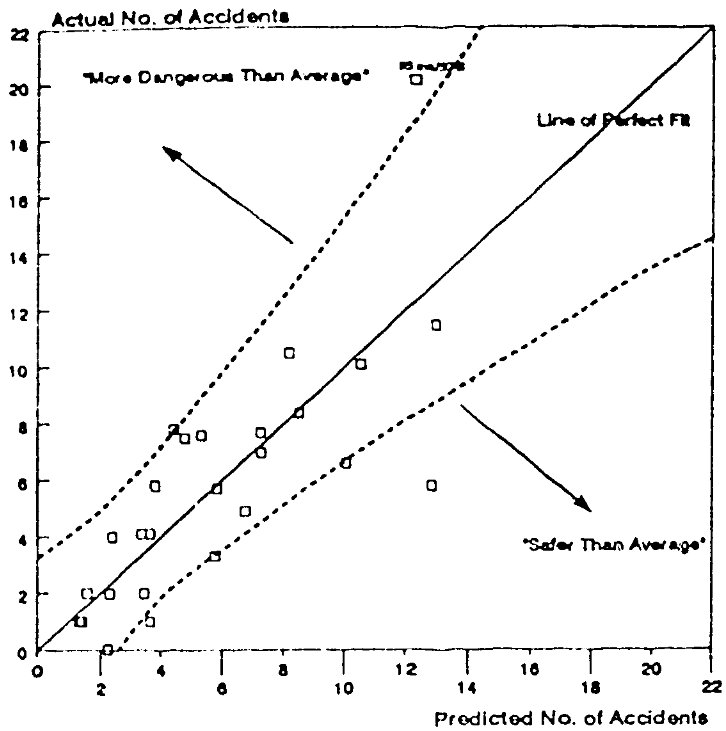


Figure 5.9 Approach in Determining "More Dangerous Than Average" Locations (Model 216)

Table 5.5 Results of Intersection Ranking

Site	Traditional Ranking Approaches		Ranking Based on Model Results (see Figure 5.9 above for example)			
	By Accident Frequency	By Accident Rate	Simple Total Delay Model	Simple Product of Flow Model	Complex Delay Based Model	Complex Product of Flow Model
93 Ave / 50 St	1	3	1	1	1	1
128 Ave/66 St	6	5	2	*	2	*
127 Ave/90 St	3	2	2	2	*	*
118 Ave/130 St	8	7	2	2	*	*
116 Ave/142 St	2	4	*	*	*	*
127 Ave/78 St	13	1	*	*	*	*

* Average safety conditions

actual accident frequency at that intersection would be significantly higher than predicted. The intersection could be considered "more dangerous than average"

The results of the various ranking approaches are shown in Table 5.5. The comparison indicates that ranking orders developed by the use of traditional non-safety parameters such as accident frequency and accident rate, do not agree well with that identified by the complex delay-based models. This is expected as the three approaches measure different aspects of the intersection safety situation.

5.7 Comparison of Parameters in the Best Models

It is proposed in the previous section that an intersection is considered "more dangerous than expected" if it has worse accident records than the other intersections with similar intersection and traffic conditions. In the accident rate ranking approach, the effect of exposure is not considered. An intersection may rank high on the list due to high exposure with a high accident rate. The actual situation at the intersection may be safer than average if the accident rate is lower than that of other intersections with similar geometry and traffic exposure. Similarly, for the accident frequency approach, a dangerous intersection may be low on the ranking list due to insignificant exposure. The fact that an intersection does not show up as a high accident frequency location because of low exposure does not mean that the intersection is safe.

To gain understanding on why the traditional ranking approaches provide different and contradicting ranking results, it is necessary to examine the parameters in the best models and compare them to accident frequency and accident rate; the two parameters most often used in traditional ranking approaches. The following model parameters are compared:

- accident frequency
- accident rates

- non-priority flow
- priority flow
- average delay to non-priority flow
- combination of the above parameters

However, it should be noted that many agencies use annual accident frequency as the ranking parameter. Significant variability is expected when a short analysis period such as one year is used.

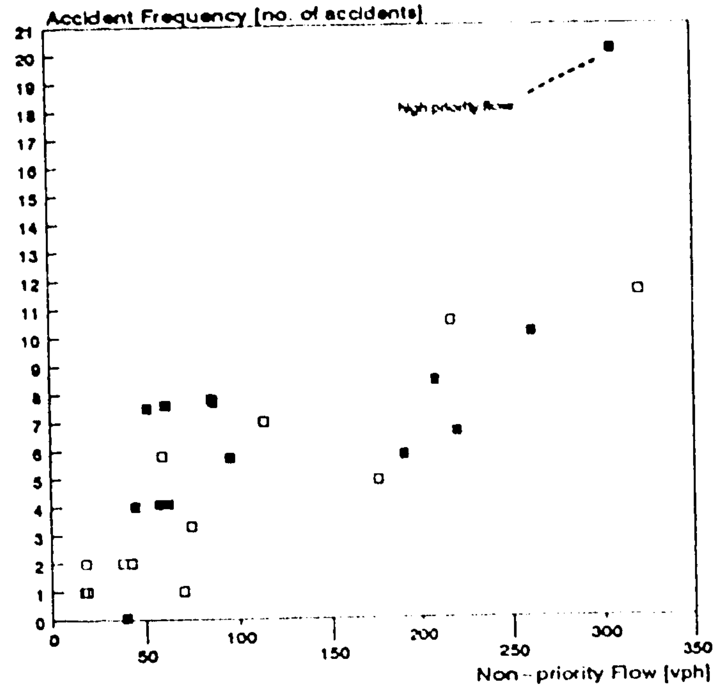
5.7.1 Examination of Accident Frequency

Accident frequency is a measure of non-safety. It indicates the number of accidents to be expected at a location during a given period of time. Figures 5.10a and 5.10b correlate the accident frequencies to the priority and non-priority traffic flow, respectively, at the 26 study locations. It appears that accident frequency (non-safety) is more directly affected by non-priority flow (exposure) and is relatively insensitive to the priority flow (risk).

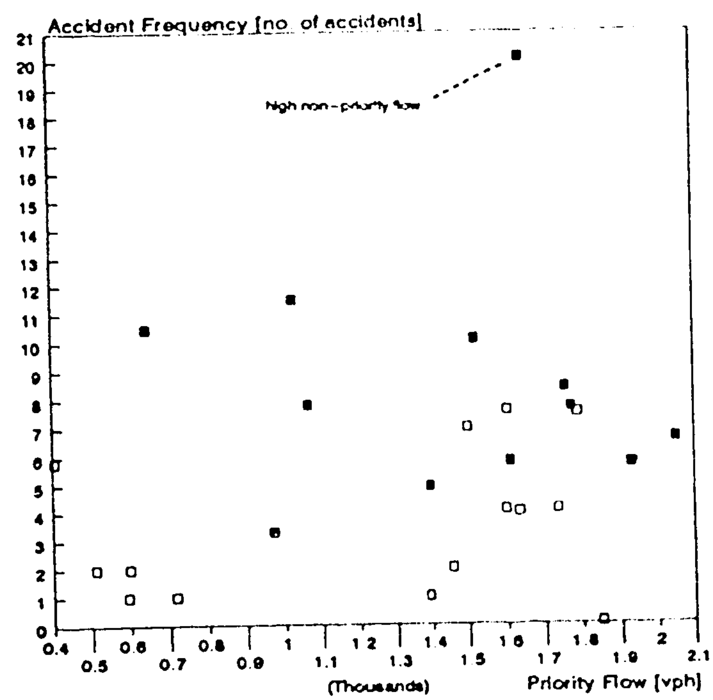
However, when the risk and exposure parameters are considered simultaneously, the correlation is strong between accident frequency and the risk and exposure measure. In Figures 5.11a and 5.11b, accident frequency is compared to the total delay as well as the product of priority and non-priority flow. The figures indicate that if the risk and exposure parameters are considered as a whole, it will provide a strong correlation to accident frequency.

5.7.2 Examination of Accident Rate

Accident rate, on the other hand, is a measure of risk. Provided that an appropriate measure of exposure is used, accident rate indicates how dangerous it is

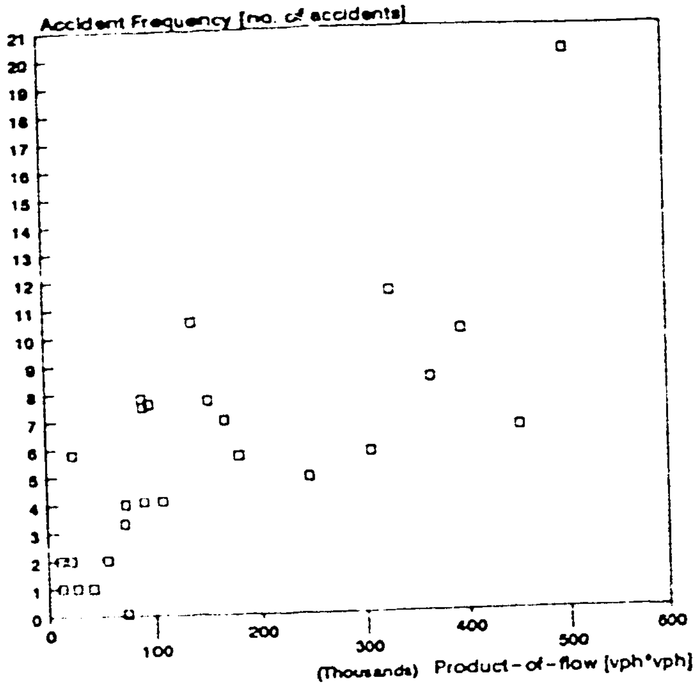


a. Accident Frequency and Non-priority Flow Plot

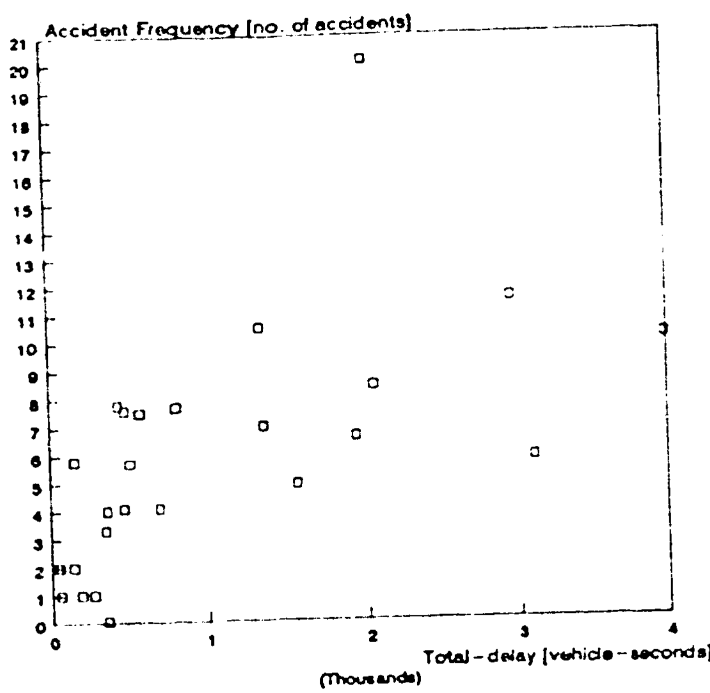


b. Accident Frequency and Priority Flow Plot

Figure 5.10 Correlation of Accident Frequencies with Priority and Non-priority Flows



a. Accident Frequency and Product-of-flow Plot



b. Accident Frequency and Total-delay Plot

Figure 5.11 Correlation of Accident Frequencies with Product-of-flow and Total-delay for the traffic going through an intersection. Accident rate can be defined by the following relationship:

ACCIDENT RATE = ACCIDENT FREQUENCY / EXPOSURE

Traditionally, the total traffic volume entering the intersection is used as the exposure measure for the accident rate parameter. As indicated in the previous section, accident frequency is not sensitive to the priority flow but is strongly correlated to the non-priority flow. The use of total entering flow is therefore questionable especially for locations with high priority flows and low non-priority flows. At those locations, the sum of entering flow is strongly correlated to the priority flow, which is shown to have poor correlation with accident frequency.

Figure 5.12a indicates that accident rate is highly correlated to the non-priority flow. Accident rate also tends to decrease with an increase in priority flow, as indicated in Figure 5.12b. The relationship, however, is not as strong as that of the non-priority flow. Figure 5.12c illustrates that accident rates increase with an increase in risk and exposure.

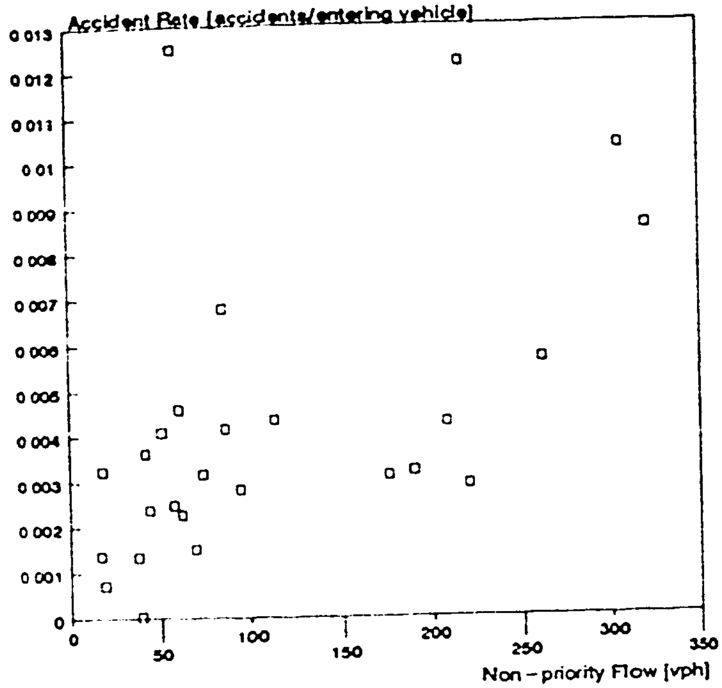
Nevertheless, the use of total-entering-flow as the exposure parameter does not distinguish the risk involved for different traffic movements at an intersection. It is intuitive that risks involved in traveling through a priority-controlled intersection along a major road is less than that of merging from a side street into the major road.

5.7.3 Examination of Modified Accident Rate Parameters

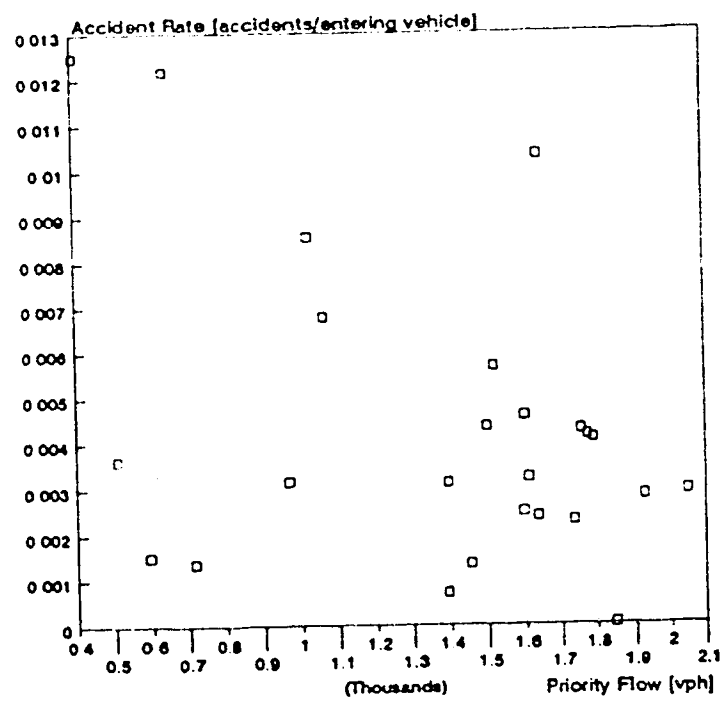
Two accident rate parameters were examined: the priority flow accident rate and the non-priority flow accident rate, as shown in the two equations below.

$$\text{PRIORITY FLOW ACCIDENT RATE} = \frac{\text{ACCIDENTS INVOLVING PRIORITY FLOW}}{\text{PRIORITY FLOW}}$$

$$\text{NON-PRIORITY FLOW ACCIDENT RATE} = \frac{\text{ACCIDENTS INVOLVING NON-PRIORITY FLOW}}{\text{NON-PRIORITY FLOW}}$$



a. Accident Rates and Non-priority Flow Plot



b. Accident Rates and Priority Flow Plot

Figures 5.12 a–b   Correlation of Accident Rates with Priority and Non–priority Flows

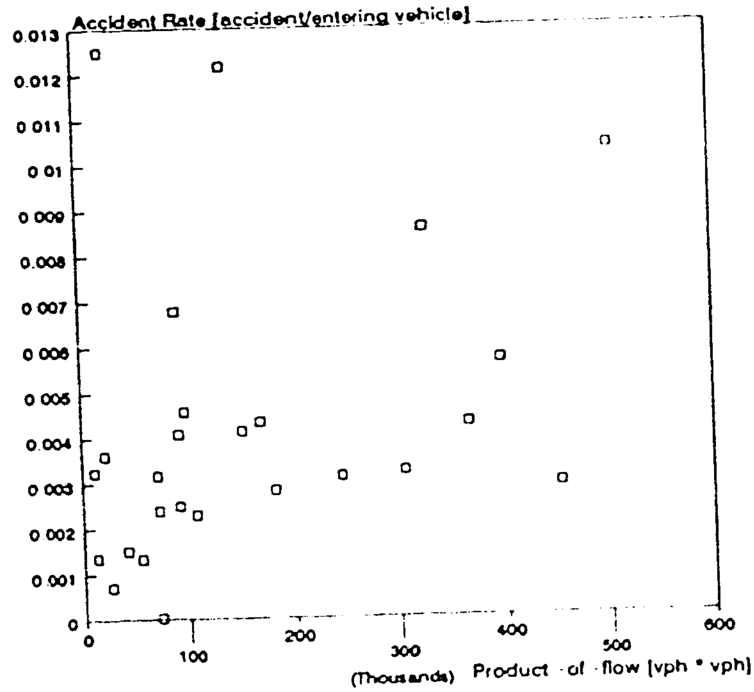


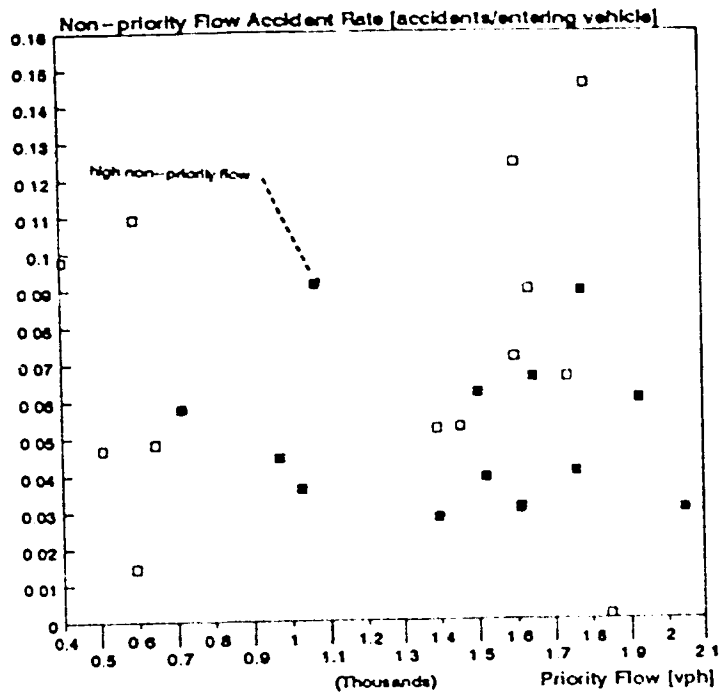
Figure 5.12 c Correlation of Accident Rates with Product of Priority and Non-priority Flows

One problem of using separate accident rates for priority and non-priority flows is that it is difficult to categorize accidents under the two types of flows. Figures 5.13a and 5.13b illustrate the effects of increase in priority flow on risk to priority flow traffic and non-priority flow traffic. In Figure 5.13a, with higher non-priority flows, the risk to priority flow decreases slightly as the priority flow increases. However, if the non-priority flow is low, the risk to priority flow remains relatively unchanged over different priority flow ranges. In Figure 5.13b, it appears that the risk to the low range non-priority flow increases slightly as the priority flow increases. The risk to higher range non-priority flows, on the other hand, appears to be relatively similar with different priority flows.

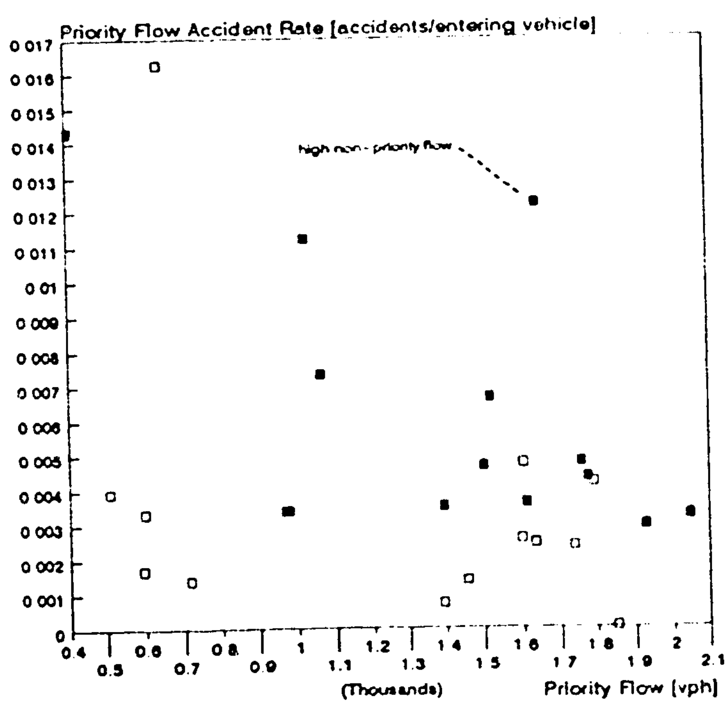
Figures 5.14a and 5.14b illustrate the effect of an increase in non-priority flow on the risk to both the priority and non-priority flows. Figure 5.14a shows that non-priority flow increase will result in an increase in risk to the priority flow. Figure 5.14b shows that increase in non-priority flow will result in a drop in risk to the non-priority flow.

5.7.4 Summary of Review on Accident Parameters

The comparison of various accident parameters indicates that accident and risk parameters, when being examined individually, do not give clear and consistent trend indications in most cases. The lack of a consistent trend indication is an evidence that other underlying factors may be present. Most accident parameters are robust measures that can be affected by a number of factors. A single parameter, when used on its own, is therefore not sufficient to explain the safety situation at an intersection. Extra caution should be applied in using a single accident parameter to determine the level of safety of an intersection.

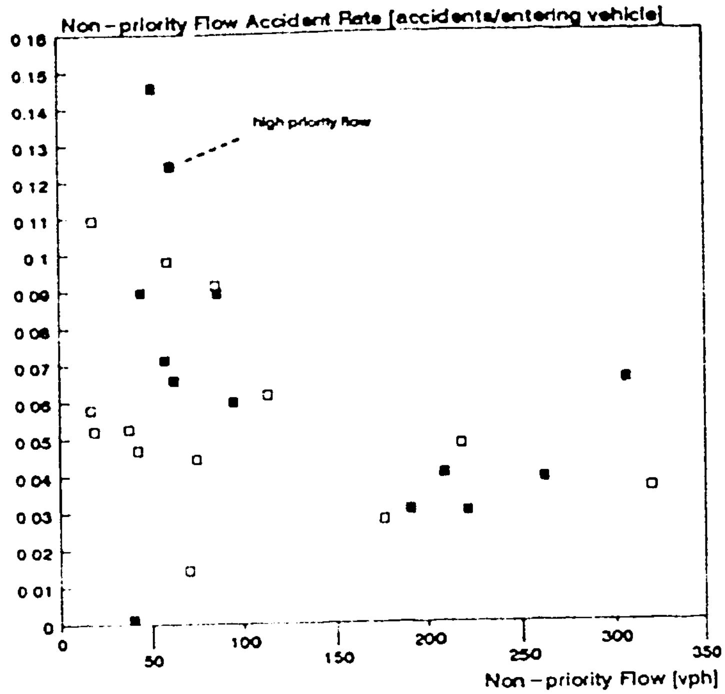


a. Non-priority Flow Accident Rate and Priority Flow Plot

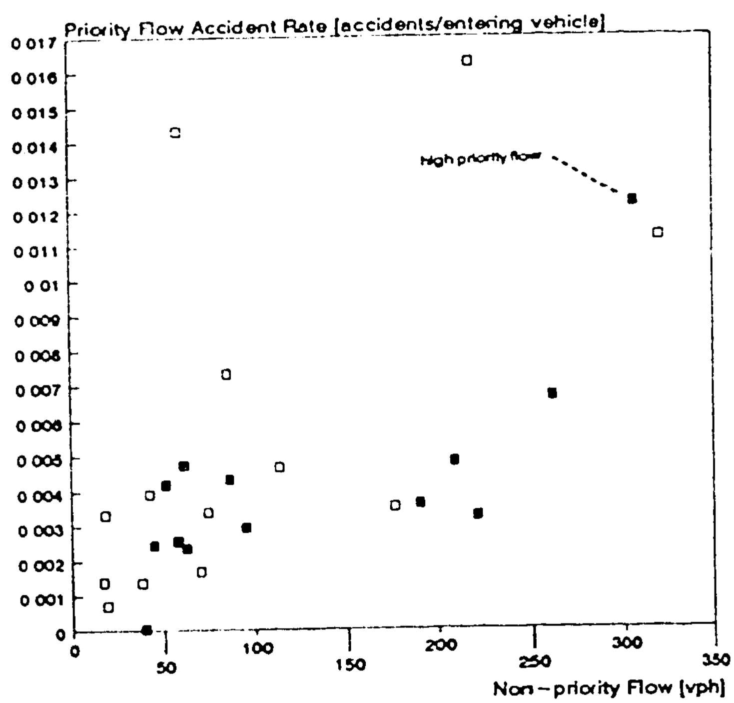


b. Priority Flow Accident Rate and Priority Flow Plot

Figure 5.13 Correlation of Priority Flow with Priority and Non-priority Accident Rates



a. Non-priority Flow Accident Rate and Non-priority Flow Plot



b. Priority Flow Accident Rate and Non-priority Flow Plot

Figure 5.14 Correlation of Non-priority Flow with Priority and Non-priority Accident Rates

6. SUMMARY OF FINDINGS AND CONCLUSIONS

Accident occurrences are rare events with a great degree of randomness. Numerous factors contribute to the occurrences of accidents. However, human factors are generally considered to be the major contributing elements.

An accident prediction model with strategically selected model parameters can contribute significantly towards the understanding of the risk presented to the drivers at an intersection. Past researchers have developed several accident prediction models. Most of these models employed some combinations or products of traffic volumes as main model parameters. Recognizing the role of human behaviour in accidents, it was hypothesized that accident prediction might be better accomplished if the parameters in the accident prediction model could reflect human behaviour better.

This research attempted to use delay experienced by motorists as a parameter in accident prediction models to reflect the effect of driver behaviour. The main objective of this research was therefore to compare the use of delay to the use of traffic volume in predicting traffic accidents. The secondary objective of the research was to examine the traditional methods of intersection safety ranking and compare that to the accident prediction modelling results.

This chapter outlines the findings, main conclusions and the practical implications of this research. Findings of the research are grouped into three main categories:

1. Accident Trends
2. Modelling Approaches
3. Accident Parameters

6.1 Findings in Accidents

6.1.1 Findings in Literature Review

1. Driver factor is the major contributing factor to accidents. It was involved in more than 94% of all roadway traffic accidents in a U.K. study.

2. The within site distribution of accidents at a particular location follows the Poisson process. The between site variation may be described by a gamma distribution. The resulting sampling distribution over all sites is a negative binomial distribution.

6.1.2 Findings in General Accident Patterns for the 429 Intersections Without Traffic Signals in Edmonton

1. There was a significant degree of variability in the occurrences of accidents.

2. Accident records in the City of Edmonton Accident database appeared to be generally reliable and valid in location, time, date and number of vehicles involved in accidents. However, information on driving lanes and road surface conditions might be of questionable accuracy.

3. Other than the average accident frequencies, accidents at 3-legged and 4-legged intersections were generally very similar in nature in yearly, monthly, daily and hourly trends. The trends were also similar in the type of pre-crash maneuvers and human actions at time of collision, and severity.

4. Accident occurrences generally peaked on Fridays and at peak hours during the remaining days of the week. The p.m. peak periods had the highest frequency of accidents.

5. Characteristics of accidents could be distinctively different, depending on the day of the week and the time of day they occurred. Daytime accidents mostly occurred (84% as compared to the average of 81%) on the weekdays, whereas a significant amount of nighttime accidents (42% as compared to the average of 19%) occurred on the weekend. On the weekends, 26% of the accidents occurred at night.

6.1.3 Findings in Specific Accident Statistics for 429 Intersections Without Traffic Signals

1. Accident frequencies were low for intersections without traffic signals with an average of 8.8 accidents/4 years from 1985 to 1988 for 429 intersections in Edmonton.

2. Of the 429 intersections, the 4-legged intersections with 10.2 accidents/4 years had almost twice as many accidents as the 3-legged intersections with 5.7 accidents/4 years.

3. The number of accidents in a year had remained relatively stable with less than 15% overall annual fluctuations from 1985 to 1988.

4. Within a year, accidents peaked in November and February, and were also high in June.

5. Analysis on weekday daytime accidents indicated that, among the 429 intersections, on the average there were 3.3 accidents/day during 5 winter months and 2.3 accidents/day during 5 summer months.

6. Of the 17 days with the highest accident frequencies between 1985 and 1988, the top 9 days were days with "severe winter driving conditions." These high accident occurrence days coincided with unusually poor travel conditions such as first day of snow, very slippery roadway surface and very heavy snowfall.

7. The most frequent types of pre-crash maneuvers were:

   - left turn
   - slowing/stopping
   - improper lane change
   - ran off road
   - improper turns
   - failed to yield to pedestrians

8. The most frequent types of human action involved were:

   - stop sign violation
   - lane changes
   - skidding/swerving
   - pedestrian crossing with right-of-way

6.2 Findings in Accident Modelling

6.2.1 Findings Related to Model Framework and Scope

1. Based on the risk and exposure concept, non-safety at an intersection would be considered to be related to the product of risk presented to the motorists and the number of motorists (exposure) that are exposed to such a risk.

2. Accident patterns at 3-legged and 4-legged intersections were highly comparable. T-intersections were therefore chosen for the study to maintain data collection and interpretation efforts at a manageable level.

3. For modelling purposes, sites of varying accident frequencies were selected to cover a wide range of intersection non-safety conditions. Random selection would have resulted in an over-representation of the low accident frequency locations.

4. Based on the travel condition modules identified from the accident and traffic patterns, modelling efforts were concentrated on the weekday off-peak analysis period to minimize variability in travel conditions. The analysis period encompassed from Monday and Friday between 6:00 a.m. and 9:00 p.m.

5. Data collection was carried out within a period from Monday to Thursday between 11:00 a.m. and 4:00 p.m.. With relatively steady traffic volumes and accident trends, as well as the long duration of the period, this analysis period was ideal for data collection.

6.2.2 Findings Related to Model Structures and Parameters

1. Historical accident data were considered superior to other parameters such as traffic conflict due to its well-established collection and documentation procedures as well as its immediate availability and accessibility for retrieval.

2. Excluding multiple vehicle accidents and single vehicle accidents, 13 major basic accident types were identified to be related to one or more non-priority movements at a T-intersection. The most frequent accident types were right-angle collisions, rear-end collisions and left-turn-cross-path accidents. Disaggregating the non-safety situations to a basic accident type level resulted in very low accident frequencies for all basic accident types at most locations.

3. Attempts to further aggregate the accidents by non-priority movements resulted in higher accident frequencies. However, difficulties were still experienced in relating some accident types to non-priority movements.

4. The delay data collected from the field displayed a high degree of variability.

5. With minor input modifications, the KNOSIMO traffic simulation program was found to generate reasonable results.

6. As accident data has a negative binomial distribution, typical classical regression models, which are based on a normality assumption, are not suitable for accident data analysis. The GLIM statistical analysis package can handle analysis for data set with negative binomial distribution and was therefore used in this research.

6.2.3 Findings Related To The Modelling Results

1. The disaggregate accident prediction model for rear-end collisions was found to have very poor predictability with an r-square value of 0.03. This was probably due to the inability of the model to reflect the actual cause-and-effect relationship in rear-end collisions.

2. The disaggregate models for right-angle accidents also had relatively poor predictability. However, its r-square value of 0.37 was much better than the 0.03 of the rear-end accident model. The better predictability was probably because it was much easier to model the accidents with the more obvious conflicting traffic movements.

3. Disaggregate models for each of the non-priority movements were generally poor in predictability, mainly because of low non-priority movement accident frequencies at several study locations. The model with the best predictability was the minor road left turn accidents model with an r-square value of 0.40.

4. The previous models had demonstrated that a highly disaggregated level of modelling was not suitable for predicting accidents at intersections without traffic signals, where accident frequencies were low. Aggregate accident prediction models using the total number of accidents which occurred at a location was preferable.

5. The aggregate sum-of-flow accident prediction model had poor predictability with an r-square value of 0.12.

6 The aggregate product-of-flow accident prediction model had much better prediction power with an r-square value of 0.59.

7 The use of non-priority flow and priority flow parameters instead of main road and minor road flows resulted in better accident trend prediction in the product-of-flow model.

8 Aggregate delay-based models were slightly better in predicting accidents than the aggregate product-of-flow models based on product of main road and minor road flows. However, if non-priority and priority flows were used in the product-of-flow model, both the total-delay and product-of-flow models had almost identical results in accident prediction. From an accident prediction modelling standpoint, using simulated delay or measured delay did not result in significant difference in accident prediction model capability.

9. For the simpler model, comparison of model parameters to their standard errors indicated that the parameters were significant. Parameters in the more complicated models were found insignificant in similar comparisons. Based on this observation, the more complicated models should not be considered more favourable than the simpler models, although they did give marginally better accident prediction results. The low accident frequencies presented a practical limit to the complexity of an accident prediction model, and therefore restricted the level of aggregation for modelling.

10. Two simple accident prediction models at aggregate level had been identified as sound and practical. They were the product-of-flow (based on priority and non-priority flows) and total-delay model. A trial by taking square roots of the

product-of-flow and total-delay parameters indicated that the performance of the two models could be enhanced with a power function model format. The actual power factor optimal for the two models should be determined based on a larger database size.

6.3 Findings in Accident Parameters

Further analysis on the two simpler models selected suggested that the accident and risk parameters examined did not provide clear trend indications in most cases. It was concluded that a single parameter was not capable to fully explain the non-safety situation as there might be numerous underlying factors in accident occurrences. Extra caution should be used in utilizing single accident parameters to determine the safety of intersections. Some of the major trends between accident parameters of the preferred models are:

1. Accident frequencies correlated much better with non-priority flow than major road flow.

2. Analysis using number of accidents over major road flow as a major road accident risk parameter indicated that:

    a. When non-priority flow was low, the effect of major road flow on major road accident risk could not be identified.

    b. When non-priority flow was high, major road risk decreased as major road flow increased.

    c. Regardless of the level of flow at the major road, major road accident risk increased with an increase in non-priority flow.

3. Analysis using the number of accidents over non-priority flow as a non-priority flow accident risk parameter indicated that:

    a. Regardless of the flow on the major road, non-priority accident risk decreased with an increase in non-priority flow.

    b. The effect of major road flow on non-priority accident risk could not be identified.

4. The above trends between accident parameters were either not always consistent or apparent, suggesting that other underlying factors might be present. As well, the non-linear characteristics of some of the plots indicated complex interactions between the parameters. This confirmed findings of Abou-Henaidy (1993) which indicated changes in drivers' gap acceptance behaviour under different road, traffic, vehicle and delay conditions.

6.4 Main Conclusions

The pilot research project succeeded in demonstrating that, due to low accident frequencies, a disaggregated level of modelling was not suitable for predicting accidents at intersections without traffic signals. Two simple accident prediction models at aggregated levels had been identified as optimal: the product-of-flow model and the total-delay model.

Both the product-of-flow model and the total-delay model performed well in predicting accidents at intersections. The two models had very similar model structures. In modelling the interaction between conflicting priority and non-priority flows at the intersection, the two models both employed traffic volume and / or delay data related to the priority and non-priority flows as model parameters. The results of accident prediction from the two models were almost identical.

The objective of developing a satisfactory accident prediction model can be met with either one of these two models. The choice of which model to be used in a particular application would depend on the specific application criteria.

Further analysis on the various accident parameters indicated that the underlying causes of safety were too complicated for an aggregate two-parameter model to explain. For example, consider the positive relationship between average delay and accident frequency, the increase in average delay may be due to several underlying factors or a combination of several factors.

As robust accident parameters can be influenced by more than one underlying factor, a heavy reliance on a single parameter accident statistic will easily mask the true cause(s) of the problem and will result in misleading conclusions. Therefore, in using single parameter accident statistics such as accident frequency and accident rate to prioritize intersections for safety improvements, the transportation engineer should be aware of the limitation of such a ranking approach.

Application of the aggregate models in a more detailed microscopic level of analysis is limited because of the robustness of the model parameters and the inability of the models to represent the safety situations at a more disaggregate level. The simple product-of-flow and total-delay models should be used primarily as a tool to indicate, on a comparative basis, general safety expectancies at intersections to pre-qualify intersections for more detailed safety analysis.

6.5 Practical Implications of Research Results

Either the product-of-flow or the total-delay accident prediction models can be used to determine the average objective safety of a T-intersection without traffic signals. For locations identified with unexpectedly high accident experiences, a more detailed level of analysis would be required to identify the actual safety problems and the underlying causes.

Increase in traffic safety can only be achieved if the right intersections are selected for improvement and the improvements are effective. It is therefore crucial to know the effectiveness of safety improvement measures under a particular combination of traffic and roadway conditions.

Safety should also be considered at a system level and be included as part of the transportation planning process. A system-wide measure of safety should be incorporated as one of the measures of effectiveness in traffic models to determine safety implication of the traffic system. In this respect, the effect of increased traffic on the traffic safety in the overall traffic system can be determined quantitatively. Accident costs can be included in the benefit / cost analysis in justifying for large scale traffic circulation changes or arterial upgrading.

6.6 Concluding Remarks

The research confirmed that a total-delay model could be used to predict accidents at intersections without traffic signals. Accident prediction results of the model were found to be almost identical to that of a product-of-flow model. Precautions for the application of both models are:

1. The models were tested using a small data size of 26 intersections where 129 accidents occurred over 4 years. The research served to investigate only the potential of the delay parameter in accident prediction.

2. The models were developed for weekday off-peak travel conditions for priority-ruled intersections. Application to other travel conditions has not been tested.

3. The models were evaluated over a range of traffic volume from 200 to 400 vehicles per hour for the priority flow (i.e., main road through volumes), and from 20 to 300 vehicles per hour for the non-priority flows (i.e., all four turning volumes). Transferability to intersections with heavier traffic volumes is not known.

4. The models were intended to be used as a robust guide to pre-qualify intersections for further safety analysis. To achieve better understanding of the safety problem at the intersection level, a more microscopic approach is recommended.

References

ABBES, C., D. JARRET and C. WRIGHT. 1987. Accidents at Blackspots: Estimating the Effectiveness of Remedial Treatment, with Special Reference to the 'Regression-to-mean' Effect. Traffic Engineering and Control, October 1987, Vol. 22 No. p.535.

ABOU-HENAIDY, M. 1993. Ph. D. Research Working Paper 5. Research Topic: Analysis of Driver Gap Acceptance Behavior Using Discrete Choice Modelling. Unpublished, Department of Civil Engineering, University of Alberta, Canada.

ABOU-HENAIDY, M. 1993. Ph. D. Research Working Paper 6. Research Topic: Analysis of Driver Gap Acceptance Behavior Using Discrete Choice Modelling. Unpublished, Department of Civil Engineering, University of Alberta, Canada.

ALBERTA TRANSPORTATION. 1985. Traffic Signal Control Display Standardization and Guidelines. Prepared by UMA Engineering Ltd. in Conjunction with Swanson Transportation Consultants Ltd.

AITKIN, M., D. ANDERSON, B. FRANCIS and J. HINDE. 1989. Statistical Modelling in GLIM. Clarendon Press. Oxford Science Publications.

AMUNDSEN, F. and C. HYDEN. 1977. Proceedings of the First Workshop on Traffic Conflicts. OSLO: Institute of Transport Economics.

BLOMQUIST, G. 1986. A Utility Maximization Model of Driver Traffic Safety Behaviour. Accident Analysis and Prevention, Volume 18, No. 15, p.371-375.

BRILON, W. 1988. Recent Developments in Calculation Methods for Unsignalized Intersections. Intersections Without Traffic Signals, Proceedings of an International Workshop 16-18 March, 1988 in Bochum, West Germany.

CHAN, P. and S. TEPLY. 1992. Simulation of Multilane Stop-controlled T-intersections by KNOSIMO in Canada. Intersections Without Traffic Signals II. Proceedings of an International Workshop on Intersections Without Traffi Signals. 18-19 July 1991 in Bochum, West Germany. Springer-Verlag Berlin Heidelberg 1991. p.308-319.

CHAN, P. and S. TEPLY. 1993. Traffic Signals At Intersections Without Traffic Signals. Proceedings of the Institute of Transportation Engineers District 7 Canada Annual Meeting in Edmonton, 1993.

CITY OF EDMONTON TRANSPORTATION DEPARTMENT DESIGN AND CONSTRUCTION BRANCH. 1986 - 1988. Arterial Roadway and Construction Schedule.

CITY OF EDMONTON TRANSPORTATION DEPARTMENT MONITORING UNIT. 1989. Accident Decoding Manual for 1985 - 1989 Computerized Data File.

ENVIRONMENT CANADA ATMOSPHERIC ENVIRONMENT SERVICE. 1989. Annual Meteorological Summary - Edmonton Municipal Airport.

GILCHRIST, R. B. FRANCIS and J. WHITTAKER (Editors). 1985. Lecture Notes in Statistics 32: Generalized Linear Models - Proceedings of the GLIM 85 Conference, Lancaster 1985. Springer-Verlag Berlin Heidelberg New York Tokyo.

GROSSMAN, L. 1954. Accident-exposure Index. Proc. of Highway Research Board. 33. 129-138. (includes discussion by O.K. Normann)

GROSSMANN, M. 1988. A Practicable Simulation Model for Unsignalized Intersections. Intersection Without Traffic Signals, Proceedings of an International Workshop 16-18 March, 1988 in Bochum, West Germany.

HAKKERT, A.S. and D. MAHALEL. 1978. Estimating the Number of Accidents at Intersections. From a Knowledge of the Traffic Flows on the Approaches. Accident Analysis and Prevention, VOlume 10, p.69-79.

HAUER, E. 1982. Traffic Conflicts and Exposure. Accident Analysis and Prevention, Volume 14, No. 5, p.359-364.

HAUER, E. J. NG and J. LOVELL. 1989. Estimation of Safety at Signalized Intersections. Transportation Research Record 1185, p.48-61.

INSTITUTE OF TRANSPORTATION ENGINEER. 1984. Canadian Capacity Guide for Signalized Intersections, First Edition. District 7, Canada. Editor: Teply, S.

KLEBELSBERG, D. 1982. Verkehrspsychologie. Springer-Verlag Berlin Heidelberg New York.

KYTE, M. and A. BOESEN. 1989. Traffic Data Input Program. Program Documentation and User's Manual. Version 2.0. Department of Civil Engineering, University of Idaho, Moscow, Idaho. April 1989.

McDONALD, J.W. 1953. Relationship Between Number of Accidents and Traffic Volumes at Divided Highway Intersections. Highway Research Board Bulletin. 74, p.7-17.

MAYCOCK, G. and R.D. HALL. 1984. Accidents at 4-Arm Roundabouts. Transport and Road Research Laboratory, TRRL Laboratory Report 1120, Crowthorne, Berkshire.

McCULLAGH, P. and J.A. NELDER. 1983. Generalized Linear Models. Chapman and Hall.

NELDER, J.A. and R.W.M. WEDDERBURN. 1972. Generalized Linear Models. Journal of the Royal Statistical Society, A(1972), 135, Part 370-383.

NUMERICAL ALGORITHMS GROUP. 1987. The Generalized Linear Interactive Modelling System - Release 3.77 Manual, User's Guide and Reference Guide. Payne C.D. (Editor).

PICKERING, D., R.D. HALL and M. GRIMMER. 1986. Accidents at Rural T-junctions. Transport and Road Research Laboratory, TRRL Research Report 65, Crowthorne, Berkshire.

PLASS, M. and W.D. BERG. 1988. Evaluation of Opportunity-based Accident Rate Expressions. Transportation Research Record 1111.

RAFF, M.S. 1953. Interstate Highway Accident Study. Highway Research Board Bulletin 74, p.18-45.

SABEY, B.E. 1983. Recent Developments and Research in Road Safety Remedial Measures. Road Safety in the 80's. Presented to Symposium, Salford, September 1983.

SABEY, B.E. and STAUGHTON, G.C. 1975. Interacting Roles of Road Environment, Vehicle and Road User in Accidents. Fifth International Conference of the International Association for Accident and Traffic Medicine, and the Third International Conference on Drug Abuse of the International Council on Alcohol and Addiction, London, September 1975.

STANLEY ASSOCIATES ENGINEERING LTD. 1983. A Rural Intersection Impedance Model. Prepared for Alberta Transportation, April 1983.

TANNER, J.C. 1953. Accidents at Rural Three Way Junctions. Journal of Institution of Highway Engineers. 2(1), 56-57, 1953.

TEPLY, S. 1987. Driver Behavior Assessment as a Basis for "Before" and "After" Studies. Traffic Engineering and Control, London, England. July/August 1987, p.402-407.

TEPLY, S. 1988. Lecture Notes for Civil Engineering 614, Transportation Engineering, a graduate course at the University of Alberta, Canada. Unpublished.

TRANSPORTATION RESEARCH BOARD. 1975. Highway Capacity Manual. Special Report 165. Chapter 10.

WILLIAMS, M.J. 1981. Validity of the Traffic Conflicts Technique. Accident Analysis and Prevention, Volume 13, p.133-145.

APPENDIX A

Simulation of Multilane Stop-controlled T-intersection by KNOSIMO in Canada

Paper presented in the Second International Conference on Intersections Without Traffic Signals, Bochum, Germany. Published in Intersections Without Traffic Signals II 1991. p.308-319.

SIMULATION OF MULTILANE STOP-CONTROLLED T-INTERSECTIONS BY KNOSIMO IN CANADA

Paul L. Chan and Stan Teply

University of Alberta,
Edmonton, Alberta, Canada

1. SUMMARY

This paper reports on an application of the KNOSIMO traffic simulation program for priority-ruled intersections in Edmonton, Alberta, Canada. This University of Alberta study was restricted to urban multilane stop-controlled T- intersections and its main objective was to investigate the relationship between delays and safety. As a result, it required a reliable representation of delays. A relatively large sample of delay data was collected, but it was necessary to supplement it by simulation. The KNOSIMO program was validated and used to that end.

The paper describes the experience with the program and comments on its applicability to various conditions at the studied T- intersections.

2. DATA

Delay data, together with volume and headway information, were collected by using a video camera in January and February 1990, at 26 T-intersections in Edmonton. The surveys were limited to off- peak periods on weekdays with normal driving conditions. In total, over 15 hours of real time events with 21,069 vehicle arrivals were analyzed and transferred to a computer data base with the assistance of the TDIP program [1].

The collected delay data had a high degree of variability. In order to improve the statistical significance of the samples, more surveys would have been necessary. This approach, however, was not considered practical.

To generate additional delay information, it was decided to employ the KNOSIMO simulation model [2]. The data measured in the field were considered sufficient to test the suitability of the program for the needs of the overall project.

3. THE KNOSIMO PROGRAM

KNOSIMO is a traffic model for intersections without traffic signals developed at Ruhr University in Germany for the Federal Minister of Transport [2,3]. It is a microscopic, event-oriented simulation program and is applicable to 3- or 4-legged intersections. Shared and exclusive turning lane situations, as well as time-varying traffic demand and different vehicle types can be considered.

The vehicles are generated in the program using hyperlang headway distributions, calibrated to German conditions to realistically represent single lane traffic streams. The operation of non-priority traffic maneuvers is determined from gap acceptance principles using erlang distributions. The program is interactive regarding data input and file handling operations.

4. DIFFERENCES IN APPLICATION CONDITIONS

In order to simulate traffic conditions for multilane intersections, it was necessary to address the differences in the conditions considered in the program and at the studied locations.

The investigation started from the premise that the program itself would not need modifications and that the input values could be appropriately adjusted.

4a. Geometric features, critical gaps and move-up times

The differences in geometric features between the KNOSIMO and the studied intersections are illustrated in Figure 1. They imply the need for a different application of the critical gap and move-up time criteria [4]. The critical gaps and move-up times used in KNOSIMO are for 2-lane conditions but longer time periods are required to cross a multilane roadway. Moreover, the critical gaps and move-up times used in KNOSIMO were calibrated for driver behaviour in Germany. Whether and how these critical gaps could be transferred to North American conditions was unknown.

Unfortunately, the scope of the overall project did not allow a detailed investigation of these problems. To accommodate them, higher main road speeds were used to raise the critical gaps used by the program. For all approaches, test results showed that a 10 km/h higher main road speed gave better delay predictions.

4b. Headway distribution

KNOSIMO "feeds" traffic from one- lane approaches only. This was a serious problem because the headway distributions for the two- to three- lane intersection approaches included in the study were different. While on single lanes zero headways are not possible, and the number of very short headways is small, they frequently occur on multilane roadways.

Studied locations KNOSIMO

Figure 1: Comparison of intersection conditions in Edmonton and in the KNOSIMO program

The most important parameters are the values of critical gaps which are normally in the range from 4 to 10 seconds. The availability of headways above the critical gap value presents the "practical headway" for the non-priority traffic, and the number of these headways in the traffic stream on the main road determines capacity and delays on the minor road.

In comparing measured multilane headways to KNOSIMO headways in the range greater than 4 seconds, it was found that the difference was relatively consistent. Figure 2 shows that, for headways greater than 4 seconds, an equivalent "one-lane volume" of 400 vph in the KNOSIMO program can be used to approximate the actual headway distribution at a multilane approach with a volume of 690 vph. Figure 3, which includes thirteen intersection approaches, exhibits a linear relation between the single lane and multilane distributions. This linear transformation was employed for input volume conversions.

4c. Traffic interactions

The way individual traffic movements influence each other at a multilane intersection is different than at a 2-lane roadway. For this reason, conflicting volumes for certain non-priority maneuvers needed further adjustments.

It was found, for example, that when the full "eastbound" through volume used for the "northbound" right-turn maneuver from the minor road, delays was overestimated. This maneuver was only slightly affected by the inside lane of the main road traffic. As a result,

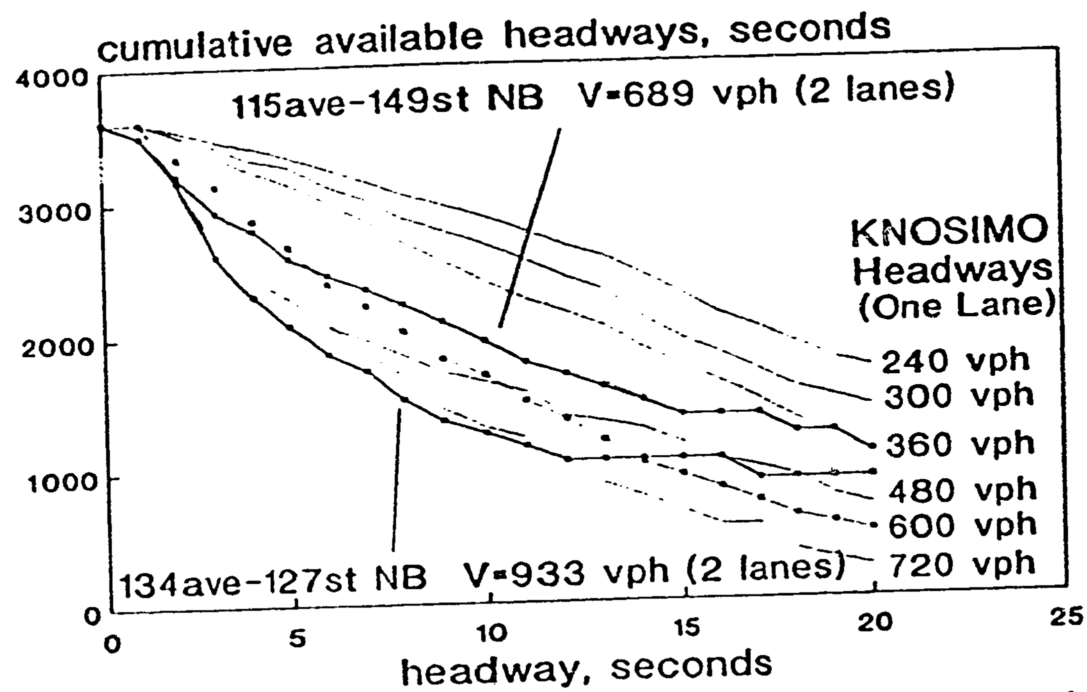


Figure 2 Comparison of measured cumulative available headways (2-lane) to various profiles assumed in KNOSIMO for different 1-lane volumes

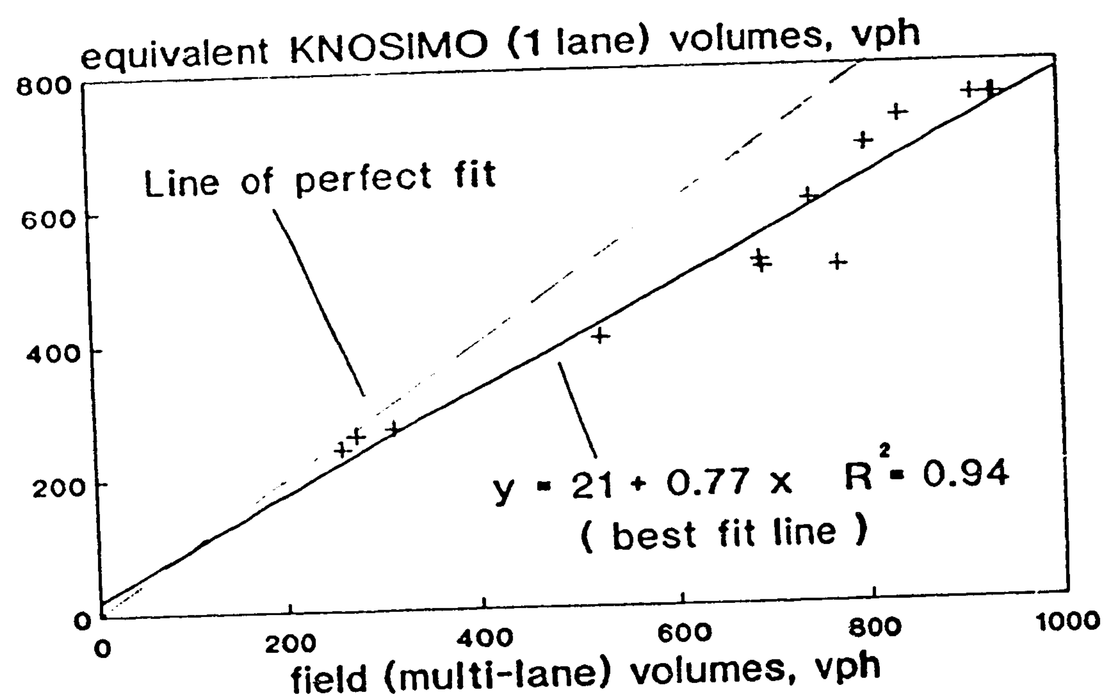


Figure 3 Correlation between actual field (2-lane) volumes and the equivalent 1-lane KNOSIMO volumes

only the curb-lane volume was applied, but since it already represented a one-lane situation, it was not adjusted further to an equivalent one-lane volume.

Nevertheless, for the "northbound" left-turn traffic from the minor road, it was necessary to keep the full volume of the "eastbound" traffic. To represent the actual conflicting "westbound" through traffic volume from the right, only the inside lane volume was used.

As a result, it was necessary to run the KNOSIMO program twice with somewhat different volumes in order to obtain proper delay generating situations.

4d. Effect of platooning

In an urban setting, priority intersections are frequently close to signalized intersections. The operation of these signals which creates platoons in the traffic stream can have a considerable effect on non- priority movements at intersections without signals. The platooning effect diminishes gradually with increasing distance downstream because of platoon dispersion.

No adjustments to the KNOSIMO input were made to accommodate these effects. Nevertheless, the results obtained from the KNOSIMO runs were compared to the measured delays, and are discussed in the next section of this paper.

4e. Type of delay

KNOSIMO calculates the average *overall* delay experienced by the drivers which includes time needed to decelerate and accelerate. The field delay data, however, were recorded as *stopped* delay. To modify the field data, the "zero- volume delays" (11 seconds for minor road left turns and 10 seconds for other movements) were added to make them compatible with delays generated by KNOSIMO.

5. RESULTS AND EVALUATION

The measured and the KNOSIMO generated delays for all of the three non- priority maneuvers at the studied T-intersections are compared in Figure 4. Considering the multiplicity of conditions and the differences among the three movements, the graph exhibits a reasonable degree of trend consistency. On average, however, the KNOSIMO generated delays are about 35% shorter than those measured. The underestimation of the actual delay is especially severe for higher values.

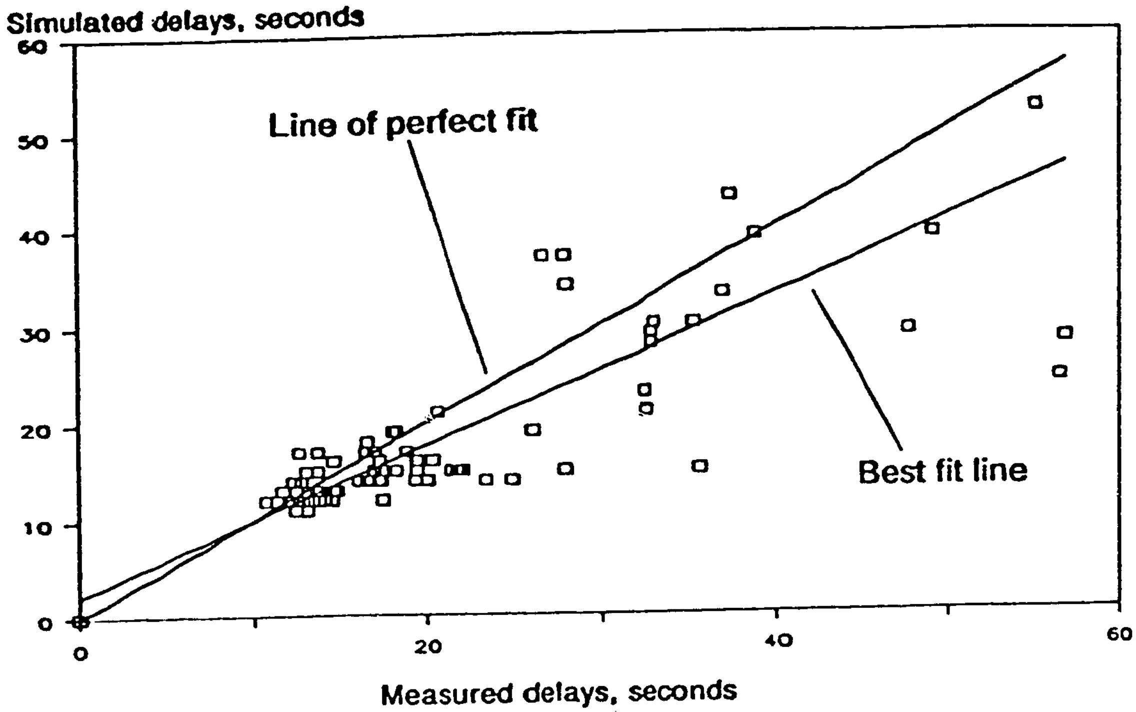


Figure 4: Comparison of measured and simulated delays for all three non-priority maneuvers

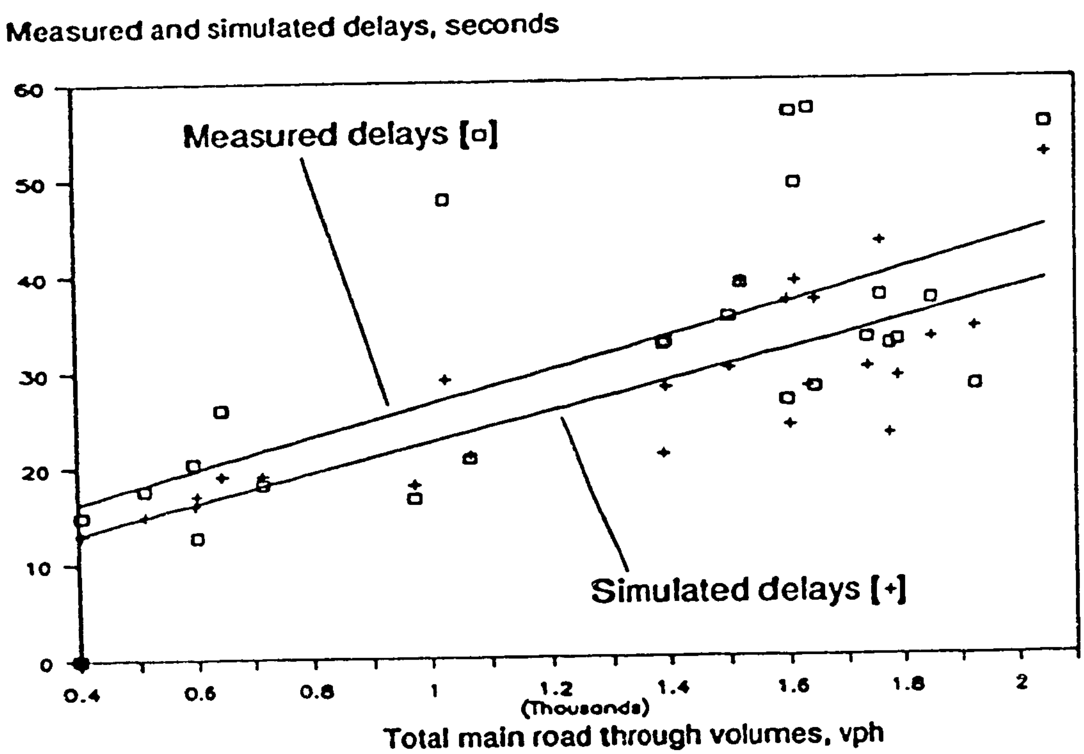


Figure 5: Variations in measured and simulated delays of minor road left turns with respect to major road through volumes

The fit, however, was quite different for individual movements. This was not surprising, since each non-priority maneuver depended on different exposure to other traffic streams, degree of priority and gap acceptance conditions which are discussed below.

5a. Comparison by the type of movement

The simulation results for the left- turn maneuver from the minor road were only about 11% lower than the measured delays in the range below 40 seconds. Above that value, however, KNOSIMO underestimated actual delay by about 60% (Figure 4).

For the right-turn movement for the minor road, KNOSIMO underestimated the actual delay quite significantly. The average was about 40% for shorter delays to over 100% for delays longer than 20 seconds. The discrepancy, however, was caused in part by the previously discussed volume adjustment, which was probably somewhat exaggerated, and partly by the impact of different sight distances and other local conditions. For instance, the two locations with the longest measured delays for minor road right-turners had sight distance problems.

The simulation results for left-turn movement from the main road across the path of the opposing main road traffic stream were very satisfactory for delays under 20 seconds. On average, they were only 8% lower than the measured values. The underestimation beyond 20 seconds, however, was higher.

5b. Comparison by volumes on the main road

Figures 5 illustrates the trend of increasing delay to minor road left-turners with increasing conflicting priority volumes. The conflicting traffic was taken as the sum of the main road volumes in both directions.

The underestimation, again, is noticeable and, as would be expected, the simulated values are more consistent than the measured data. The relatively constant randomness of the differences along all volume ranges suggests that the discrepancies cannot be explained by different volume categories. Similar trends were observed for minor road right- turners and major road left- turners.

Figure 6 shows that the measured probability of short headways is higher for approaches with higher traffic volumes. As a result, there are less of the available headways, resulting in longer delays. The mode of the distribution shifts to the left with decreasing volumes which means that there were less followers in the traffic stream.

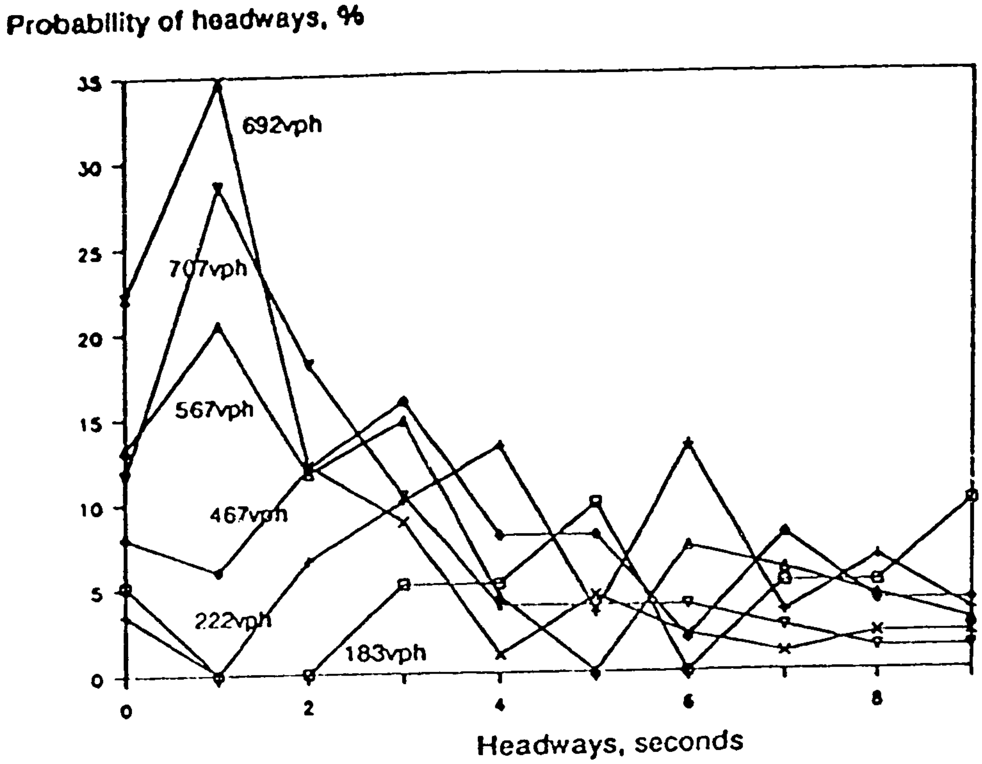


Figure 6 Comparison of short headway distributions for various approach volumes.

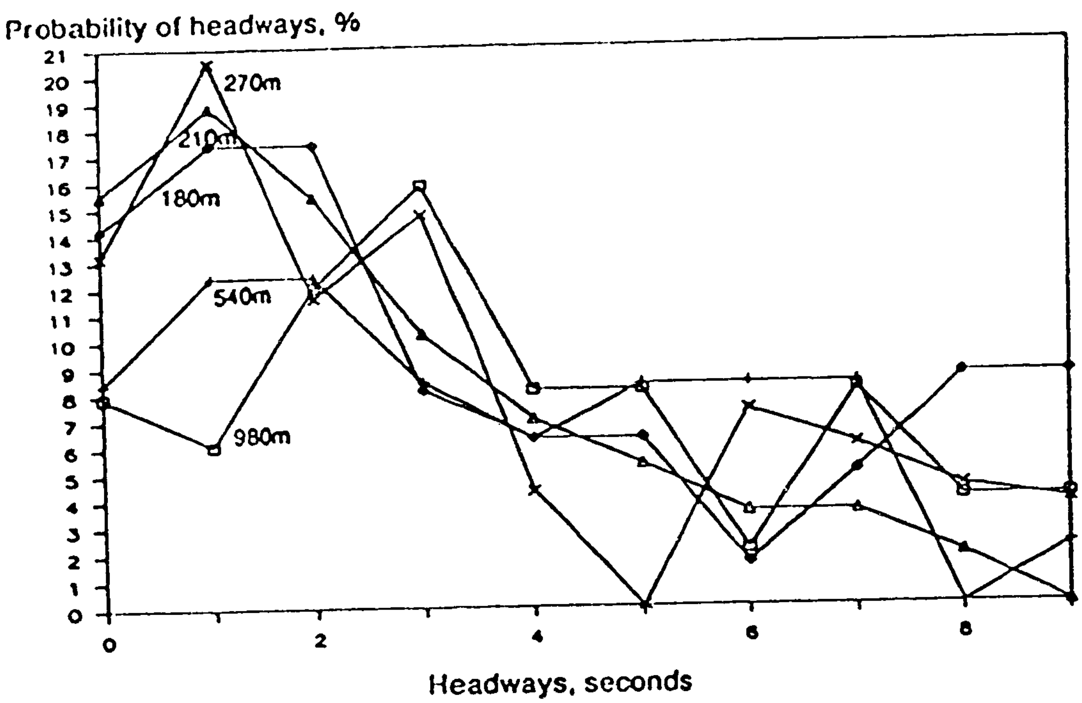


Figure 7: Comparison of short headway distributions for various distances from traffic signals (for 450-650vph)

5c. Comparison by distance from upstream traffic signals

It was observed that headway distributions at intersections with low volumes were almost unaffected by the distance to the nearest traffic signal. The relative uniformity of the distributions was interesting.

The greatest impact of distance on headway distributions was for intersections with medium volumes between 400 to 700 vph, as illustrated in Figure 7. Similar to Figure 6, the mode of the distributions is shifted to the left with increasing distance from traffic signals. This shift indicates a substantial reduction of the proportion of short headways resulting from platoon dispersion. Examination of headway distribution data at other locations with high volumes exhibited a similar, but less pronounced, mode shift.

The time shift between the arrivals of the main platoons in the two opposing traffic streams on the main road also influenced non-priority maneuvers. Delays for minor road left- turners were much longer if the two platoons in the opposite directions on the main road arrived in a staggered fashion.

Figure 8 shows the effect of the volume in both main road directions on available headways for left- turn traffic from the minor road. The low points on the graph indicate that at these locations there were less available gaps than at the high point locations. The intersections in these two groups are significantly different. For example, two of the four low points represent a street (99th Street) with short blocks, close spacing of traffic signals, high level of business activity and more trucks, compared to the street with five of the seven high point locations (149th Street).

6. COMPARISON WITH HIGHWAY CAPACITY MANUAL

All of the 26 locations were also analyzed using the 1985 Highway Capacity Manual (HCM) [5] method. The outcome was compared with the results of the simulation and surveys.

HCM gives reserve capacity as the output for each non-priority maneuver and relates its value to a particular Level of Service. There is no direct link to the corresponding delay. Brilon [4] estimated the upper delay limits for different reserve capacity levels and Tracz [6] defined a set of functions representing the relationship between reserve capacity and delay. These suggested transformations, together with the correlation of the measured delay to the calculated reserve capacity of each non-priority maneuver are shown in Figure 9.

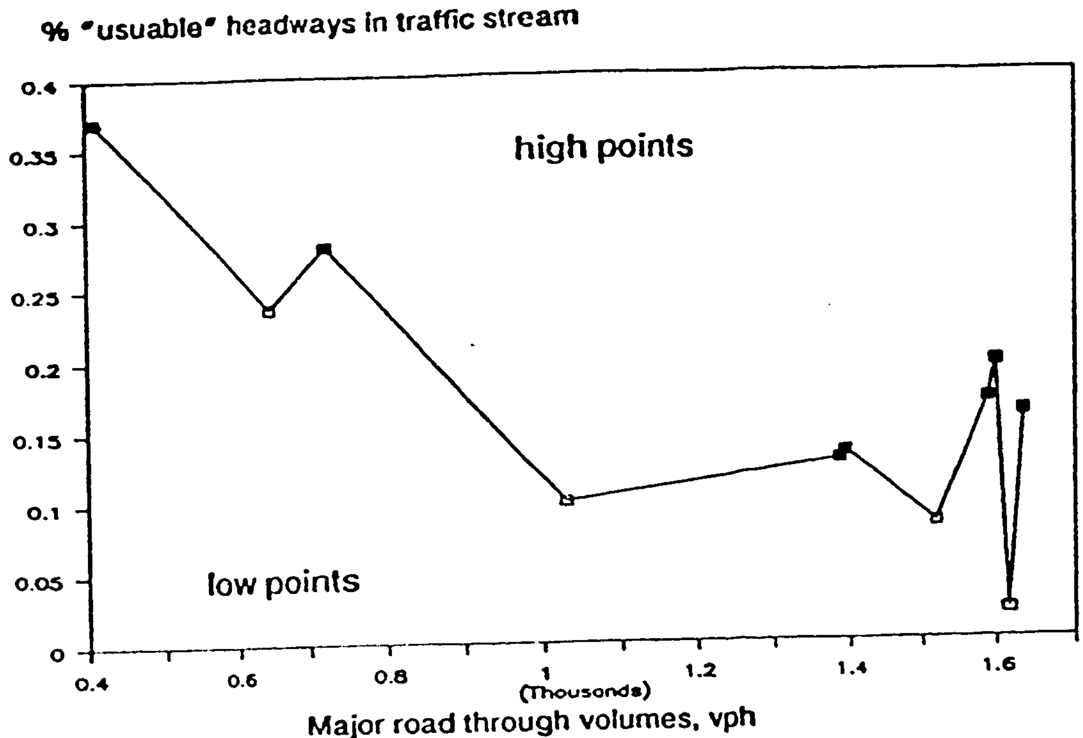


Figure 8: Effects of main road traffic interaction on availability of "usuable" headways for minor road left turn traffic

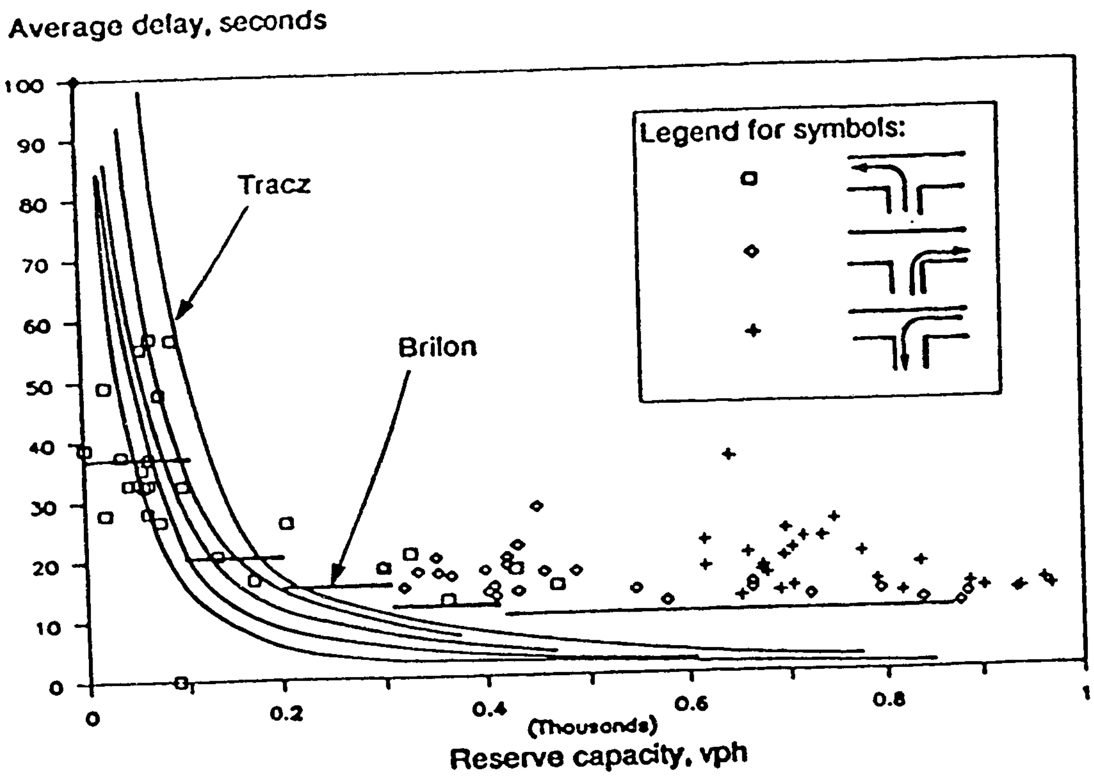


Figure 9: Comparison of reserve capacity [5] and measured delays. Relationships suggested by Brilon [4] and Tracz [6] are included.

In order to convert the HCM reserve capacity levels to delay, the suggested limits were applied, but the results did not agree well with the measured delay. The minor road left turns, with relatively low reserve capacities, were an exception. For the other two non-priority maneuvers, with reserve capacities often over 400 vph, the measured delays are considerably higher than the suggested delay levels. It appears, from this limited evidence, that the HCM method provides only a general indication of the degree of difficulty in making a non-priority maneuver. Figure 9 shows that similar measured delays do not necessarily coincide with similar reserve capacity. Moreover, definitive trends for all three movements are noticeble, different from the relationships proposed by Brilon and Tracz. It seems that a relationship based on a single average delay for all maneuvers may not be appropriate.

Another comparison showed that KNOSIMO generally gives good estimates when compared to the measured delay, except in cases when the reserve capacity is below 100 vph, or for the right turns from the minor road, possibly because the problems of our liberal input modification.

7. CONCLUSIONS

Traffic volumes, sight distance, degree of platooning generated by upstream signals, and other conditions at the 26 studied T-intersections varied considerably. Although many of these variables could not be considered in the KNOSIMO program and its application, the results of the verification were considered very satisfactory. As a consequence, KNOSIMO was used to generate additional information for the main part of the project, i.e. an investigation of the relationship between delay and safety.

Adjustment of arrival volumes and speeds on the main road proved to be a good instrument to account for the effect of local geometric conditions and, possibly, for the different driver behaviour. These adjustments successfully affected main road arrival headway distribution and minor road gap acceptance in the critical ranges.

Naturally, this study had a number of limitations. The major constraint was the fact that the verification of the KNOSIMO program formed only a minor objective of the overall project. Nevertheless, the study demonstrated that KNOSIMO is a sufficiently robust investigation tool and can be forced by relatively simple input adjustments to perform well in a variety of conditions. Additional output information on queues provided by the program was found useful, although it was not employed in this project.

8. ACKNOWLEDGMENTS

The support for this project has been provided by the Natural Sciences and Engineering Research Council of Canada as a part of an operational grant, and by the City of Edmonton in the form of traffic and other data. The authors also wish to acknowledge gratefully the help of Professor Brilon in providing the access to KNOSIMO and for additional comments, Professor Kyte for the TDIP program, and Ms. Butt for the assistance in the translation of the KNOSIMO User Guide.

9. REFERENCES

1. Kyte,M., Boesen, A., Traffic Data Input Program", Program documentation and Iser's Manual, Version 2.0, Department of Civil Engineering, University of Idaho, Moscow, Idaho, April 1989

2. Grossmann, M.," KNOSIMO - A Practicable Simulation Model for Unsignalized Intersections", Proceedings of an International Workshop on Intersections Without Traffic Signals, W. Brilon, editor, Springer Verlag, 1988, pp. 263-273

3. "KNOSIMO Version 3.4, Benutzeranleitung", (KNOSIMO User Guide, English translation by Chan, P. and Butt, K.), Lehrstuhl fuer Verkehrswesen, Ruhr-Universitaet Bochum, September 1988

4. Brilon, W., "Recent Developments in Calculation Methods for Unsignalized Intersections in West Germany", Proceedings of an International Workshop on Intersections Without Traffic Signals, W.Brilon, editor, Springer Verlag, 1988, pp. 111-153

5. "Highway Capacity Manual", Special Report 209, Transportation Research Board, Washington, D.C., 1985, Chapter 10.

6. Tracz,M., Chodur,J., Gondek,S., "The Use of Reserve capacity and Delay as the Complementary Measures of Junction Effectiveness", Proceedings of the Eleventh International Symposium on Transportation and Traffic Theory, Elsevier Science Publishing Co., 1990.

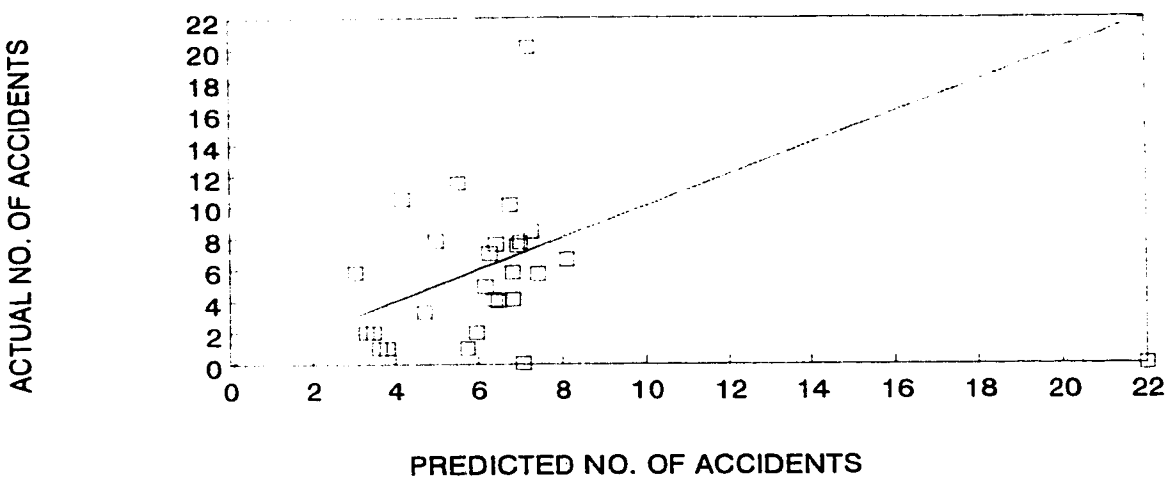
APPENDIX B

Goodness-of-fit Plots

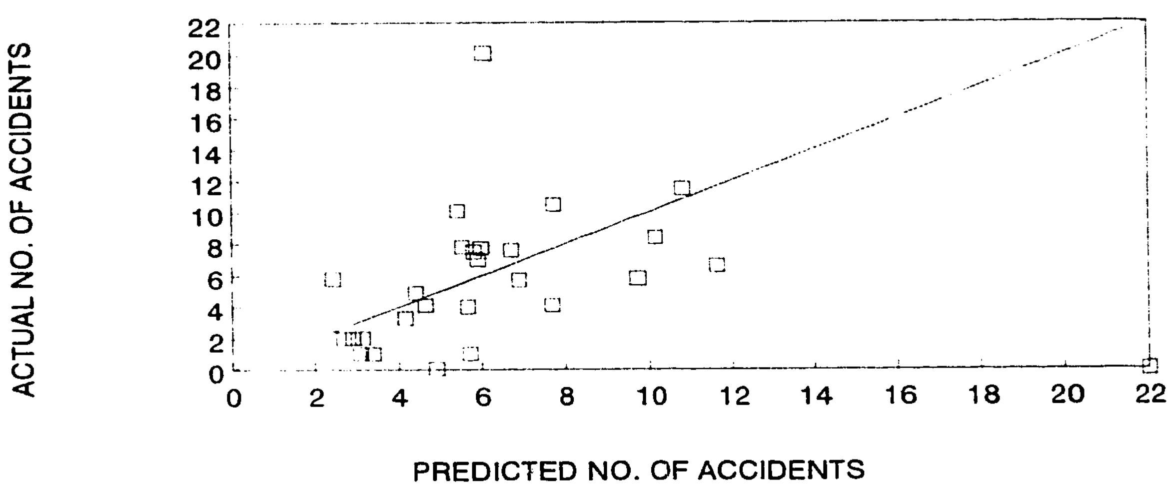
Summary of Parameter Estimates

Selected GLIM Listings

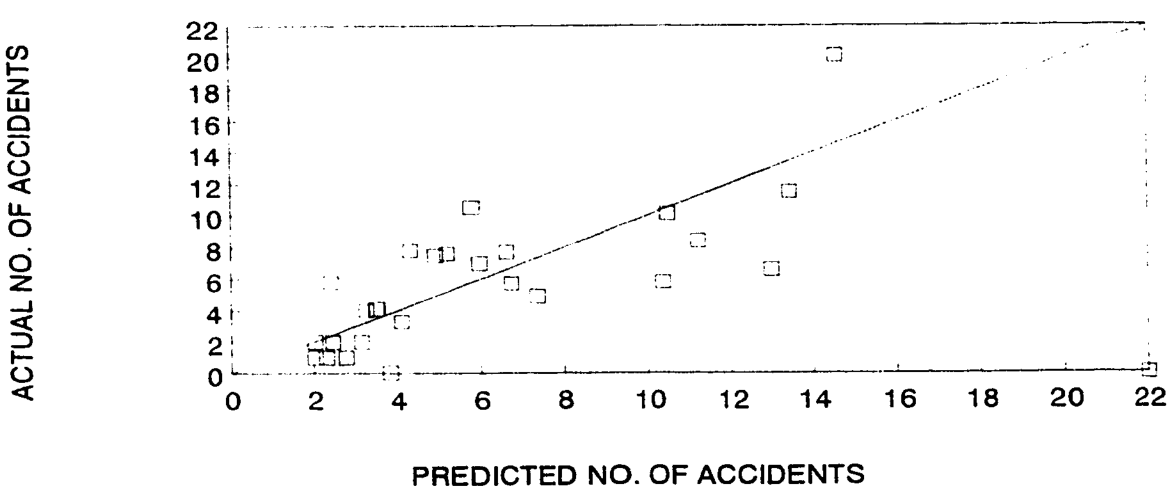
MODEL 111

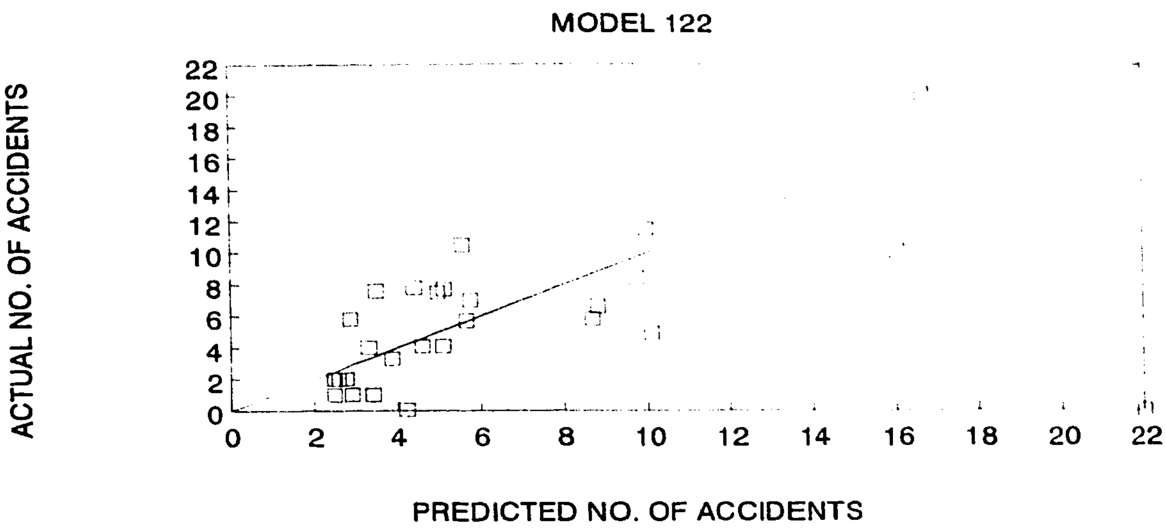


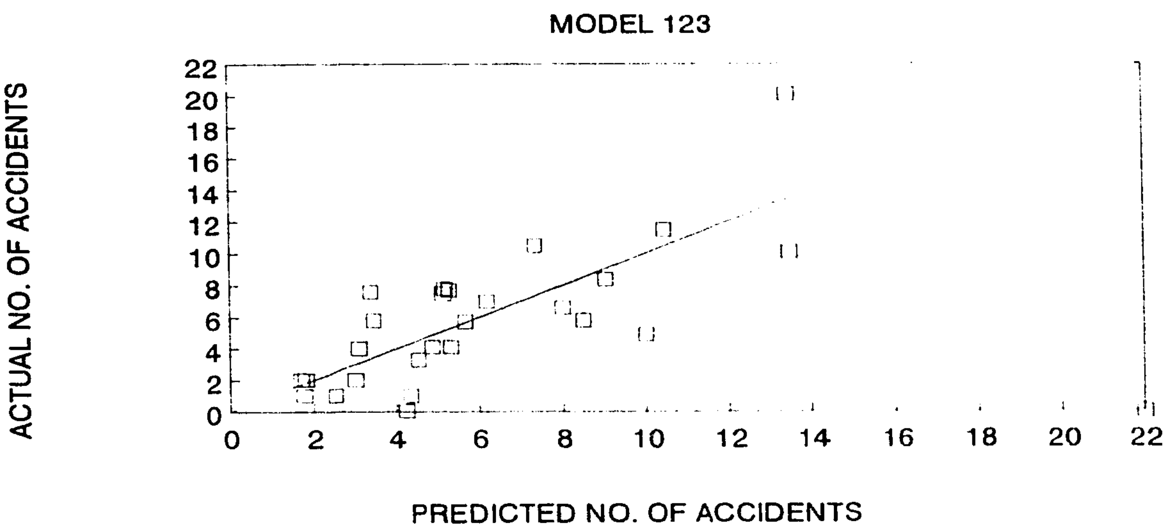
MODEL 112

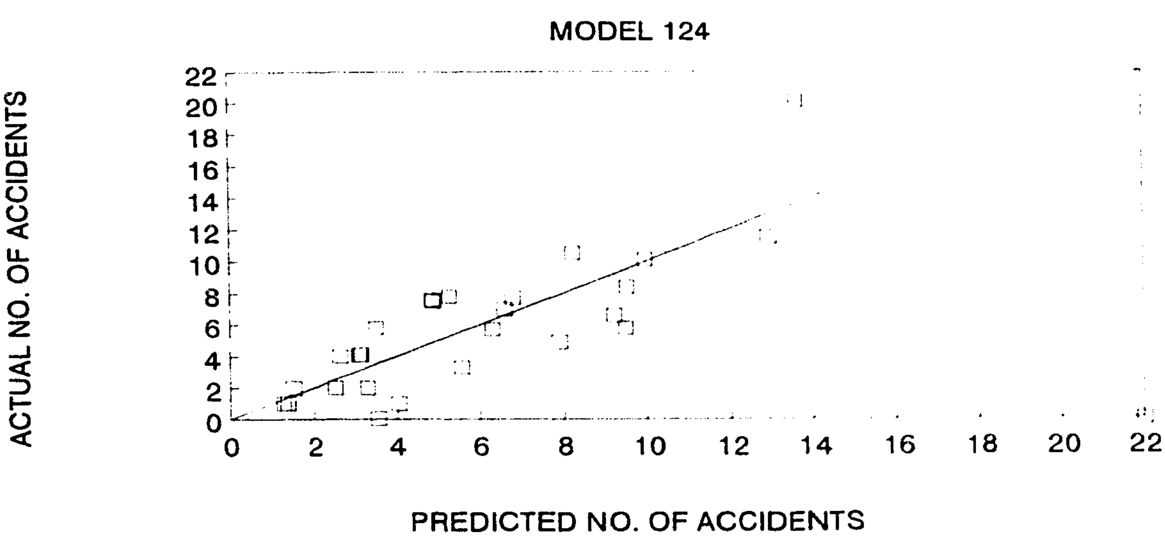


MODEL 121

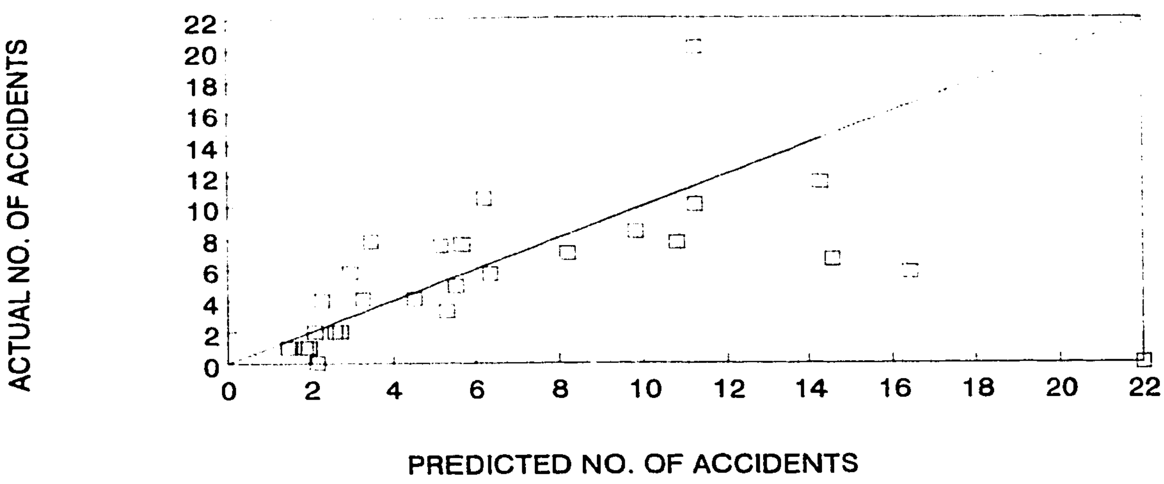



MODEL 122
ACTUAL NO. OF ACCIDENTS
PREDICTED NO. OF ACCIDENTS

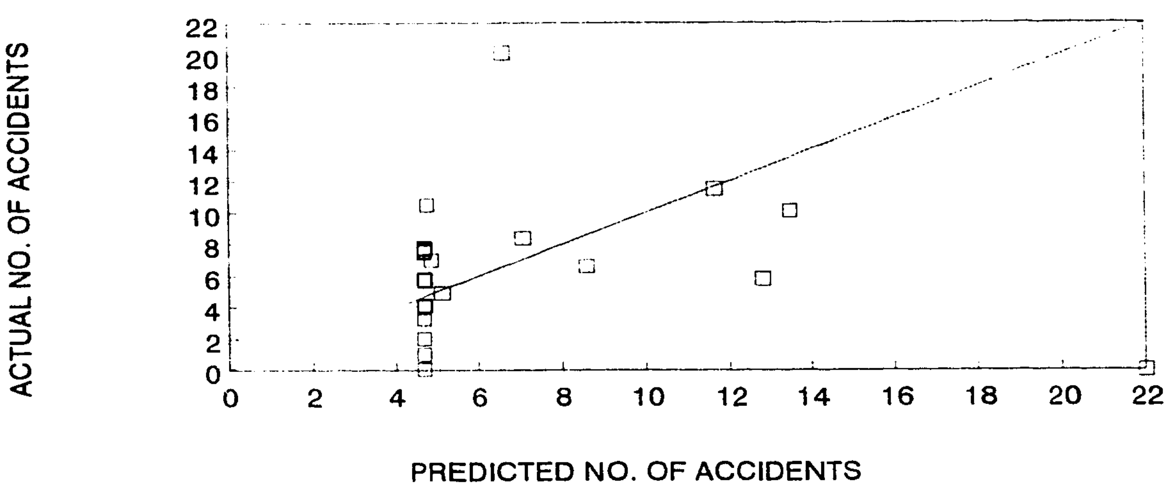

MODEL 123
ACTUAL NO. OF ACCIDENTS
PREDICTED NO. OF ACCIDENTS


MODEL 124
ACTUAL NO. OF ACCIDENTS
PREDICTED NO. OF ACCIDENTS

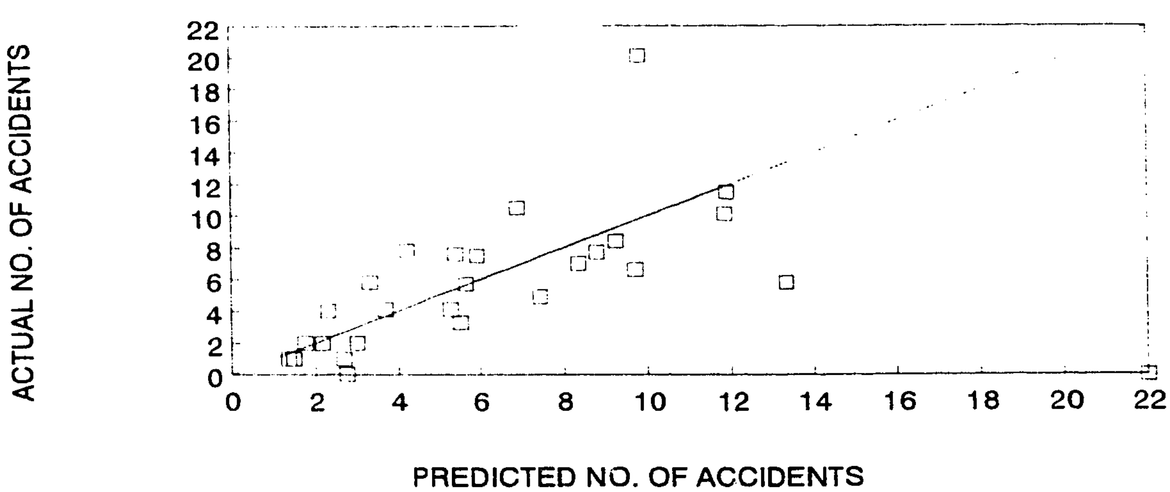
MODEL 211

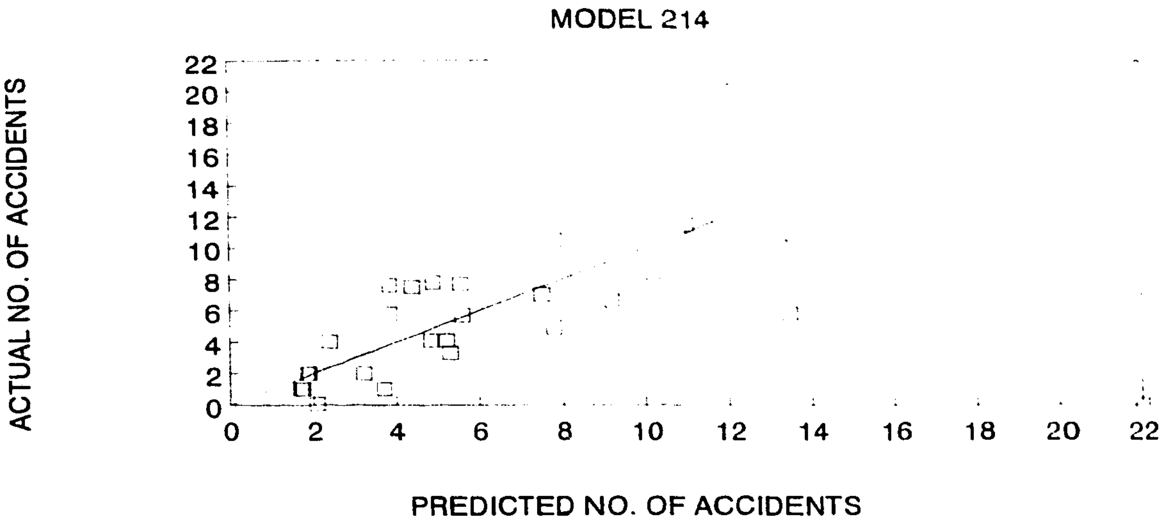


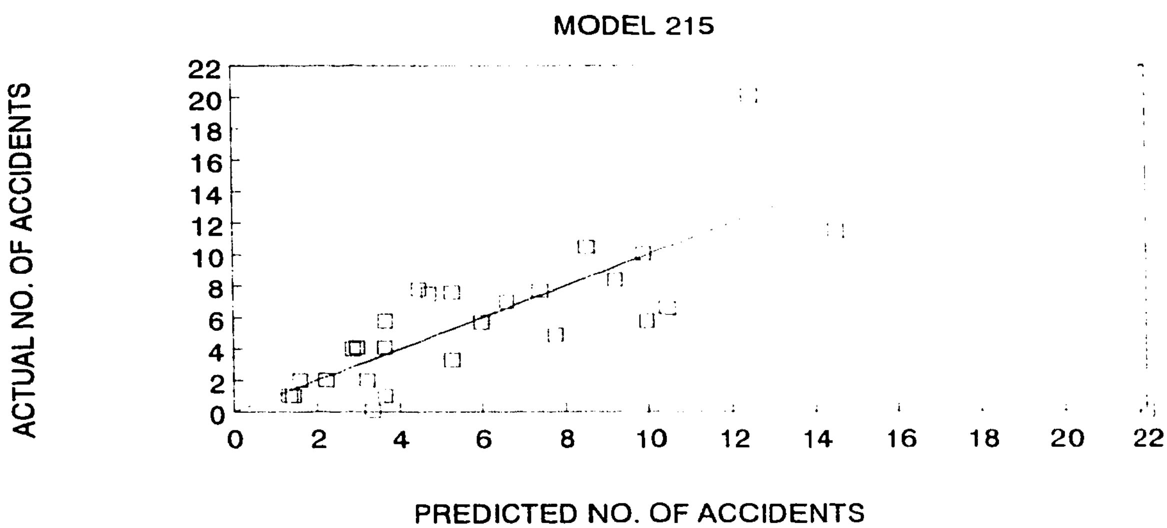
MODEL 212

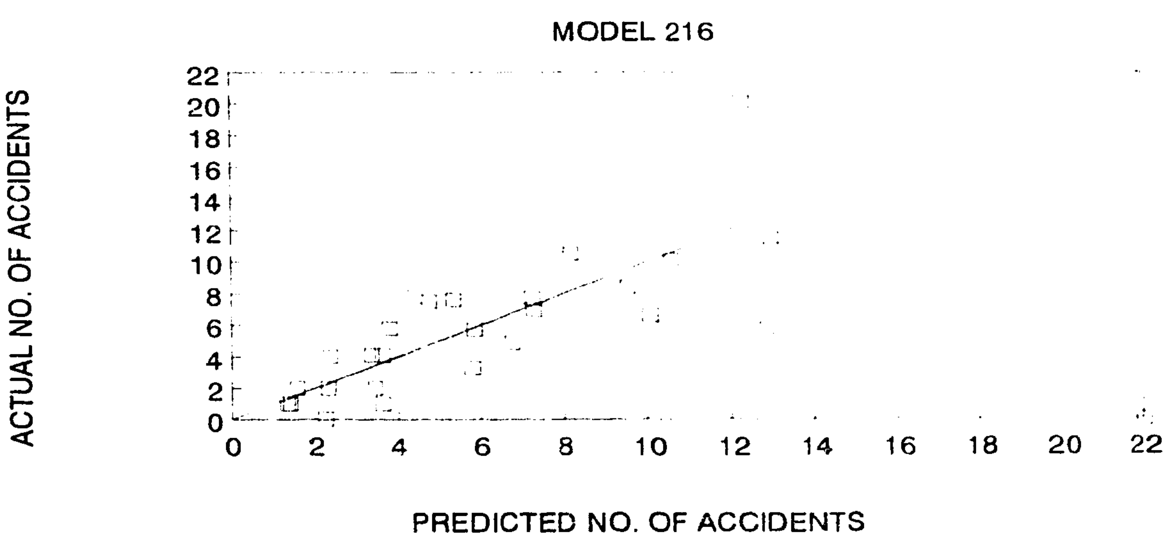


MODEL 213

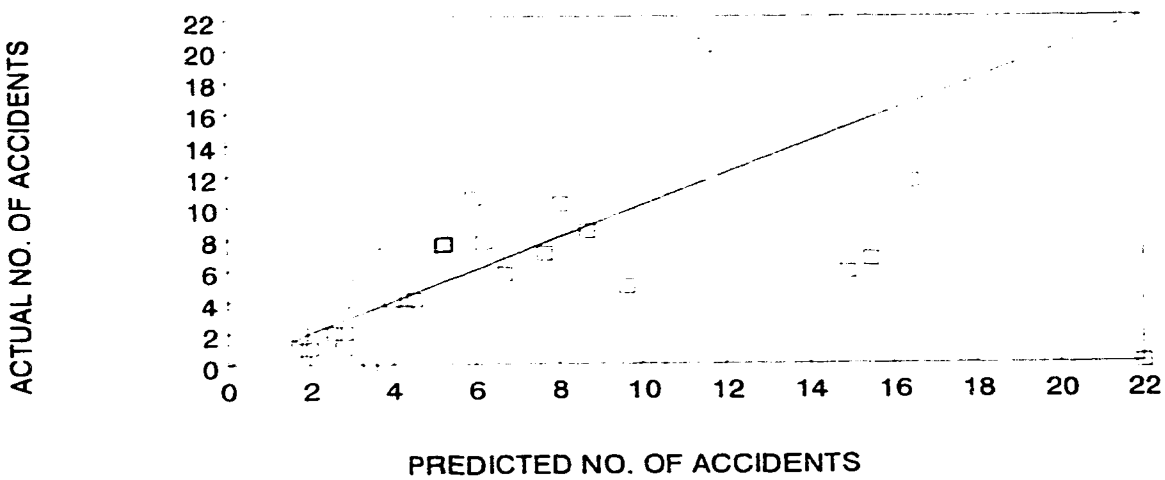



MODEL 214
ACTUAL NO. OF ACCIDENTS
22
20
18
16
14
12
10
8
6
4
2
0
0 2 4 6 8 10 12 14 16 18 20 22
PREDICTED NO. OF ACCIDENTS

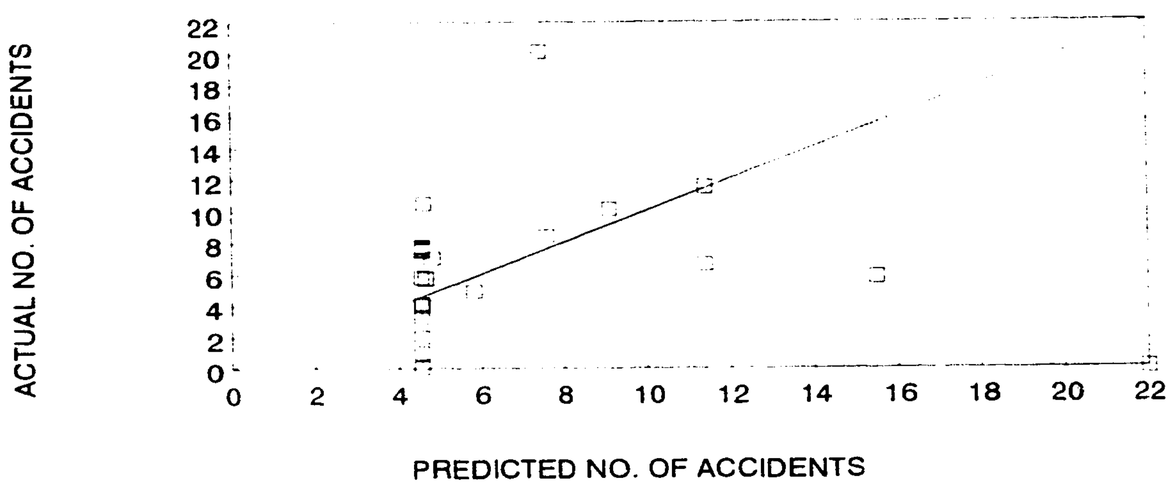

MODEL 215
ACTUAL NO. OF ACCIDENTS
22
20
18
16
14
12
10
8
6
4
2
0
0 2 4 6 8 10 12 14 16 18 20 22
PREDICTED NO. OF ACCIDENTS


MODEL 216
ACTUAL NO. OF ACCIDENTS
22
20
18
16
14
12
10
8
6
4
2
0
0 2 4 6 8 10 12 14 16 18 20 22
PREDICTED NO. OF ACCIDENTS

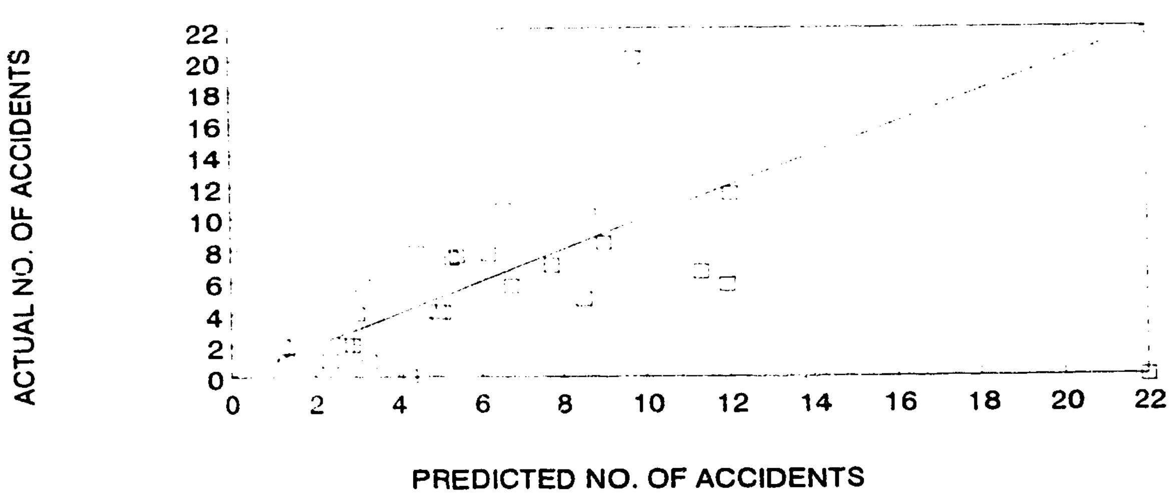
MODEL 221

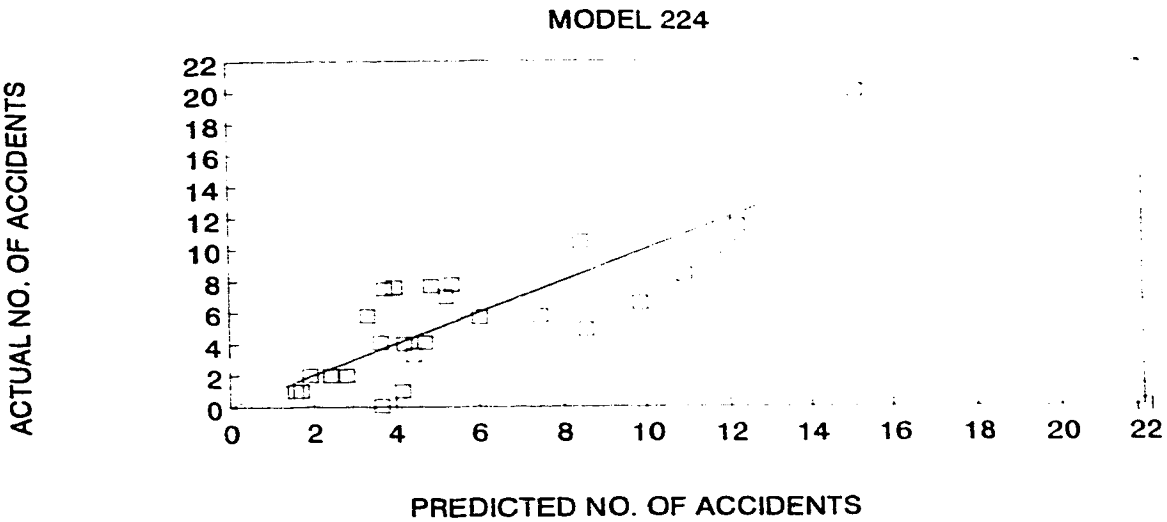


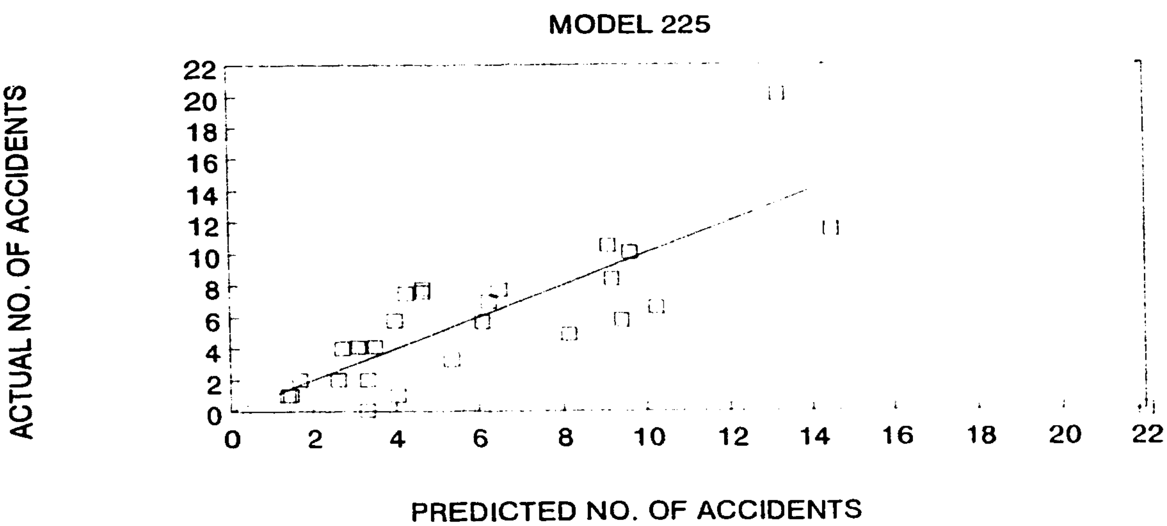
MODEL 222

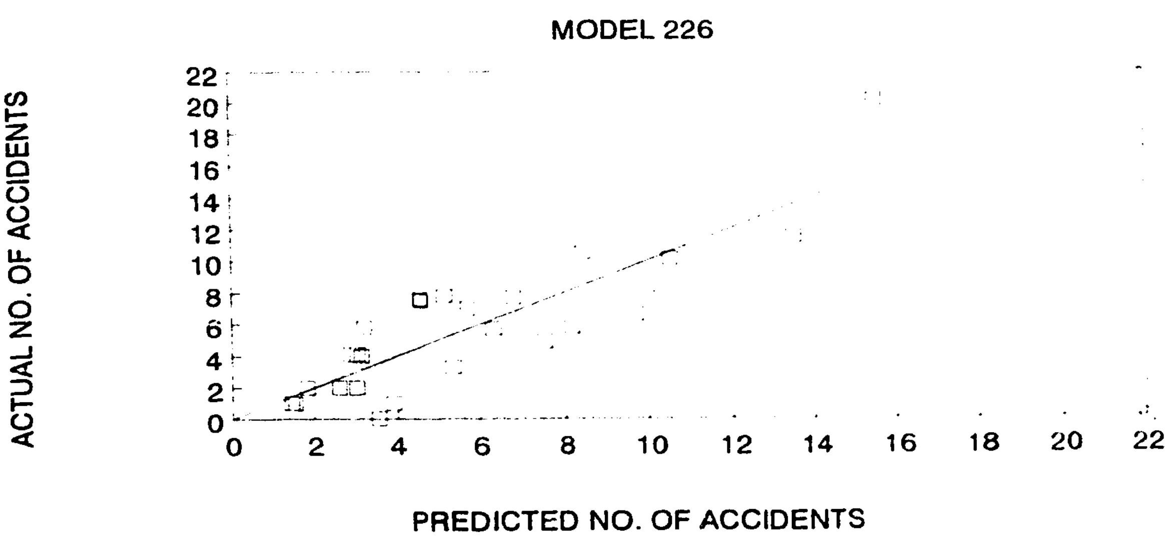


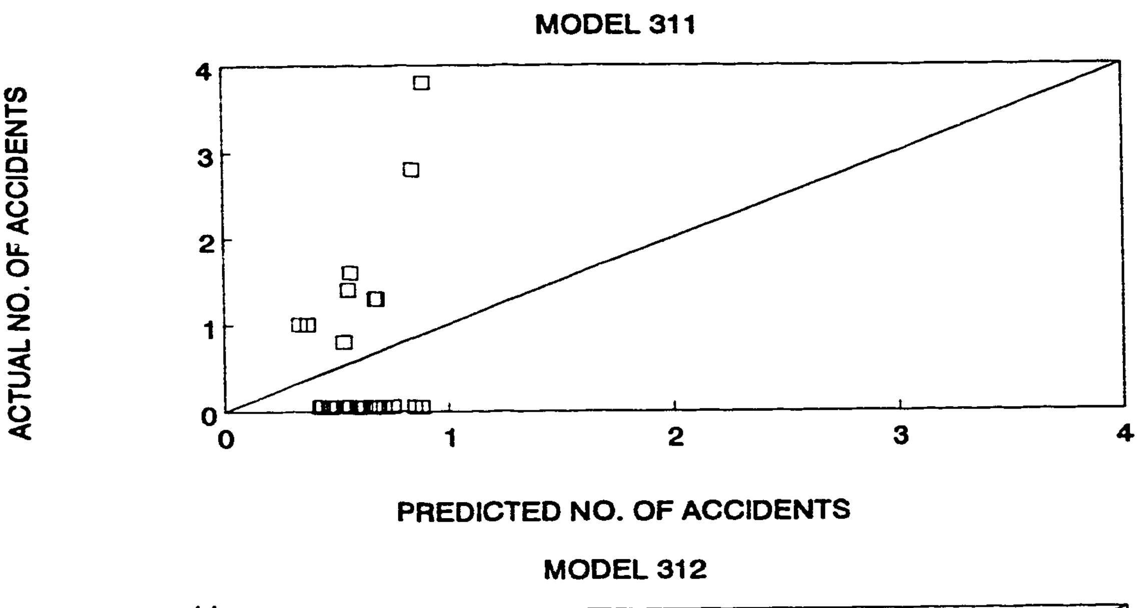
MODEL 223

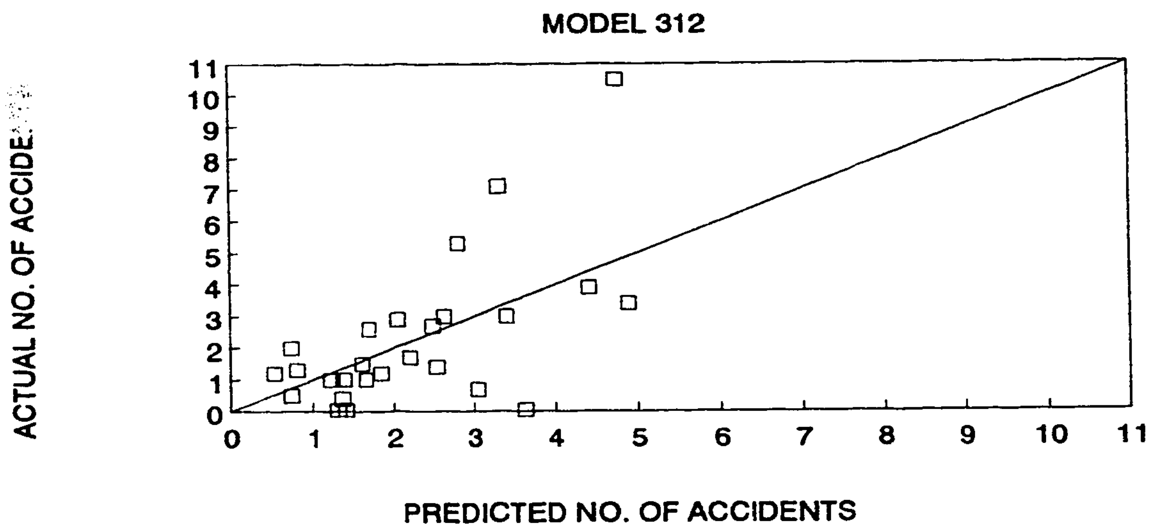


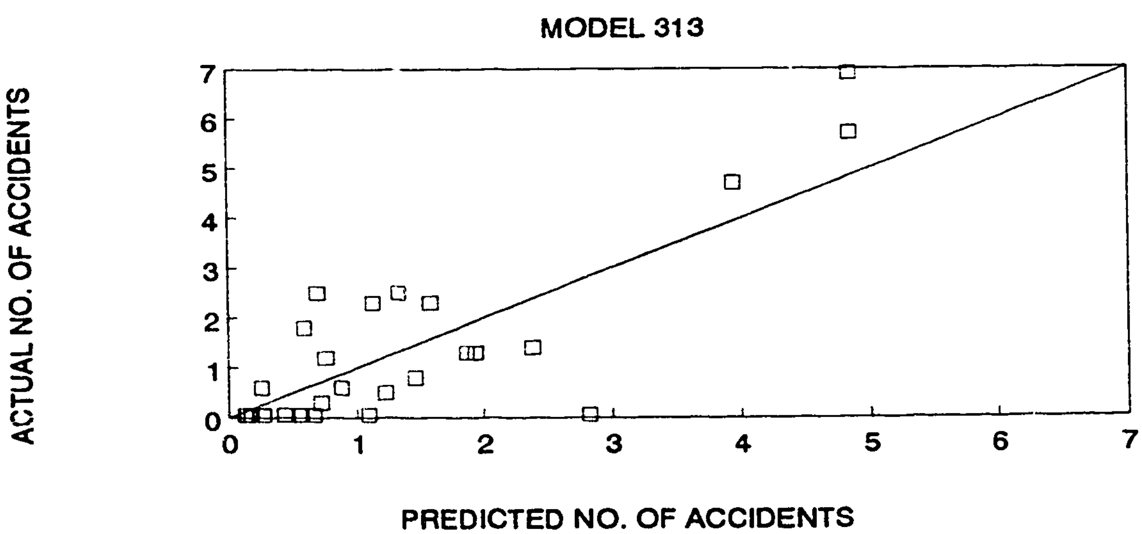

MODEL 224
ACTUAL NO. OF ACCIDENTS
0 2 4 6 8 10 12 14 16 18 20 22
PREDICTED NO. OF ACCIDENTS


MODEL 225
ACTUAL NO. OF ACCIDENTS
0 2 4 6 8 10 12 14 16 18 20 22
PREDICTED NO. OF ACCIDENTS

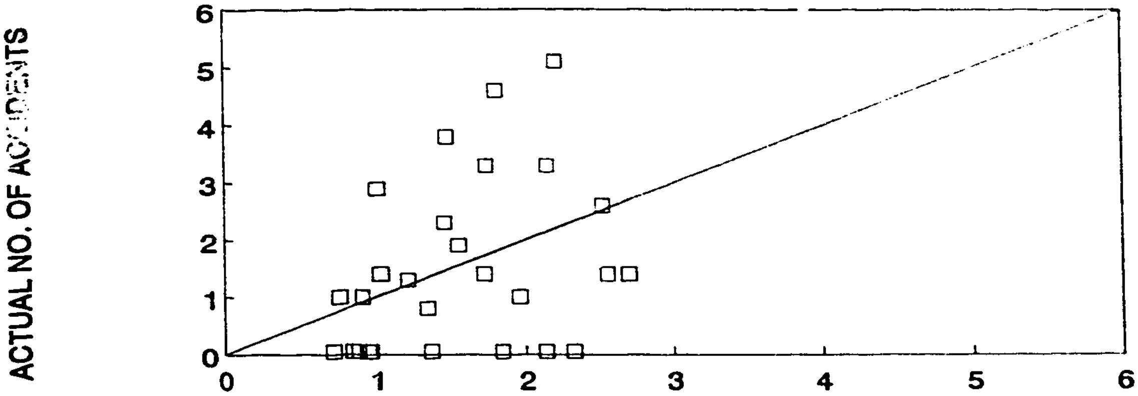

MODEL 226
ACTUAL NO. OF ACCIDENTS
0 2 4 6 8 10 12 14 16 18 20 22
PREDICTED NO. OF ACCIDENTS


MODEL 311
ACTUAL NO. OF ACCIDENTS
0
1
2
3
4
0
1
2
3
4
PREDICTED NO. OF ACCIDENTS


MODEL 312
ACTUAL NO. OF ACCIDE
0
1
2
3
4
5
6
7
8
9
10
11
0
1
2
3
4
5
6
7
8
9
10
11
PREDICTED NO. OF ACCIDENTS

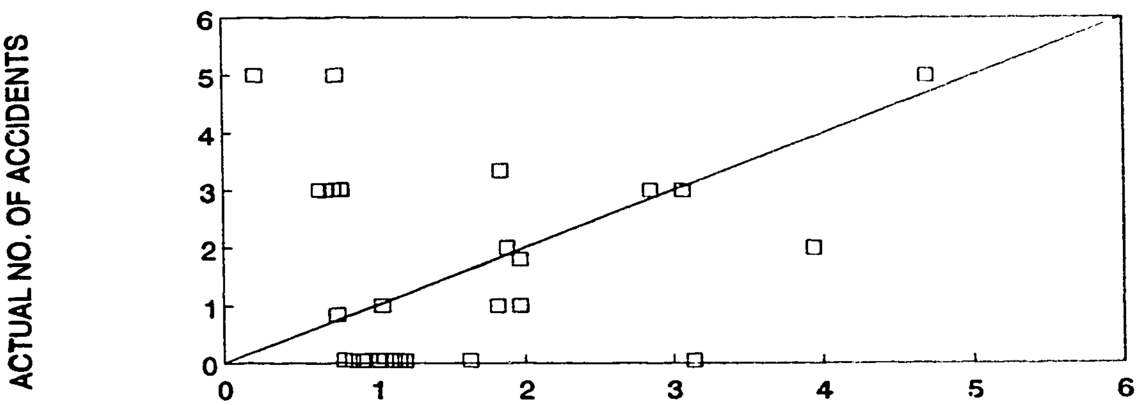

MODEL 313
ACTUAL NO. OF ACCIDENTS
0
1
2
3
4
5
6
7
0
1
2
3
4
5
6
7
PREDICTED NO. OF ACCIDENTS

MODEL 314



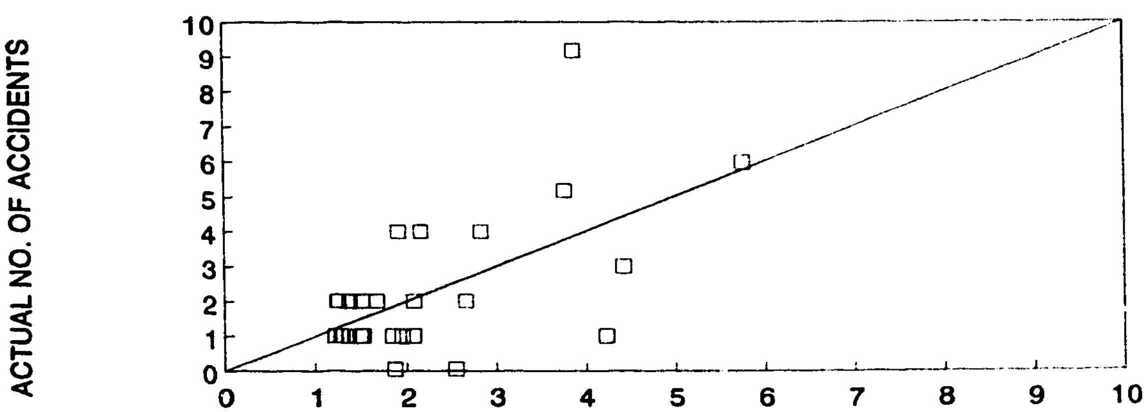
PREDICTED NO. OF ACCIDENTS

MODEL 512



PREDICTED NO. OF ACCIDENTS

MODEL 612



PREDICTED NO. OF ACCIDENTS

GLIM PROGRAM LISTING FOR MODEL 122

– SIMPLE PRODUCT–OF–FLOW MODEL

```
$units 26
$data site s v2 v3 v4 v6 v7 v8 d4 d6 d7 sa sb
$dinput 19
$yvariate s
$calc e1=v2+v3+v7+v8
$calc e2=v4+v6
$calc e=e1*e2
$mac m1 $calc %fv=%lp $endmac
$mac m2 $calc %dr=1 $endmac
$mac m3 $calc %va=%fv*(%fv+%d)/%d $endmac
$mac m4 $calc %di=2*(%yv*%log(%yv/%fv)-(%yv+%d)*%log((%yv+%d)/(%fv+%d)))
$endmac
$own m1 m2 m3 m4
$calc %lp=%log(%yv+1.5)
$calc %d=4
$fit e
$display m e r d
$calc %d=3 $fit.
$display m e r d
$calc %d=2 $fit.
$display m e r d
$calc %d=1 $fit.
$display m e r d
$calc %d=.75 $fit.
$display m e r d
$calc %d=.5 $fit.
$display m e r d
$calc %d=.25 $fit.
$display m e r d
$stop
```

GLIM PROGRAM LISTING FOR MODEL 216

– COMPLEX DELAY–BASED MODEL

```
$units 26
$data site s v2 v3 v4 v6 v7 v8 d4 d6 d7 sa sb
$dinput 19
$yvariate s
$calc e1=%log(v3)
$calc e2=%log(v4)
$calc e3=%log(v6)
$calc e4=%log(v7)
$calc r1=%log(v2)
$calc r2=%log(d4)
$calc r3=%log(d6)
$calc r4=%log(d7)
$mac m1 $calc %fv=%exp(%lp) $endmac
$mac m2 $calc %dr=1/%fv $endmac
$mac m3 $calc %va=%fv*(%fv+%d)/%d $endmac
$mac m4 $calc %di=2*(%yv*%log(%yv/%fv)-(%yv+%d)*%log((%yv+%d)/(%fv+%d)))
$endmac
$own m1 m2 m3 m4
$calc %lp=%log(%yv+1.5)
$calc %d=4
$fit r1+r2+r3+r4+e1+e2+e3+e4
$display m e r d
$calc %d=3 $fit.
$display m e r d
$calc %d=2 $fit.
$display m e r d
$calc %d=1 $fit.
$display m e r d
$calc %d=.75 $fit.
$display m e r d
$calc %d=.5 $fit.
$display m e r d
$calc %d=.25 $fit.
$display m e r d
$stop
```

Model Parameters – Models 311 to 314

Model No.	Parameters	Parameter Estimates	Standard Errors	Deviance
311	Model for Main Road Right Turn Movements:			21.65
1	%gm	−3.563E+00	4.157	
2	e1	−1.450E−01	3.170E−01	
3	r1	5.380E−01	6.710E−01	
312	Model for Minor Road Left Turn Movement:			14.96
1	%gm	2.820E−01	4.180E−01	
2	e2	5.960E−01	1.820E−01	
3	r2	−4.020E−01	2.130E−01	
313	Model for Minor Road Right Turn Movement:			13.07
1	%gm	−3.684E+00	9.370E−01	
2	e3	1.019E+00	2.470E−01	
3	r3	2.420E−01	3.240E−01	
314	Model for Main Road Left Turn Movement:			20.73
1	%gm	−1.161E+00	8.990E−01	
2	e4	2.870E−01	2.370E−01	
3	r4	3.890E−01	3.630E−01	

Model Parameters – Models 111–112, 121–124

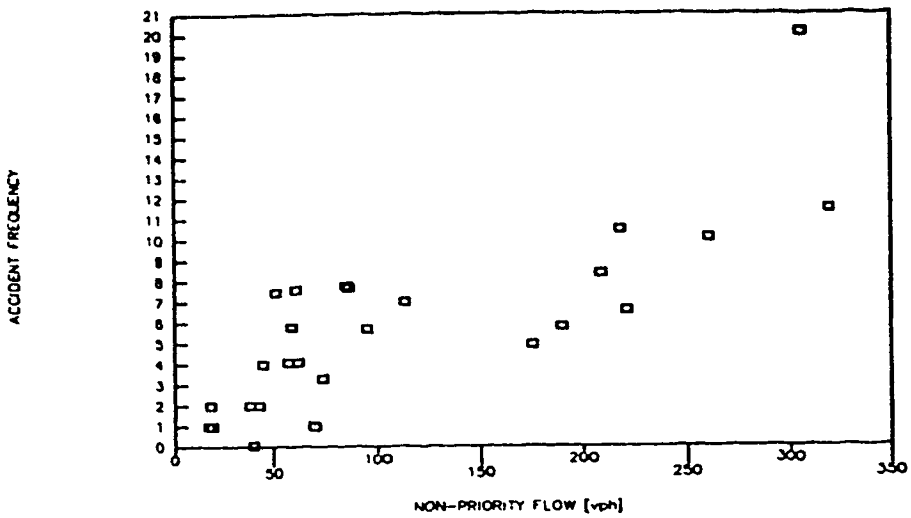
Model No.	Parameters	Parameter Estimates	Standard Errors	Deviance
111	Sum of Flow Model 1:			25.53
1	%gm	1.736E+00	1.837E+00	
2	e	2.820E−03	1.400E−03	
112	Sum of Flow Model 2:			17.03
1	%gm	1.604E+00	1.582E+00	
2	e1	−5.290E−03	5.570E−03	
3	e2	−5.160E−02	3.500E−02	
4	e3	6.600E−02	3.720E−02	
121	Product of Flow Model 1			10.79
1	%gm	1.715E+00	6.180E−01	
2	e1	−4.580E−06	5.200E−05	
3	e2	8.420E−05	9.590E−05	
4	e3	9.970E−05	5.540E−05	
122	Product of Flow Model 2			12.34
1	%gm	2.189E+00	6.690E−01	
2	e	4.890E−05	1.310E−05	
123	Product of Flow Model 3			11.09
1	%gm	−2.515E+00	1.903E+00	
2	e1	2.870E−01	2.710E−01	
3	e2	5.790E−01	1.340E−01	
124	Product of Flow Model 4			8.95
1	%gm	−1.561E+00	2.567E+00	
2	e1	3.543E+00	8.208E+00	
3	e2	−1.050E−02	1.440E−01	
4	e3	4.350E−01	2.870E−01	
5	e4	4.000E−01	2.950E−01	
6	e5	−3.162E+00	8.478E+00	
7	e6	−2.490E−01	1.806E+00	

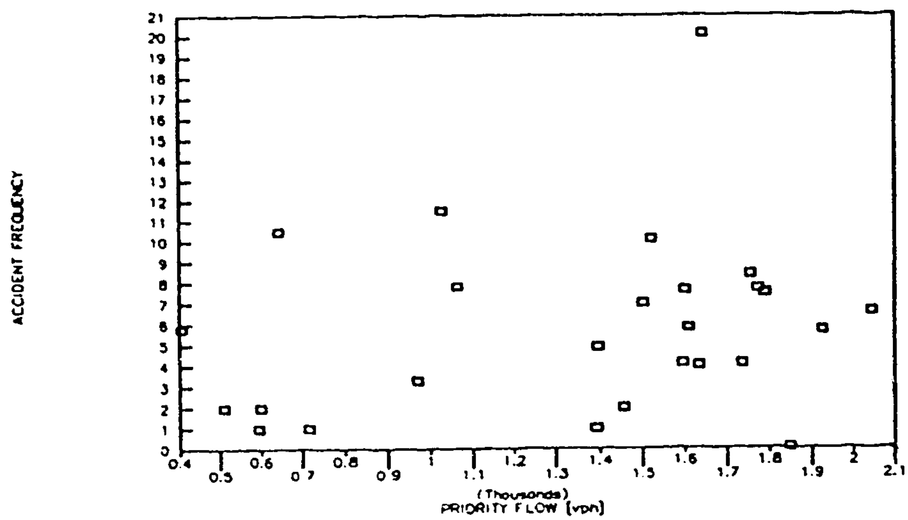
Model Parameters – Models 221 – 226

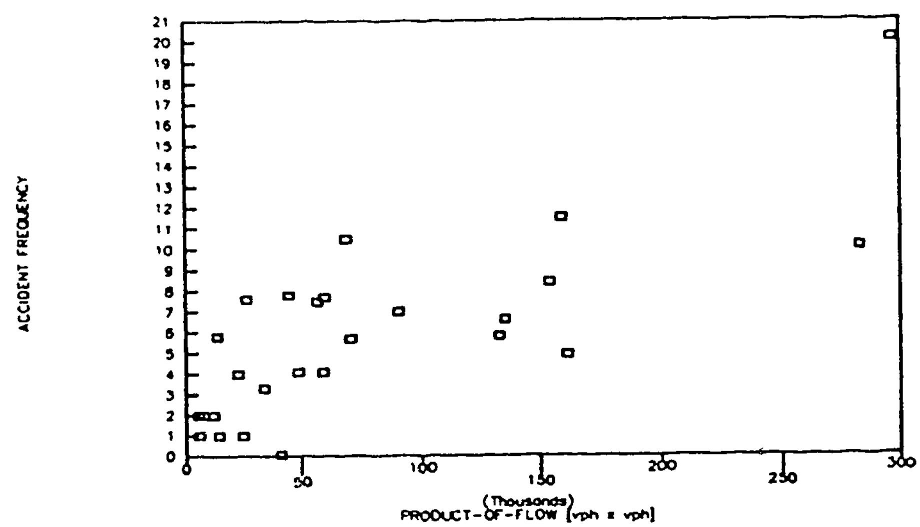
Model No.	Parameters	Parameter Estimates	Standard Errors	Deviance
221	Delay (simulated) Based Model 1:			12.04
1	%gm	1.533E+00	6.590E−01	
2	q3	2.780E−05	1.050E−04	
3	q4	−1.790E−03	3.000E−03	
4	q6	2.100E−02	1.600E−02	
5	q7	1.090E−02	8.290E−03	
222	Delay(simulated) Based Model 2:			19.10
1	%gm	4.560E+00	7.850E−01	
2	e	7.330E−13	3.880E−13	
223	Delay (simulated) Based Model 3:			11.38
1	%gm	−9.140E−01	1.245E+00	
2	q1	3.850E−02	2.180E−01	
3	q2	9.890E−03	6.800E−02	
4	q3	2.280E−01	2.800E−01	
5	q4	2.410E−01	2.230E−01	
224	Delay (simulated) Based Model 4:			9.81
1	%gm	−2.707E+00	2.676E+00	
2	e	1.720E−01	4.520E−02	
3	r1	4.520E−01	4.750E−01	
4	r2	−1.290E−01	1.400E−01	
5	r3	−3.540E−01	4.870E−01	
6	r4	1.530E−01	4.930E−01	
215	Delay (simulated) Based Model 5:			9.22
21	%gm	−9.950E−01	7.560E−01	
2	r	2.950E−02	6.020E−02	
3	e1	4.220E−02	2.630E−01	
4	e2	−2.830E−02	1.320E−01	
5	e3	3.880E−01	2.940E−01	
6	e4	3.070E−01	2.500E−01	
226	Delay (simulated) Based Model 6:			8.86
1	%gm	−3.120E+00	3.062E+00	
2	r1	4.250E−01	5.440E−01	
3	r2	1.520E−03	2.140E−01	
4	r3	−1.160E−01	5.420E−01	
5	r4	−1.350E−01	5.660E−01	
6	e1	−5.940E−03	2.970E−01	
7	e2	−2.050E−03	1.990E−01	
8	e3	3.970E−01	3.260E−01	
9	e4	3.490E−01	2.760E−01	

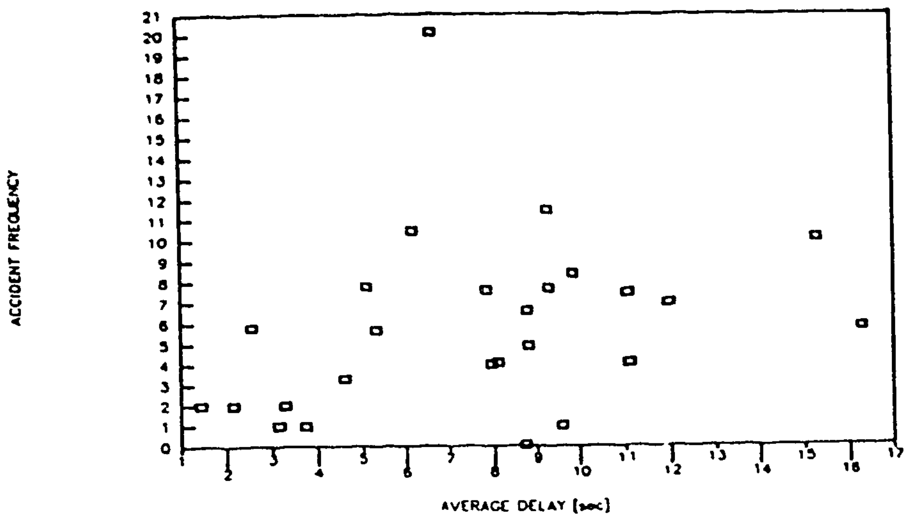
APPENDIX C

Correlation Plots of Parameters in the Best Models

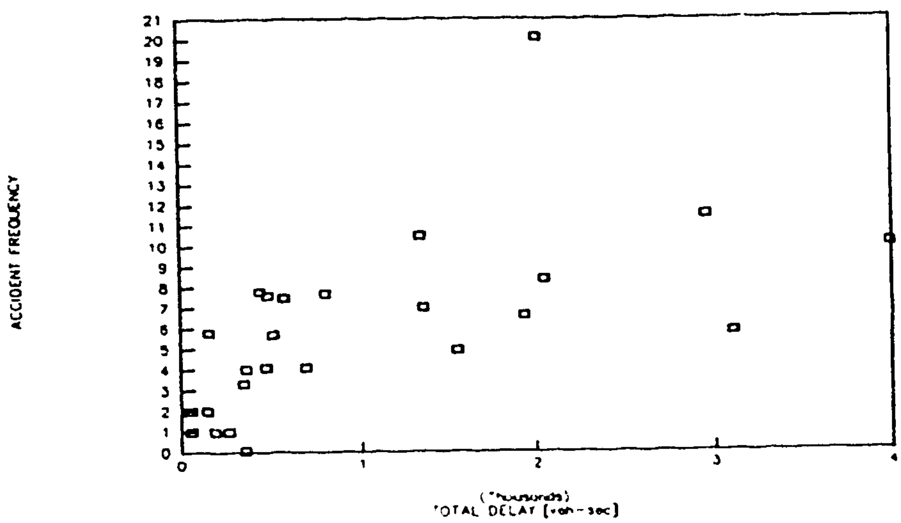

ACCIDENT FREQUENCY
NON-PRIORITY FLOW [vph]


ACCIDENT FREQUENCY
(Thousands)
PRIORITY FLOW [vph]

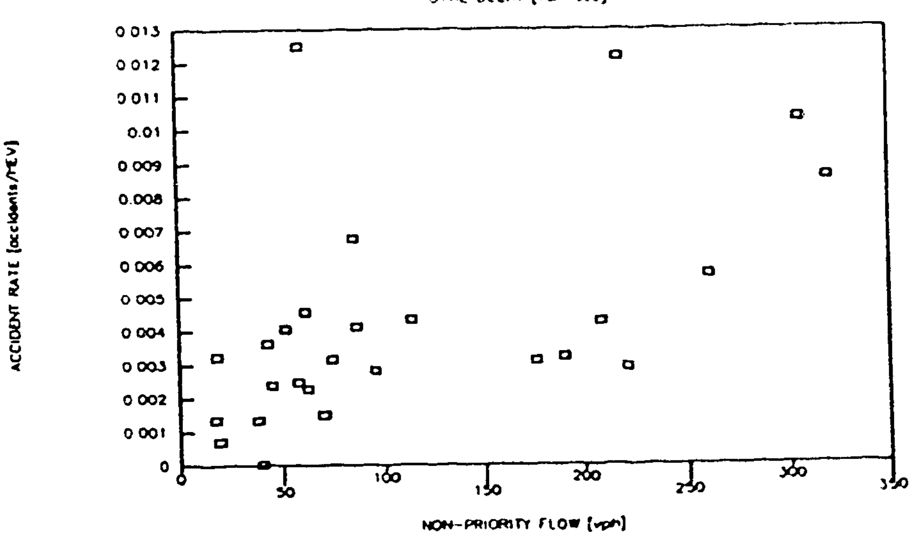

ACCIDENT FREQUENCY
(Thousands)
PRODUCT-OF-FLOW [vph x vph]



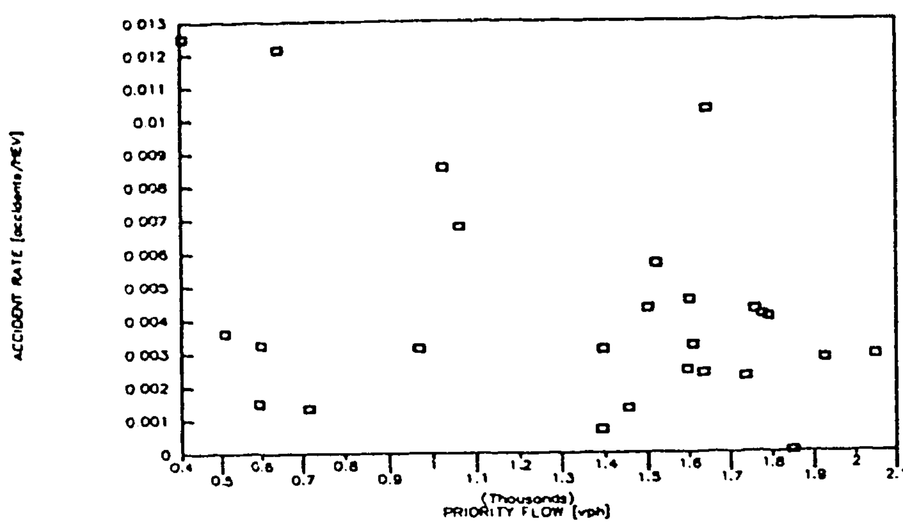
ACCIDENT FREQUENCY
AVERAGE DELAY (sec)

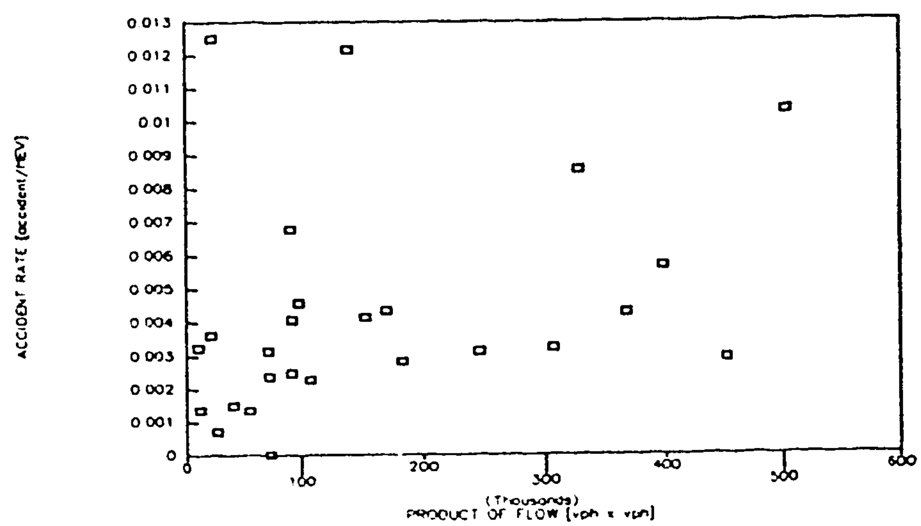


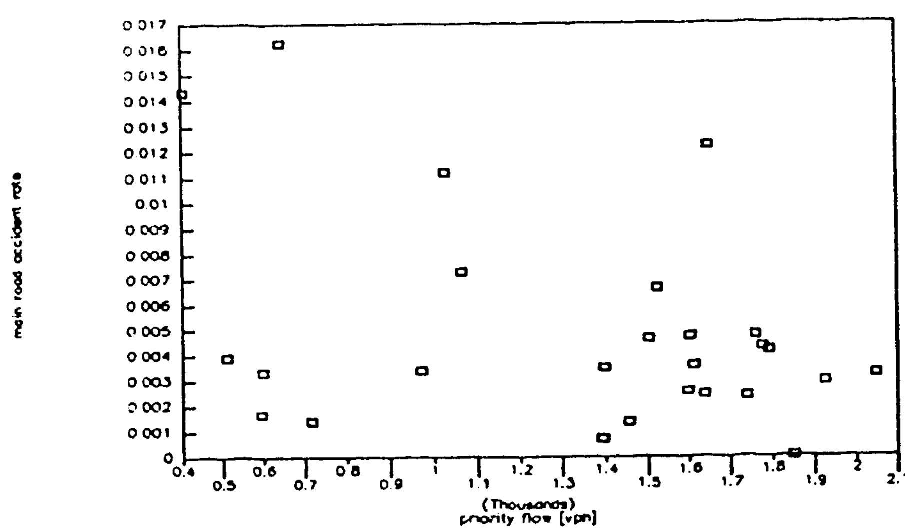
ACCIDENT FREQUENCY
(Thousands)
TOTAL DELAY (veh-sec)

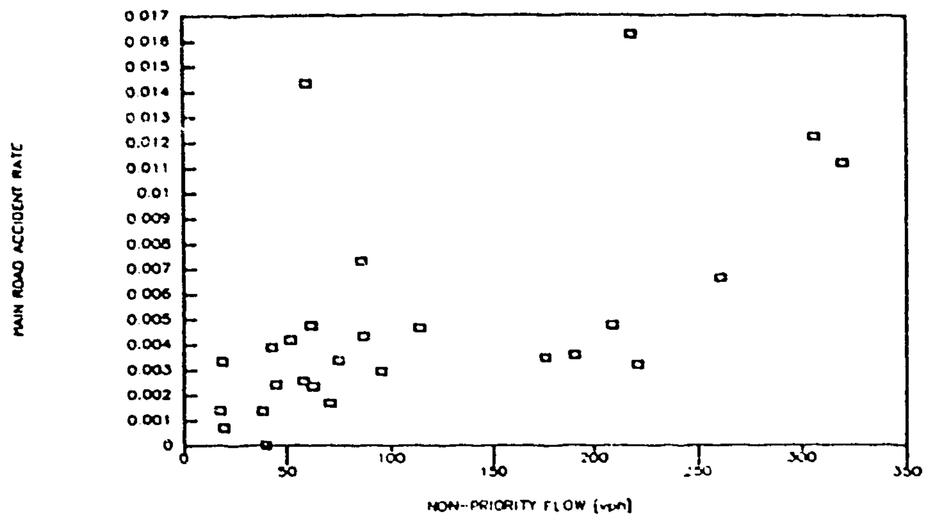


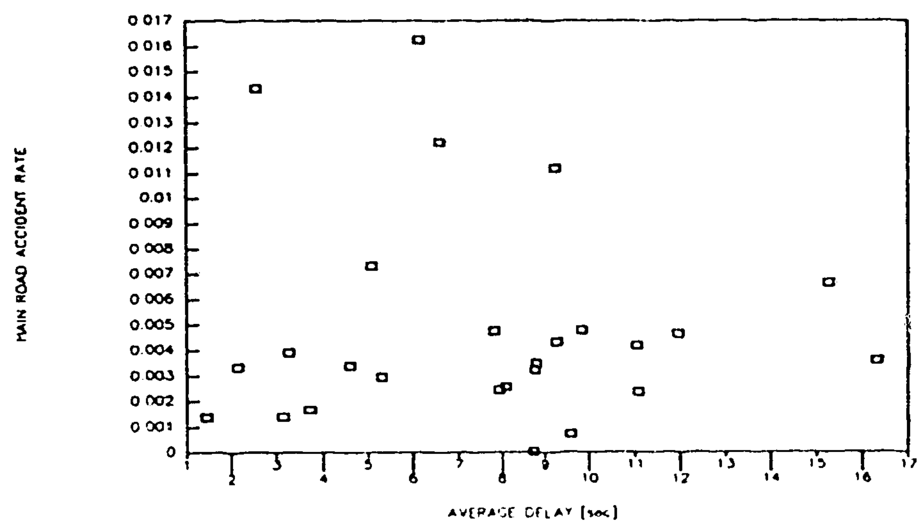
ACCIDENT RATE (accidents/MEV)
NON-PRIORITY FLOW (vph)

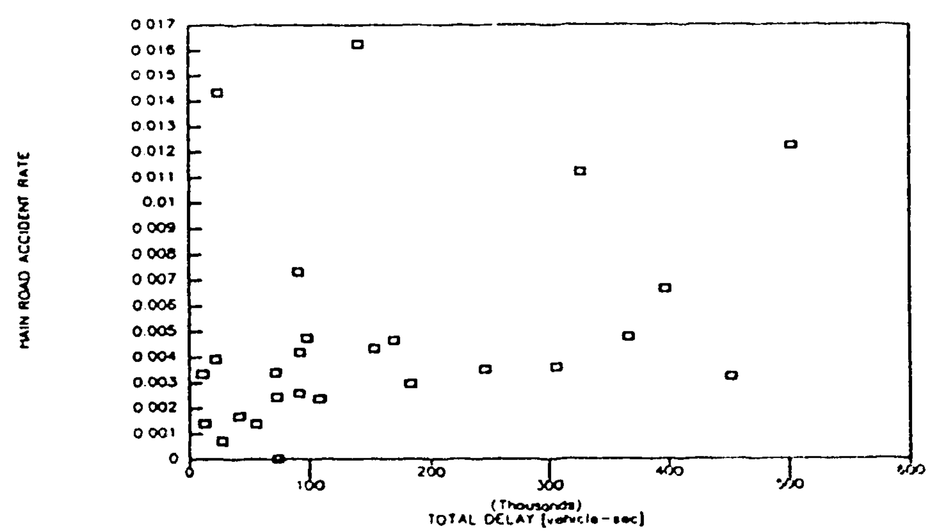

ACCIDENT RATE [accidents/MEV]
(Thousands)
PRIORITY FLOW [vph]

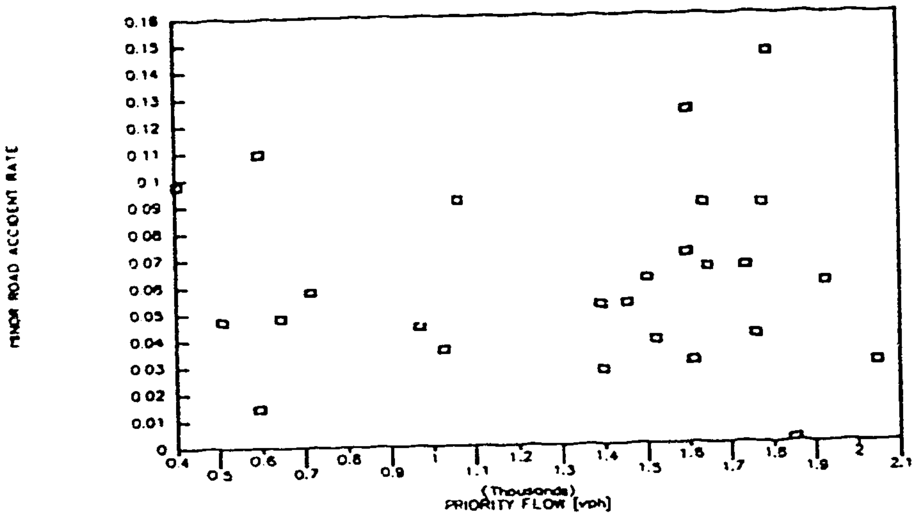

ACCIDENT RATE [accident/MEV]
(Thousands)
PRODUCT OF FLOW [vph x vph]

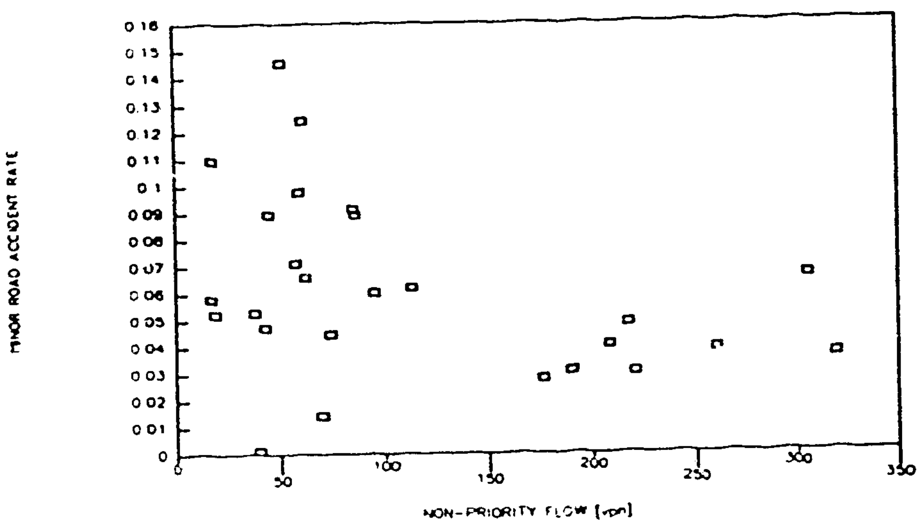

main road accident rate
(Thousands)
priority flow [vph]

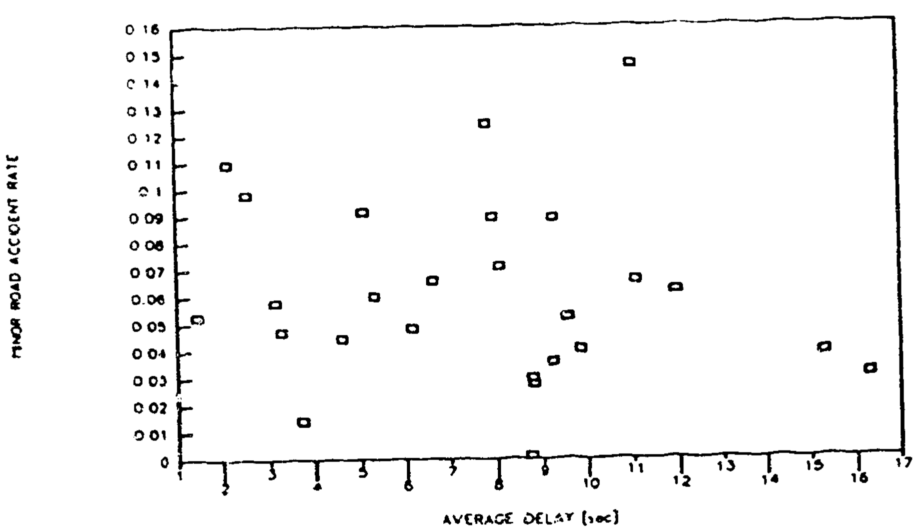

MAIN ROAD ACCIDENT RATE
NON-PRIORITY FLOW (vph)

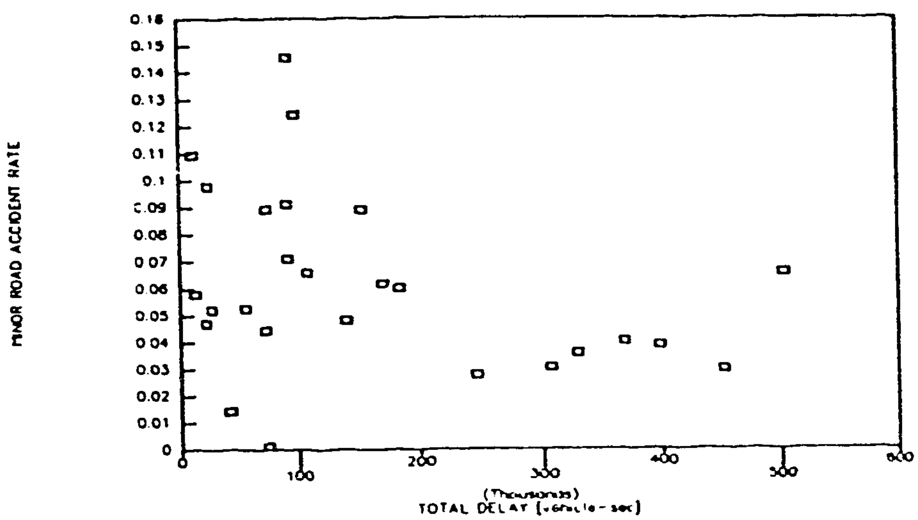

MAIN ROAD ACCIDENT RATE
AVERAGE DELAY (sec)


MAIN ROAD ACCIDENT RATE
(Thousands)
TOTAL DELAY (vehicle-sec)


MINOR ROAD ACCIDENT RATE
(Thousands)
PRIORITY FLOW [vph]


MINOR ROAD ACCIDENT RATE
NON-PRIORITY FLOW [vph]


MINOR ROAD ACCIDENT RATE
AVERAGE DELAY [sec]


MINOR ROAD ACCIDENT RATE
(Thousands)
TOTAL DELAY (vehicle-sec)

APPENDIX D

Traffic Safety at Intersection Without Traffic Signals

Paper presented at the 1993
Annual Conference on the
Institute of Transportation Engineers
District 7 - Canada in Edmonton.
Compendium of Technical Papers.
p.169-186.

Traffic Safety at Intersections Without Traffic Signals

Paul Chan* and Stan Teply

BACKGROUND

Considerable efforts have been directed towards the monitoring of roadway safety through accident record-keeping. Statistical data are used in determining priorities for the allocation of resources for safety improvements. This approach to traffic safety may be termed "reactive". On the other hand, methods which identify locations with higher accident potential in advance on the basis of indirect measures, may allow a timely action before the accident issue becomes critical, i.e. they are proactive.

In both instances, the success of a safety improvement program relies heavily on the technique used in identifying locations with higher accident potential. Accident statistics are frequently used directly to determine locations that are "more dangerous than average". The most often employed parameters related to safety are accident frequencies and accident rates. They reflect two different aspects of the actual safety situation and safety improvement priority lists developed on the basis of each of them often do not agree well with each other.

A proactive traffic safety approach begins with a review of appropriate measures of safety. The difficulty lies in the fact, that although the number of accidents and the associated human suffering and economic loss is enormous, the number of comparable accidents is statistically small. The intricacies of small-number statistics require a scientific approach to build a sound knowledge foundation for safety engineering. This paper and the underlying research are an attempt to contribute toward that goal.

A good accident prediction model can significantly enhance the understanding of traffic operations at a location and the risk to which drivers are exposed. Most of the models developed in the past use some combinations of the sums or products of traffic volumes as independent parameters.

Accidents are events with a high degree of randomness. Human factor is generally considered the most significant of all contributing issues but also the least predictable. Driving a vehicle is a complex task involving perception, information processing, evaluation, decision-making, attitudes, emotions and skills. In-depth accident studies (Sabey and Staughten 1975, 1983) by "on-the-spot" investigation teams indicated that the human element was dominant in 95% of roadway accidents. The other system components are usually grouped in two categories: vehicle and environment.

Since human factors are major contributors to accidents, it follows that a model which includes some representation of driver behavior, or a parameter which influences human behavior, would have a better chance of correctly estimating the degree of safety. Based on this hypothesis, we have decided to investigate the impact of delay on safety, with a focus on the location ranking. In order to reduce the number of influencing variables, the scope of this research was limited to T-intersections, the simplest intersection type.

ACCIDENT DATA

A total of 3795 accident records from 1985 to 1988 for 429 priority-ruled intersections were obtained from the City of Edmonton Accident Inventory File. Of the 429 intersections selected, 129 are 3-legged and 300 are 4-legged.

In general, accident records in the City of Edmonton Accident Database appeared reliable and valid in location, time, date and number of vehicles involved in accidents. However, information on driving lanes and road surface conditions may be of questionable accuracy.

General Trends

Overall, accident frequencies were low with an average of 8.8 accidents/4 years from 1985 to 1989 for the 429 intersections. However, the average was 10.2 accidents/4 years for 4-legged intersections and 5.7 accidents/4 years for 3-legged intersections. There was also a significant degree of variability in the occurrences of accidents, probably due to the low accident frequencies.

Other than the low average accident frequencies, accidents at 3-legged and 4-legged intersections were very similar in nature in yearly, monthly, daily and hourly trends; and in types of pre-crash maneuvres and human actions at time of collision, as well as severity (Figures 1.a to 1.f).

The number of accidents in a year had remained relatively stable with less than 15% overall annual fluctuations from 1985 to 1989. Within a year, accidents peaked in November and February, and were also high in June. Analysis on weekday daytime accidents for the 429 intersections indicated that, on the average, there were 3.3 accidents/day during 5 winter months and 2.3 accidents/day during 5 summer months in the database.

At a daily level, accident occurrences generally peaked on Friday and during the day at the peak hours. The p.m. peak had the highest frequency of accidents.

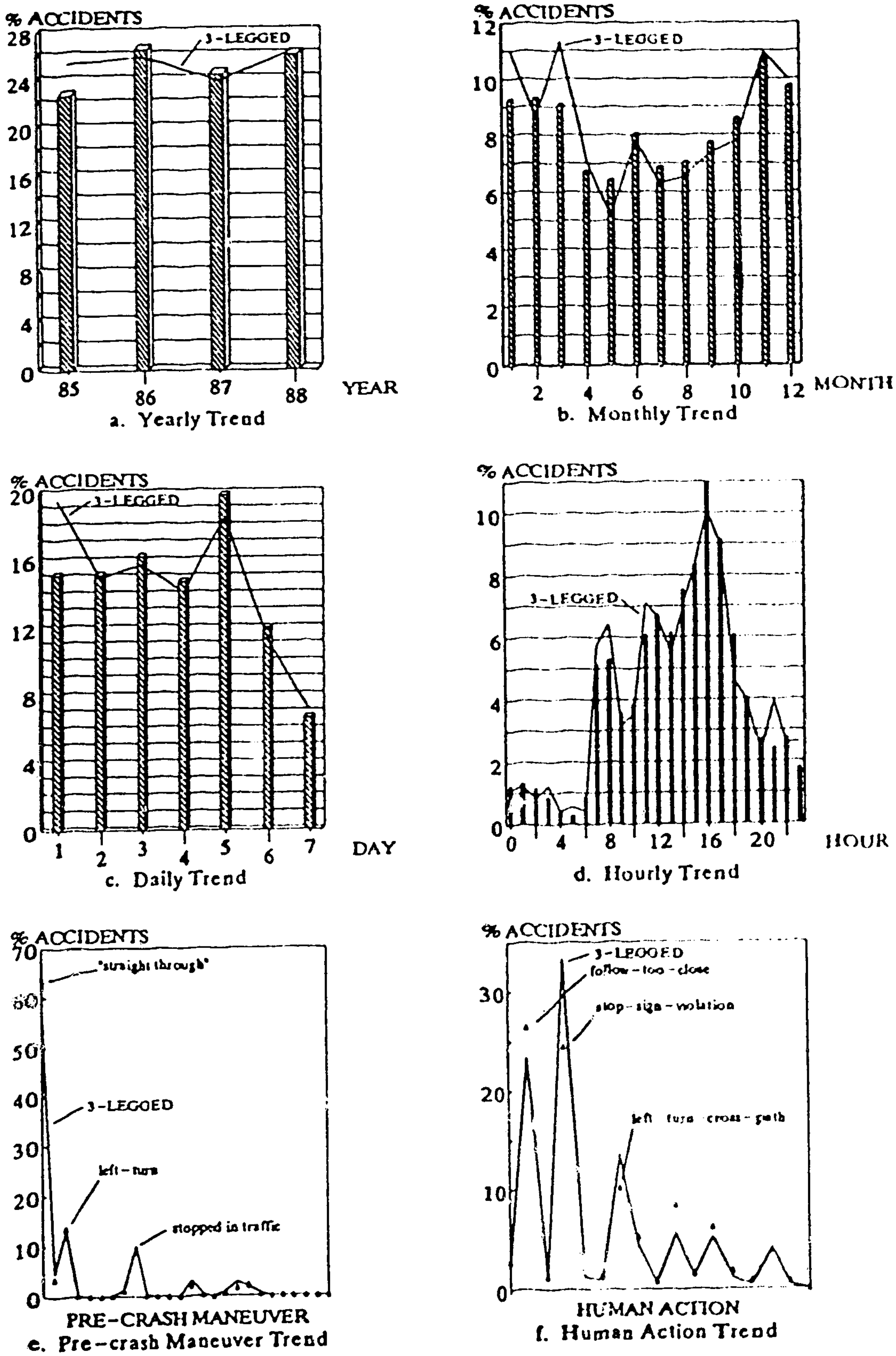


FIGURE 1 ACCIDENT TRENDS AT 429 INTERSECTIONS WITHOUT TRAFFIC SIGNALS

Travel Condition Modules

Characteristics for accidents that occurred in the weekday/weekend and in daytime/nightime were found to be distinctively different. Daytime accidents mostly occurred (84% as compared to the average of 81%) on the weekdays, whereas a significant amount of nightime accidents (42% as compared to the average of 19%) occurred on the weekend. On the weekends, 26% of the accidents occurred at night.

It appears that there were a number of different time periods in which accidents were significantly different. The distinctive difference in accident trends among the different patterns may be attributed to the fact that the underlying conditions within each module were significantly different from each others.

These time periods can be considered as unique " travel condition modules " in which each module consisted of a different combination of driver, vehicle and environment factors and resulted in different prevailing conditions from a traffic operation and safety standpoint.

ACCIDENT PREDICTION MODELLING

Risk and Exposure Concept

Formulation of the accident prediction model was based on the risk and exposure concept. Hauer (1982) made a clear explanation of the concept of risk and exposure by relating it to the philosophy of chance:

> "A unit of exposure corresponds to a trial. The result of such a trial is the occurrence or non-occurrence of an accident (by type, severity, etc). The chance set up is the transportation system (physical facilities, users, and environment) which is being examined and risk is the probability (chance) of accident occurrences in a trial".

The following relationship represents the fundamental structure of a typical risk and exposure accident prediction model:

$$\text{SAFETY} = \text{RISK} \times \text{EXPOSURE}$$

or

$$S = R \times E,$$

where

- S = Measure of intersection safety during a period of time
- R = Measure of the average risk presented to drivers during the same period of time
- E = Exposure, i.e. the number of drivers that are presented the risk during the same period of time

Data Collection

Based on the travel condition modules identified from the accident and traffic patterns, modelling efforts were concentrated on the weekday off-peak time period to minimize variability in travel conditions. The selected module encompassed the period between 1 p.m. and 4 p.m. from Monday to Thursday. Over 38% of the total accidents occurred within this module. With relatively steady traffic volumes and accident trends, as well as the long duration of the module, this module was ideal for data collection.

As accident patterns at 3-legged and 4-legged intersections were found to be highly comparable. T-intersections were therefore chosen for the study to maintain data collection and interpretation efforts at a manageable level. Sites of varying accident frequencies were selected to avoid over-representation of low accident frequency locations.

Delay data, together with volume, headway and site attributes information, were collected by using a video camera in January and February 1990, at 26 priority-ruled T-intersections in Edmonton. The surveys were limited to the off-peak periods on weekdays with normal driving conditions. Figure 2 summarizes the survey conditions and the accident, delay and volume data collected for the 26 study locations.

Delay and volume data collected were transferred to a computer database with the assistance of the TDIP program (Kyte and Boesen 1989). The delay data collected had a high degree of variability. The KNOSIMO simulation model (Grossmann 1988) was employed to generate additional delay information to complement the measured delay data In applying the KNOSIMO model, modifications were made at the data input level so that the program was applicable for urban 4 lane conditions (Chan and Teply 1991). Figures 3 and 4 compare the intersection conditions assumed in the KNOSIMO program and at the 26 study locations.

Analytical Approach

The within site distribution of accidents at a particular location follows the Poisson process. The between site variation may be described by a gamma distribution. The resulting sampling distribution over all sites is a negative binomial distribution (Maycock and Hall 1984).

Since accident data have a negative binomial distribution, typical classical regression models, which were based on a normality assumption, were not suitable for accident data analysis. The GLIM statistical analysis package (Payne 1987) can handle analysis for data sets with negative binomial distribution and was therefore used in this research.

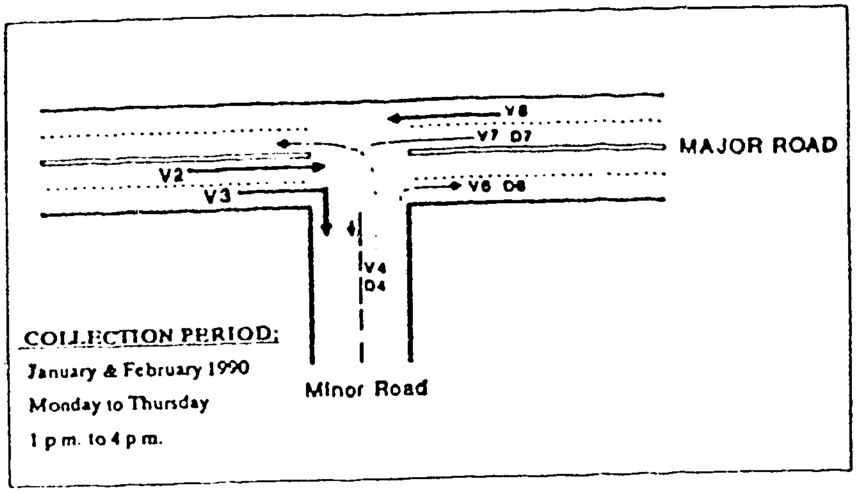


FIGURE 2 SURVEY CONDITIONS AND DATA COLLECTED

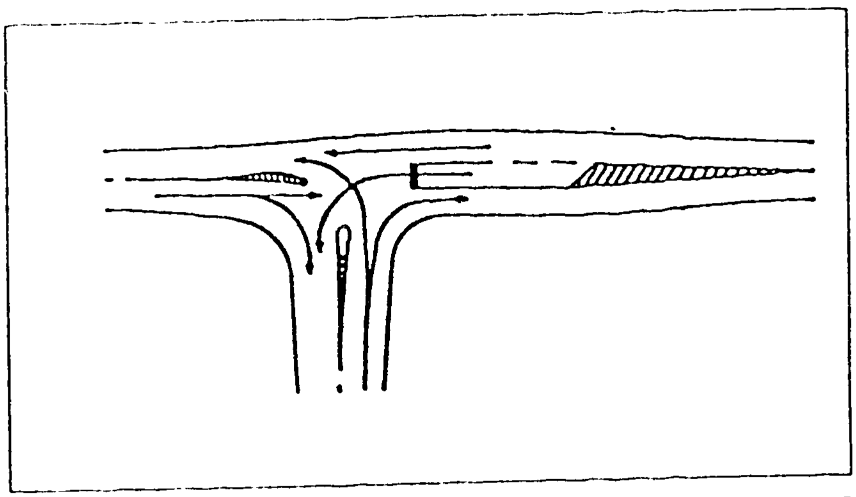


FIGURE 3 INTERSECTION CONDITIONS ASSUMED IN KNOSIMO

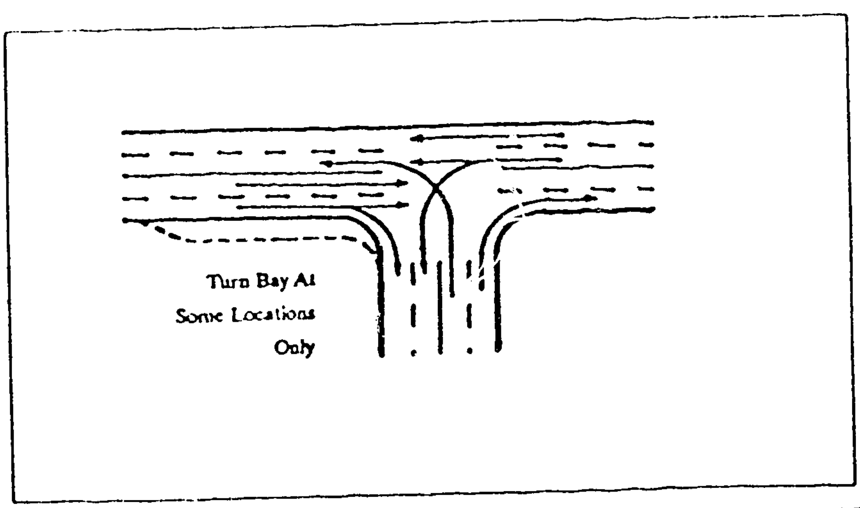


FIGURE 4 INTERSECTION CONDITIONS FOR THE STUDY LOCATIONS

Model Formulation

Excluding multiple vehicle accidents and single vehicle accidents, there were 13 major types of accidents that can be associated with one or more non-priority movements at intersections without traffic signals (Figure 5).

There were altogether 129 accidents for the 26 study locations from 1985 to 1989. The most frequent accident types were right-angle collisions, rear-end collisions and left-turn-cross-path accidents. However, the frequencies for most accident types are low at several intersections.

Attempts to further aggregate the accidents by non-priority movement types were only mildly successful, as it was difficult to categorize some of the accidents by non-priority movement types.

Model Development

Four groups of accident prediction models were developed by using the GLIM analysis package:

a. Disaggregate accident prediction models for rear-end and right-angle collisions

b. Disaggregate accident prediction models for each non-priority traffic movements

c. Aggregate accident prediction models based on sum-of-flow and product-of-flow approaches

d. Aggregate accident prediction models based on different combinations of traffic volumes and delays parameters

Table 1 summarizes the parameters used in the various accident prediction models. A total of 26 models were developed. Table 2 summarizes some selected examples of model form of the 26 models.

MODEL COMPARISON

Comparison on accident trend prediction was made on the four groups of accident prediction models developed. It was found that models formulated at a more aggregate level provide better trend prediction results than the disaggregate models. The best accident prediction models were the product-of-flow models and the delay-based models.

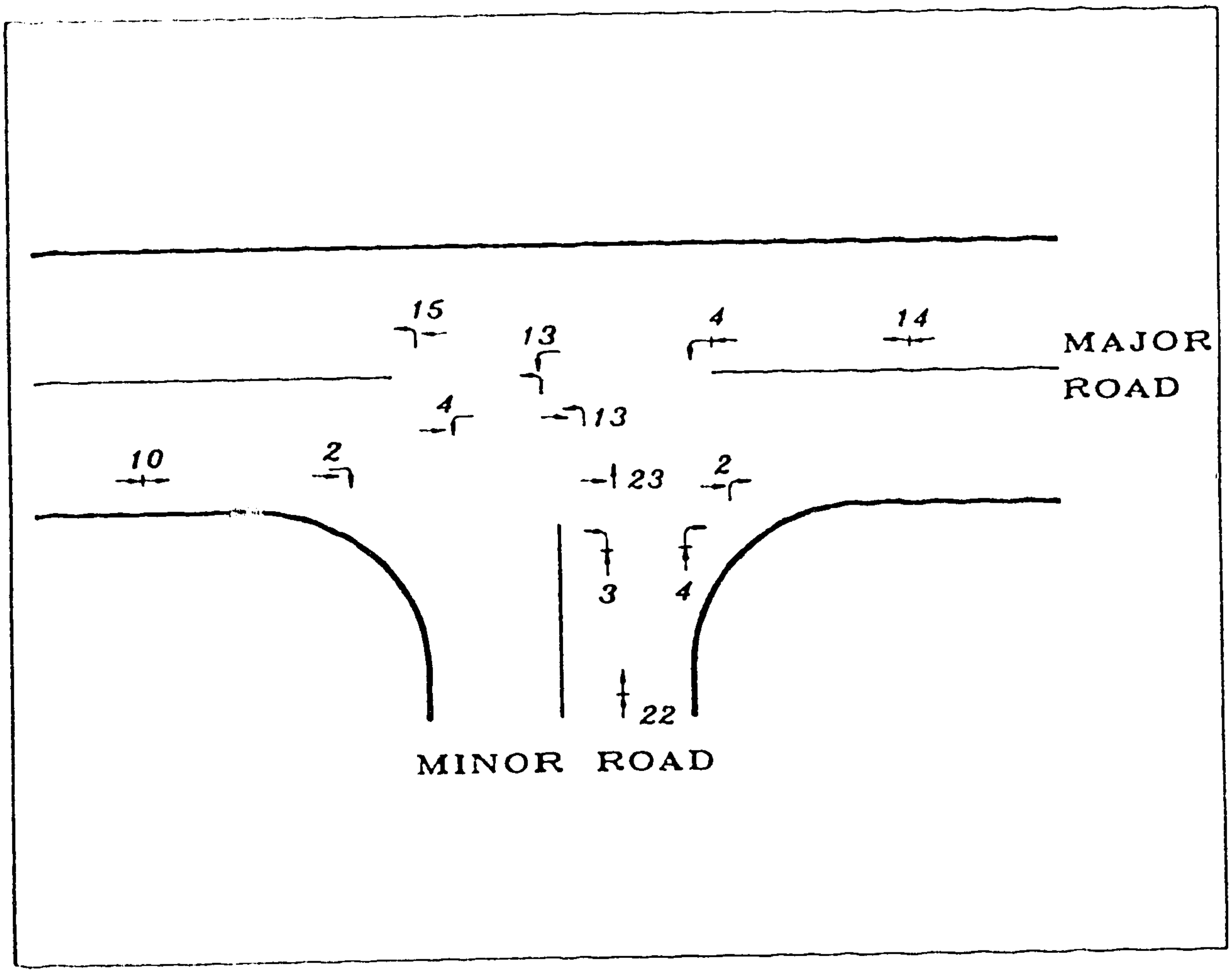


FIGURE 5 13 ACCIDENT TYPES AND FREQUENCIES AT THE 26 STUDY LOCATIONS

TABLE 1 ACCIDENT PREDICTION MODEL PARAMETERS

Parameter	Description
S	Overall safety (Objective safety) at an intersection = Expected no. of accidents over a 4 years period
S rc	No. of rear-end collisions over 4 years
S ra	No. of right-angle accidents over 4 years
S3	No. of accidents (over 4 years) related to main road right-turn movements
S4	No. of accidents (over 4 years) related to minor road left-turn movements
S6	No. of accidents (over 4 years) related to minor road right-turn movements
S7	No. of accidents (over 4 years) related to major road left-turn movements
V2	Main road through volume from left of minor road
V3	Main road right-turn volume
V4	Minor road left-turn volume
V6	Minor road right-turn volume
V7	Main road left-turn volume
V8	Main road through volume from right of minor road
D4	Average stopped delay [illegible]or minor road left-turn flows
D6	Average stopped delay for minor road right-turn flows
D7	Average stopped delay for main road left-turn flow

TABLE 2 SELECTED EXAMPLES OF MODEL FORMS

Sum-of-F[illegible]ent Prediction Model:

$$S = a_0 + a_1 \left[V_2+V_3+V_4+V_6+V_7+V_8\right]$$

Product-of-Flow Accident Prediction Model:

$$S = e^{a0} \left(V_2+V_3+V_7+V_8\right)^{a1} \left(V_4+V_6\right)^{a2}$$

Delay-based Accident Prediction Model:

$$S = e^{b0} \left(V_2^{b1} * V_3^{c1}\right) \left(D_4^{b2} * V_4^{c2}\right) \left(D_6^{b3} * V_6^{c3}\right) \left(D_7^{b4} * V_7^{c4}\right)$$

Rear-end and Right-angle Accidents

Accident prediction models for rear-end accidents were found to have poor predictability, which was probably because the model could not reflect the actual cause and effect relationship present in rear-end accidents. Models for right-angle accidents had better predictability.

Disaggregate Models by Non-priority Movements

Disaggregate models for each of the non-priority movements were generally poor in predictability, mainly due to low accident frequencies within the disaggregate models. The model with the best predictability was the minor road left-turn accidents model.

Sum-of-flow Models

Models based on parameters comprised of the sum of entering flows had poor predictability. As illustrated in Figure 6, there was no obervable trend in correlating accident frequencies to the sum of the traffic volumes entering the intersection.

Product-of-flow Models

Models based on parameters comprised of the products of conflicting flows had much better prediction power. Figure 7 illustrates the result of accident trend prediction for a simplistic basic product-of-flow model. The trend prediction power of the simple model was only slightly lower than that provided by much more complicated product-of-flow models. It appeared that an added level of model complexity did not result in a corresponding improvement in predictability.

Delay-based Models

Models based on both traffic volume and delay parameters were slightly better in predicting accidents than those based on product-of-flow parameters. Figure 8 shows the accident trend prediction result of a full delay-based model with several model parameters.

It did not appear to be significantly different when simulated delay was used instead of measured delay for modelling.

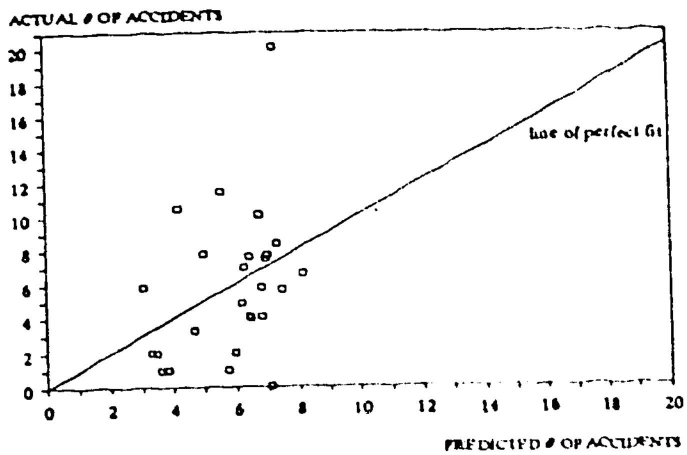


FIGURE 6 PREDICTION RESULTS OF SUM-OF-FLOW MODEL

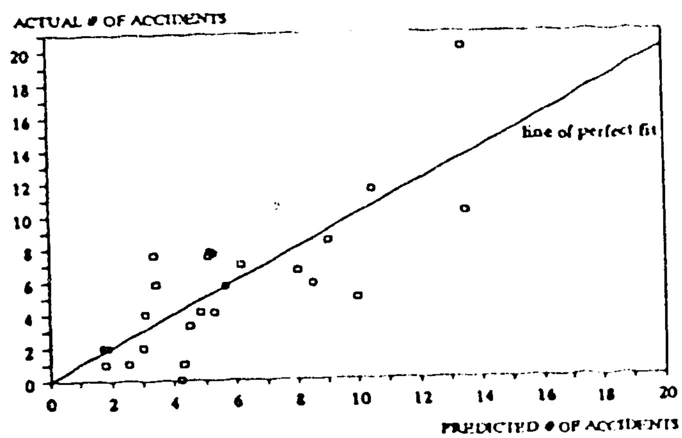


FIGURE 7 PREDICTION RESULTS OF PRODUCT-OF-FLOW MODEL

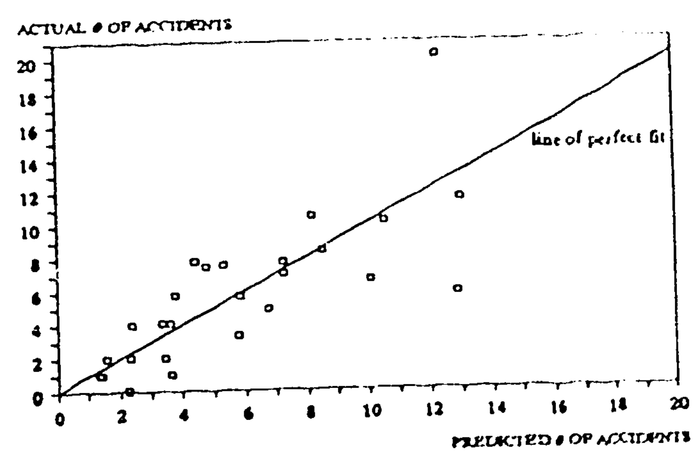


FIGURE 8 PREDICTION RESULTS OF DELAY-BASED MODEL

FINDINGS

Level of Aggregation

For the simpler accident prediction models, comparison of model parameters to their standard errors indicated that the parameters were significant. However, parameters in the more complicated models were found to be not significant in similar comparisions. Based on this observation, the more complicated models should not be considered more favourable than the simpler models, even though they did give marginally better accident prediction results. The low accident frequencies at priority-ruled intersections apparently limited the practical level of aggregation for modelling.

Product-of-flow Model vs. Total-delay Model

The delay-based model was more compact and provided a better defined accident trend prediction than the product-of-flow model. In both models, accident frequencies increased at a slower rate than the total delays or the product-of-flow (Figures 9 and 10).

The square root of both product-of-flow and the total delay provided the best accident prediction results (Figures 11 and 12).

Other Accident Parameters

Further analysis on the model parameters of the product-of-flow and total-delay models indicated that accident and risk parameters, when being examined individually, did not give clear trend indications in most cases. The lack of trend indication was an evidence that other underlying factors may be present.

Some of the major observed trends were:

a. Accident frequency correlated better with non-priority flow than major road flow.

b. Overall accident rate parameters which use either product-of-flow or total delay as the exposure unit were found to be poor as they were not sensitive to any other measures.

c. Correlation of various accident rates to either the major road flow or the non-priority flow were generally inconclusive.

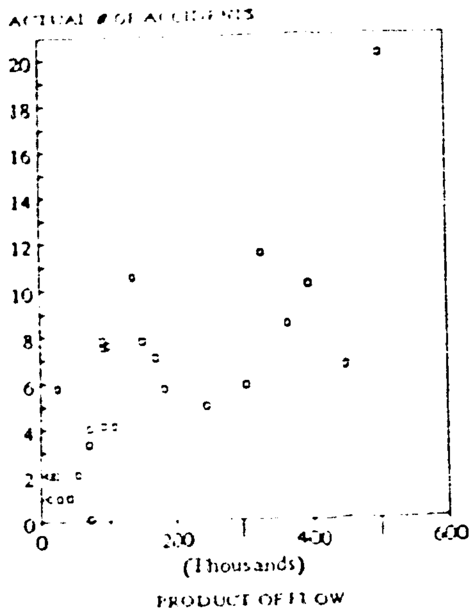


FIGURE 9 PRODUCT--OF--FLOW MODEL

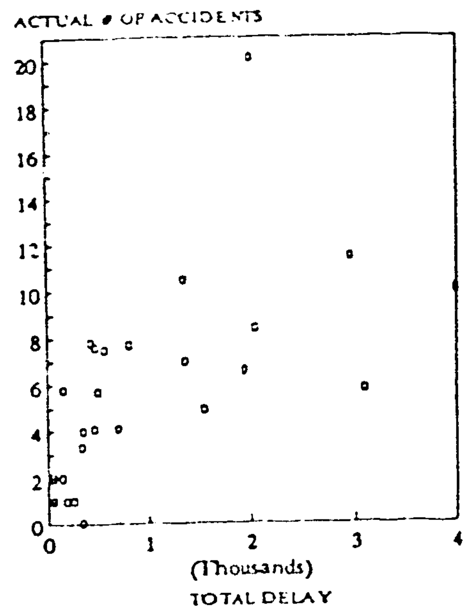


FIGURE 10 TOTAL DELAY MODEL

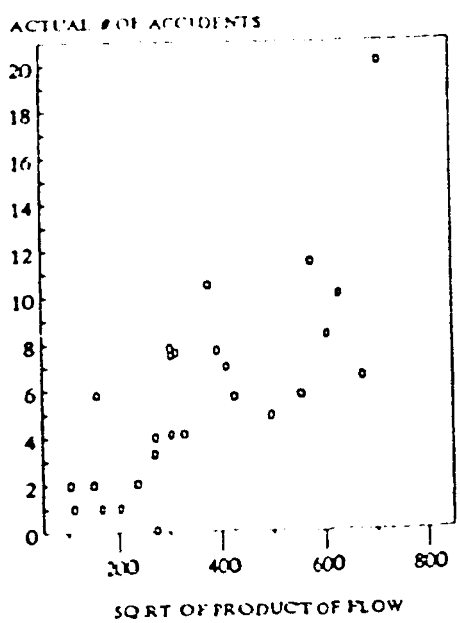


FIGURE 11 SQ RT. OF PROD. OF FLOW MODEL

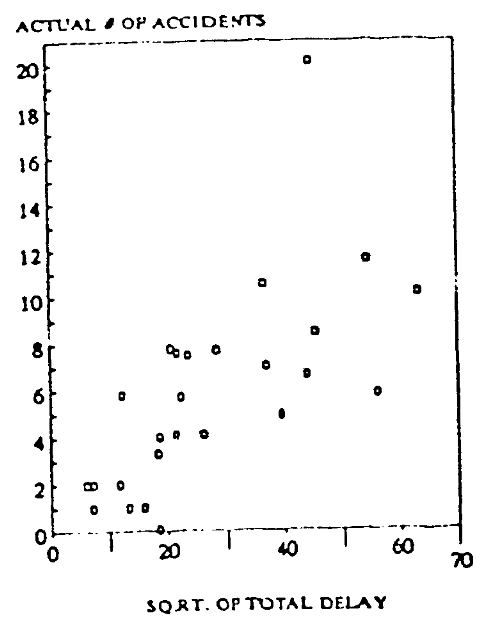


FIGURE 12 SQ.RT. OF TOTAL DELAY MODEL

Most accident parameters are robust measures that are affected by a number of factors. A single parameter, when used on its own, was not sufficient to explain the whole relationship. It confirmed that extra caution should be applied in using a single accident parameter to determine the safety of an intersection

CONCLUSIONS

This pilot research project had demonstrated that a highly disaggregated level of modelling was not suitable for predicting accidents at priority-ruled intersections. It showed that the sum of entering flows was not a good basis to determine accident rates and the priorities among accident locations.

Two relatively simple models at aggregate level had been identified as sound and practical: the product-of-flow model and total-delay model. Although both performed well and the difference in accident predictive power was small, the total-delay model was slightly better and showed more compact prediction trends.

Since only T-intersections with a limited range of weekday off-peak traffic and accident data were investigated, caution in the applications of specific results was advised. Nevertheless, the models can serve as comparative tools to indicate safety expectancy at intersections without traffic signals.

ACKNOWLEDGEMENT

The authors would like to express their gratitude to Professor W. Brilon and Dr. M. Grossmann of the Ruhr University in Bochum for providing the KNOSIMO intersection program and advice. Ms. K. Bustin assisted with the translation of the KNOSIMO manual. Dr. Fung of the University of Calgary helped with the GLIM computer package. The cooperation of Mr. R. Strynadka and Mr. B. Shacker of the City of Edmonton with regard to accident data is gratefully acknowledged. The Natural Sciences and Engineering Research Council of Canada contributed to this project in a form of an operational research grant.

REFERENCES

CHAN, P. AND TEPLY, S. Simulation of Multilane Stop-controlled T-intersections by KNOSIMO in Canada. Presented in the Second International Conference on Intersections Without Traffic Signals. 1991.

GROSSMANN, M. KNOSIMO - A Practicable Simulation Model for Unsignalized Intersections. Intersections Without Traffic Signals, Proceedings of an International Workshop 16-18 March, 1988 in Bochum, West Germany.

HAUER, E. Traffic Conflicts and Exposure. Accident Analysis and Prevention, Volume 14, No. 5, pp 359-364, 1982.

KYTE, M. AND BOESEN, A. Traffic Data Input Program Program Documentation and User's Manual Version 2.0. Department of Civil Engineering, University of Idaho, Moscow, Idaho. April 1989

MAYCOCK, G. AND HALL, R.D. Accidents at 4-Arm Roundabouts. Transportation and Road Research Laboratory, TRRL Laboratory Report 1120, Crowthorne, Berkshire, 1984

PAYNE, C.D. (Editor) The Generalized Linear Interactive Modelling System - Release 3.77 Manual, User's Guide and Reference Guide. Numerical Algorithm Group. Royal Statistical Society. 1987.

SABEY, B.E. AND STAUGHTON, G.C. Interacting Roles of Road Environment, Vehicle and Road User in Accidents. Fifth International Conference of the International Association for Accident and Traffic Medicine, and the Third International Conference on Drug Abuse of the International Council on Alcohol and Addiction, London, September 1975.

SABEY, B.E. Recent Developments and Research in Road Safety Remedial Measures. Road Safety in the 80's. Presented to Symposium, Salford, September 1983.