ACCIDENT PREDICTION MODELS FOR SIGNALIZED INTERSECTIONS

BY

SHIRLEY FENG

B.Sc, Beijing University of Aeronautics and Astronautics, 1988

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Department of Civil Engineering)

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

December 21, 1998

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her reporesentatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Civil Engineering
The university of British Columbia
Vancouver, Canada

Date April 27, 98

ABSTRACT

This thesis describes the development of accident prediction models for signalized intersections in the Greater Vancouver Regional District (GVRD). The traffic and road-related factors which appeared to underlie the occurrence of accidents are examined and models which explain, in a statistical sense, the generation of accidents as a function of these factors are developed. Recognizing the statistical and practical shortcomings associated with the use of the Conventional Linear Regression approach to develop accident prediction models, it was decided to utilize the Generalized Linear Regression Models (GLIM) approach. This approach addresses and overcomes the error structure problems that are associated with the conventional linear regression theory and allows for the use of nonlinear relationships in the model. In addition, the safety predictions obtained from the GLIM models can be refined using the Empirical Bayes' approach to provide, more accurate, site-specific safety estimates. The use of the complementary Empirical Bayes approach can significantly reduce the regression to the mean bias that are inherent in observed accident counts.

The study made use of sample accident, traffic and intersection design data corresponding to signalized intersections located in the Greater Vancouver Region. The accident data set contained 67 urban intersections from the City of Richmond and 72 urban intersections from the City of Vancouver giving a total of 139 intersections. Three different types of models were developed: (1) models relating the total number of accidents to traffic

volume; (2) models relating accidents of a specific type to traffic volume; and (3) models incorporating other geometric design variables such as the existence of left turn lanes, right turn lanes, pedestrian crossings, etc. The goodness of fit of the models was evaluated using two statistics: the Scaled Deviance (*SD*) and the Pearson χ^2 statistics. The overall fit of the models was adequate. Three applications of the GLIM models and the Empirical Bayes refinement process were described. The first related to the identification of accident prone locations. The second related to the before and after safety analysis and the third to safety planning. The usefulness of the GLIM model estimates in accounting for the randomness inherent in the accident occurrence process and the regression to the mean bias was documented and discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Acknowledgment

It gives me great pleasure to thank my supervisor Dr. Tarek Sayed. Throughout my research, he carefully and patiently directed every step. He spent countless hours discussing the problem with me, and made numerous constructive suggestions to keep me in progress.

I must also thank the Insurance Corporation of British Columbia (ICBC) for the financial support. I also thank Hamilton Associates and the Engineering Division of City of Vancouver, City of Richmond for their data supply.

Finally, I thank my husband Jing Qu for his support and encouragement. This work is respectfully dedicated to Jing Qu.

1.0 INTRODUCTION

1.1 Background

Traffic safety in British Columbia is a serious concern of traffic analysts and road authorities. In 1995, there were 534 fatalities, about 50,000 injuries and 180,000 property damage only accidents. The direct annual cost to the province exceeds 2.0 billion dollars (ICBC 1995 Annual Report, Vancouver, B.C.).

Recognizing these safety problems and the need to reduce the social and economic costs associated with them, road safety authorities have established Road Safety Improvement Programs (RSIPs). The objective of these programs is to monitor traffic conditions, collect and analyze accident data, locate trouble spots with abnormally high accident occurrences and implement appropriate and effective countermeasures in order to improve the safety potential of these sites. The success of these Road Safety Improvement Programs can be enhanced by developing reliable accident prediction models, which provide accurate estimates for the long-term safety potential at the locations under study. The development of such models is the focus of this thesis.

1.2 Aims and Objectives

The main objective of this thesis is to develop accident prediction models for estimating the long-term safety potential for signalized intersections in the Greater Vancouver

Regional District (GVRD). This research stemmed from the apparent need to estimate the safety potential of these intersections on the basis of their traffic and road geometric design characteristics. The aim, therefore, is to examine the traffic and road-related factors that appeared to underlie the occurrence of accidents and explain, in statistical sense, the generation of accidents as a function of these factors. The specific objectives of the study evolved as follows:

- Analyze the road, traffic and accident data available for signalized intersections under study in order to establish statistical models describing the empirical relationships between accident measures and the road and traffic variables. The accuracy and significance of each model and its coefficient estimates are investigated and discussed.

- Investigate the importance of various road geometric design variables in affecting the safety of signalized intersections. Some of these variables include: number of lanes, existence of left/right turning lanes and pedestrian crosswalk.

- Develop time-specific accident prediction models. These models may serve to address safety issues such as, comparing the intersection safety during peak hour to off peak hour, night to day, or AM peak hour vs. PM peak hour. Such time-specific models may be useful for evaluating strategies that affect traffic volumes or safety during certain time periods of the day and identifying potentially hazardous operating conditions.

- Develop type-specific accident prediction models. These models may be useful in estimating the occurrence of specific types of accidents such as left-turn and rear-end accidents. Often these specific types of accidents can be targeted by specific measures. For instance, the safety of an intersection that exhibits an abnormally high occurrence of left-turning accidents may be improved by implementing a left-turn lane or phase.

- Assess the reliability and interpret the significance of various models and detect any patterns in accident occurrence, identify and relate high-risk traffic and road characteristics to unsafety, estimate the safety potential of the implementation of certain road features and provide a global as well as local assessment of traffic safety of the GVRD signalized intersections.

There are several potential applications for the proposed accident prediction models. These applications are outlined in the next section.

1.3 Application of Accident Prediction Models

The development of reliable accident prediction models for signalized intersections offers a useful tool in a number of respects. Some of the potential applications of such models include:

- Identifying accident prone signalized intersections: the occurrence of a relatively high number of accidents at a particular location, though undesirable, does not always mean that the location concerned is a blackspot or would benefit from remedial treatment. Also, due to the regression to the mean phenomenon (as will be explained), an abnormally high number of accidents may not reflect the long-term accident occurrence at the location, and could in fact be followed by a relatively low number of accidents over a similar succeeding period of time even if no changes were introduced to the location. Therefore, in identifying accident prone locations, the true accident potential for each site should be estimated. This can be achieved by using accident prediction models.

- Before and after studies: to estimate the effectiveness of a safety measure, the expected number of accidents had the measure not implemented need to be estimated. This can be achieved by using accident prediction models and the Empirical Bayes refinement as will be discussed later.

- In safety planning through identifying the traffic and geometric variables which have the most impact on the safety performance of signalized intersections so that the road authorities can focus their attention and investment on targeting these variables. Furthermore, the change of safety associated with the change in any traffic or geometric design variables can be estimated (e.g. the change of safety associated with increased traffic volume, etc.)

1.4 Thesis Structure

Chapter One provides an overview of the thesis and its structure. A literature review of the various techniques for developing accident prediction models and the main concepts behind the Generalized Linear Regression Modeling technique is outlined in Chapter Two. Chapter Three describes the accident data and the accident prediction models developed. Chapter Four discuses three applications of accident prediction models. Chapter Five provides suggestions for follow up work and the summary and conclusion of the thesis.

2.0 LITERATURE REVIEW

The past decade has seen significant developments and advances in accident data analysis and modeling. Accident prediction models are no longer limited to the conventional linear regression assumptions, as more suitable and less restrictive nonlinear models can now be considered. The use of the Bayes' theory for developing accident prediction models has also been an important development in the data analysis literature. In this section, a brief outline of the major developments in accident prediction models will be presented.

2.1 Conventional Linear Regression Models

Conventional Linear Regression (CLR) models were initially used to describe the empirical relationships between accidents and traffic and road geometric design variables, such as traffic exposure (volume), horizontal curvature, vertical grade, lane width, and others (Zegeer et al. , 1987, Miaou et al., 1991). The underlying assumption of these CLR models is that the number of accidents, Y_i at a site i, in a reference population of size n, varies linearly with a set of m traffic and road geometric variables, { $x_{i1}, x_{i2},\ldots,x_{im}$ } as follows:

$$Y_i = a_0 + \sum_{j=1}^{m} a_j x_{ij} + e_i, \qquad i=1,2,\ldots,n. \tag{2.1}$$

where a_0, a_1, ..., a_m are the model parameters to be estimated by the least squares or the maximum likelihood method. These estimates can be obtained by using any of the numerous standard linear regression computer software available. The term e_i measures the error, i.e. the difference between the model estimate and the observed accident number at the site. It is assumed that the error terms e_i, $i=1,2,\ldots n$, at all sites of the reference population are independently and normally distributed random variables with zero mean and variance $\sigma^2 \geq 0$. As well be explained later, this assumption is restrictive and may not be suitable for the typically discrete and non-negative accident data.

Lau and May (1988) investigated the use of conventional linear regression model to describe a simple relationship between the number of injury accidents per year and the average yearly traffic volume for a population of 2,488 intersections in California. It was demonstrated for this data set that traffic exposure was the most important single factor in predicting injury accidents, and the following model was found to provide the best fit:

$$\textit{Injury accidents/year} = a_0 + a_1.(\textit{millions of entering vehicles/year}) \quad (2.2)$$

where a_0 and a_1 are the model parameters . The model was also used to predict the number of property damage only (PDO) and fatal accidents.

Conventional linear regression models were also used in discriminate analysis to investigate whether a site is predicted to exhibit a potentially high number of certain types of accident. Al-Senan et al (1987) used the discriminate analysis technique to identify the significant predictors of head-on accident on highway sections. Although head-on crashes are relatively rare, this class of vehicular accidents accounted for 14.6% of highway fatalities in the

United States for the period 1982 to 1984 (Accident Facts, 1985). The authors proposed prediction models for head-on accidents sites to identify sites that are prone to head-on accidents based on traffic and road design variables.

The following traffic and road features were found to be significant predictors of head-on accident proneness of a highway section: pavement width, shoulder width, pavement-shoulder combination, horizontal alignment, vertical alignment, combined horizontal-vertical alignment, roadside elements, traffic control features. The discriminate analysis was used to distinguish between two groups of sites, head-on crash sites and control sites (which have less potential for head-on crashes), based on the above geometric and traffic control features. The discriminate analysis uses the following conventional linear regression model to describe the empirical relationship between the discriminate function *D* and a set of traffic and road geometric design variables $\{V_1, V_2, \ldots, V_p\}$ as follows:

$$D = a_0 + a_1 V_1 + a_2 V_2 + \ldots + a_p V_p \tag{2.3}$$

The section is then assigned to the head-on group if the resulting discriminate score, *D* is less than zero; otherwise it is assigned to the control group. Discriminate analysis was shown to be a logical and convenient way to differentiate between sites with different accident contributing factors.

2.2 Shortcomings of Conventional Linear Regression

Conventional linear regression (CLR) models were initially widely used to develop accident prediction models, mainly because of their simplicity. However, their success has been generally limited. It has been demonstrated that CLR models lack the distributional property to adequately describe random, discrete, non-negative, and typically sporadic vehicle accident events on the road (Jovanis and Chang, 1986, Saccomanno and Buyco 1988, Miaou and Lam 1993).

Many difficulties were associated with the use of conventional linear regression to build accident prediction models for accident data (Persaud 1992). First, it is assumed, a priori, that accidents are proportional to traffic volume and the accident rate is used as the dependent variable. However, there has been much research to suggest that this assumption is not only inaccurate but it also leads to contradicting results (Mahalel, 1985). Secondly, most conventional regression modeling software assumes that the dependent variable has a normally distributed error structure. For accident counts, which are discrete and nonnegative, this is clearly not the case; in fact a negative binomial or Poisson error structure has been shown to be more appropriate (Persaud, 1989).

Another difficulty with the use of the conventional regression models lies in the unreliability of the estimates. Regression estimation suggests that two sites of the same reference population and which have similar independent variables (i.e. traffic and road geometric factors) values would have the same accident potential. However, in general, this is not the case, since it is not possible to account for all the factors that cause differences in accident potential among similar sites. The need to overcome these difficulties was fundamental to developing a more reliable and accurate regression model describing the relationship between the unsafety and the traffic and geometric traits.

The use of Generalized Linear Regression Modeling (GLIM) theory and software to develop accident prediction models has been investigated (Hauer et al., 1988; Persaud, 1989). The GLIM approach addresses and overcomes the error structure problems that are associated with the conventional linear regression theory and allows for the use of nonlinear relationships in the model. The use of the GLIM models for accident prediction is the topic of the following section. To distinguish between the safety estimates of similar sites, an Empirical Bayesian method that combines the regression model prediction with the observed short-term accident count at each location will also be discussed.

2.3 Generalized Linear Regression Models with Empirical Bayes Refinement

The need to overcome the statistical and practical shortcomings associated with the use of conventional regression models for accident prediction was fundamental to the development of realistic and more general empirical relationships between the accident potential, traffic and road design variables. To this end, use was made of a statistical generalized linear regression modeling package (GLIM) that allows for the flexibility of nonlinear accident-traffic relationship and user specific error structure for the dependent variable. Use was also made of a complementary Empirical Bayesian procedure for improving the accuracy of the regression model accident predictions. The theory of these developments will be briefly outlined next.

2.3.1 Generalized Linear Regression Models

The Generalized Linear Regression Method (GLIM) allows for non-linear empirical relationships between the dependent and the independent variables that can be linearized by taking its logarithmic function. GLIM is a recently developed statistical software package (Baker and Nelder, 1978) that is being widely used in accident data analysis. It also allows the specification of a negative binomial or a Poisson error structure for the dependent variable, which as noted earlier, is more appropriate for accident counts than the traditional normal distribution (Persaud, 1989).

The GLIM approach utilized in this study is based on the work of Kulmala (1995) and Hauer et al. (1988). Assume that Y is a random variable describing the number of accidents at an intersection in a specific time period, and y is the observation of this variable during a period of time. The mean of Y is Λ which is itself can be regarded as a random variable. Then for $\Lambda = \lambda$, Y is Poisson distributed with parameter λ:

$$P\left(Y - y \middle| \Lambda = \lambda\right) = \frac{\lambda^{y} e^{-\lambda}}{y!}; E(Y|\Lambda = \lambda) = \lambda; \ \mathrm{Var}(Y|\Lambda = \lambda) = \lambda \qquad (2.4)$$

Hauer et al. (1988) have shown that Λ follows a gamma distribution (with parameters κ and κ/μ), where κ is the shape parameter and μ is the mean of the distribution:

$$f(\lambda) = \frac{(\kappa/\mu)^{\kappa}\,\lambda^{\kappa-1}\,e^{-(\kappa/\lambda)\lambda}}{\Gamma(\kappa)};\;\; E(\Lambda) = \mu;\;\; \mathrm{Var}(\Lambda) = \frac{\mu^2}{\kappa} \tag{2.5}$$

Kulmala (1995) has also shown that the point probability function of *Y* based on (2.4) and (2.5) is given by the negative binomial distribution:

$$P(Y = y) = \frac{\Gamma(\kappa + y)}{\Gamma(\kappa)\,y!}\left(\frac{\kappa}{\kappa+\mu}\right)\left(\frac{\mu}{\kappa+\mu}\right)^{y} \tag{2.6}$$

with an expected value and variance of:

$$E(Y) = \mu;\;\; \mathrm{Var(Y)} = \mu + \frac{\mu^2}{\kappa} \tag{2.7}$$

As shown in equation (2.6), the variance of the expected number of accidents is generally larger than its expected value. The only exception is when $\kappa \to \infty$, where the distribution of   is concentrated at a point and the negative binomial distribution is identical to the Poisson distribution (Kulmala, 1995).

As described earlier, for the generalized linear regression modeling approach, the error structure is usually assumed to be Poisson or negative binomial. The main advantage of the Poisson error structure is the simplicity of the calculations (the mean and the variance are equal). However, this advantage is also a limitation. It has been shown (Kulmala, 1995) that most accident data will likely to be over dispersed (the variance is greater than the mean) which indicates that the negative binomial distribution is the more realistic assumption. However, the difference between the model parameters estimated using the Poisson and the negative binomial models was found to be very small (Kulmala, 1995). Therefore, the simpler Poisson error structure assumption will be used in this study.

The first step in developing a GLIM model is to subgroup similar sites that share a set of traffic or geometric characteristics in a reference population. The empirical relationship between the unsafety potential, measured in terms of the number of accidents at the location, and the traffic and road geometric design variables $\{x_{1,} x_{2,} ..., x_m\}$ is then described in an equation form as follows:

$$\text{Accidents / year} = \textit{function}\,(\text{traffic and road geometric design variables; } x_1, x_2, \ldots x_m)$$

where the form of the *function* is appropriately chosen so that it can be linearized by taking the logarithm. The above regression estimate would not be useful if no measure of its variability is known. It has been shown that, using the generalized linear regression, the variance of the regression estimate is directly proportional to the square of the model safety estimate in the following fashion (Persaud, 1989):

$$variance(regression\ estimate) = \frac{(regression\ estimate)^2}{k} \tag{2.8}$$

where the value of k depends of the structural error assumptions of GLIM. The estimation of the variance from the GLIM regression estimates is difficult since both k and the variance in equation (2.4) are unknown. Hauer et al. (1988) suggested an iterative process to calculate k using the maximum likelihood method. Kulmala (1995) showed that the method of moments (assuming a Poisson model) can also be used to calculate k as follows:

$$k = \frac{\frac{1}{n}\sum_{i=1}^{n} E^2(m_i)}{\frac{1}{n}\sum_{i=1}^{n}\left[\left(X_i - E(m_i)\right)^2 - E(m_i)\right]} \tag{2.9}$$

where *E(m)* is the regression estimate of the number of accidents at the location, and *Xi* is the location observed number of accidents.

Kulmala (1995) found that the method of moments produced accurate enough estimates that deviated less than 5% from those produced by the maximum likelihood method. He also indicated that the iterative process suggested by Hauer et al. (1988) was time consuming and changed the value of k only to a minor extent.

One of the main conceptual difficulties associated with the GLIM model involves the choice of the form of the *function* that describes the dependence of the unsafety potential of the site on the traffic and road variables. Unlike conventional regression models, which restrict the function to be linear, The GLIM model allows nonlinear forms. Often, a scatter plot of the number of accidents as a function of the traffic and road geometric variables may provide some hints for choosing the function form. However, the choice of the form of the *function* is generally intuitive and lacks theoretical justifications. Since the choice of the function form is a matter of judgment, different analysts may make different choices, and therefore, arriving at discrepant estimates for the unsafety of the same entity.

Another difficulty with the use of regression models to estimate the unsafety of a site is that the above equation suggests that two sites in the same reference population that have similar values for the independent variables, $x_1, x_2, \ldots, x_m,$ are expected to have similar unsafety estimates. However, in general, this is not the case, since it is not possible to account for all the factors that cause differences in accident potentials among similar sites. The use of the Empirical Bayesian method, which combines the regression model safety estimate with the sites-specific accident count, to refine the GLIM estimates could alleviate these shortcomings.

2.3.2 The Empirical Bayes Refinement

There are two types of clues to the unsafety of an entity: its traffic and road geometric design characteristics, and its historical accident data (Hauer, 1992). The GLIM method provides an unsafety estimate of a site based on the first type of clues. The Empirical Bayes (EB) approach to unsafety estimation makes use of both kinds of clues. The EB method combines the GLIM unsafety estimate with the site-specific accident history and yields better estimates of unsafety.

To illustrate, suppose that an entity is located in a reference population characterized by a set of traits. Let *pred,* and *var (pred)*, be the GLIM model estimates for the location's safety and its variance respectively, as described in the previous section. Suppose that the observed number of accidents at this site during the specified period of time is given by *count.* The Empirical Bayesian method combines the GLIM model estimates with the observed accident count to obtain a more refined, site-specific unsafety estimate, in the following fashion (Hauer, 1992):

$$EB\ safety\ estimate = \alpha . pred + (1 - \alpha).count, \quad where \quad \alpha = \frac{1}{1 + \frac{var(pred)}{pred}} \tag{2.10}$$

$$where \quad \alpha = \frac{1}{1 + \frac{var(pred)}{pred}} \tag{2.11}$$

The form of the above equation is consistent with the following reasoning: if the reference population is homogeneous (i.e. its sites are very similar), then one would expect small variations among the GLIM model safety estimates (i.e. *var(pred)* *0*), therefore *1*, and the above equation yields an *EB safety estimate* that is close to the GLIM model estimate, *pred*, as it should be since the effects of the accident counts should not influence the estimate as differences in accident counts among sites may be attributed to chance variations. On the other hand, if the reference population is heterogeneous (i.e. the sites are diverse), then one would expect a relatively high GLIM estimates variations, *var(pred)*, and *a»0*.

The above equation yields an EB estimate that is close to the observed accident *count* and what is known about the reference population exerts little influence on the estimation. This again is as x should be, since the locations are very diverse, then differences in accident counts should be attributed to the differences among sites and not to chance variations. In view of the absence of a theoretical justification for the above equation, it is important to know that the EB method combines the regression estimates with the site-specific accident counts in a practical manner (Hauer, 1992).

In addition to combining the two types of safety clues and providing site-specific safety estimates, it has also been shown that the EB method significantly reduces the regression to the mean effects that are inherent in observed accidents count (Brude and Larsson, 1988).

The regression to the mean is a statistical phenomenon by which a randomly large number of accidents for a certain entity during a before period is normally followed by a reduced number of accidents during a similar after period, even if no measures have been implemented (while the opposite applies in the case of a randomly small number of accidents). As mentioned earlier, before and after studies are one of the main applications of the accident prediction models. For instance, in a before and after study of the effect of a particular action, one should not compare the after period accident counts to the before period accident counts, this generally leads to misleading assessment of the safety improvement effect afforded by the undertaken measures. Instead, one should compare the after period accident counts to the prediction model safety estimates, had no measures been implemented.

2.3.3 Applications

The above two-step procedure, where the GLIM estimates are refined by using the Empirical Bayesian method has been used in the literature to estimate the unsafety of road sections and intersections. Next a brief review of some of these applications for developing the models for intersections will be presented.

2.3.3.1 Accident Prediction Models for Intersections

Although many studies have addressed the relationship between accidents and the traffic and geometric factors at road segments (Jovanis and Chang, 1986, Saccomanno and Buyco 1988, Miaou and Lam 1993), only a few studies have addressed this relationship for signalized intersections (Poch and Mannering, 1996). This is surprising given that accidents at signalized intersections represent a significant proportion (more than half) of the total accident population, especially in urban areas. Thus, the development of accident prediction models for intersections is of great importance. These accidents usually involve multiple vehicles. Numerous models to estimate the safety potential of an intersection on the basis of its traffic flow, geometric design features and accident history have been suggested over the years. Initially, it was suggested that the number of all accidents at an intersection is proportional to the sum of flows that enter the intersection. The merit of this approach is its simplicity. However, it has several shortcomings. Such models assume a uniform traffic flow through the intersection. However, this is usually not the case, since most intersections consist of a major road, with a higher average annual daily traffic volume (AADT) crossing a minor one with a lower (AADT).

A commonly used GLIM accident prediction model for intersections that makes the distinction between the major and minor road traffic flow, relates the unsafety potential of the intersection as a function of the AADT for the minor and the major roads as follows:

$$Accidents \ / \ year \ = \ a_0 \times (AADT_{major\ road})^{a_1} \times (AADT_{minor\ road})^{a_2} \qquad (2.12)$$

Webb (1955) used data for rural signalized intersections in California developed the following model:

$$Accident \ / \ Year = 0.007(AADT_{majorroad})^{0.51}(AADT_{minorroad})^{0.29} \qquad (2.13)$$

Lau et. al. (1989) developed separate models for fatal, injury, and property damage only (PDO) accidents as follows:

PDO Accidents / year = 4.63 + 0,514 (the sum of entering vehicles in million)

$$\textit{Injury accidents / year} = 0.62 + 0.169\ \textit{(the sum of entering vehicles in million)}$$

$$\textit{Fatal Accidents / year} = 0.018 \qquad (2.14)$$

Bonneson et al. (1993) developed a similar model to that of Webb (1955) using data for non-urban signalized intersections:

$$Accident\ /\ Year = 0.00703(AADT_{major\ road})^{0.721}(AADT_{minor\ road})^{0.366} \qquad (2.15)$$

Bonneson and McCoy (1993) have found that this model provided a best fit for accident data from 125 rural unsignalized intersection in Minnesota for the three years 1985-1987. The same model was found to fit accident data from signalized urban intersections in Philadelphia (Persaud et al 1995), accident data from 149 unsignalized rural intersections in Quebec (Belanger, 1995), and accident data from signalized rural intersections in Virginia (Hanna, 1976).

Brude and Larsson (1988) investigated a different form of combining the minor road and major road traffic flows. They suggested the following GLIM prediction model for four-legged signalized junctions:

$$Accidents/year = a_0 \times (AADT_{major\ road} + AADT_{minor\ road})^{a_1} \times \left(\frac{AADT_{minor\ road}}{AADT_{major\ road} + AADT_{minorroad}}\right)^{a_2} \quad (2.16)$$

This model was found to provide a better fit for the accident data set under consideration than the previous model which relates the number of accidents to the traffic flow on the minor and the major road in a similar manner.

Kulmala (1992) used the following GLIM model to estimate the expected number of accidents for highway junctions in a reference population as a function of the total number of vehicles entering the junction, the minor road's portion of entering traffic and a set of variables describing the geometry and the environment of the junction, $\{x_1, x_2, \ldots x_n\}$, as follows:

$$Accidents / year / km = a_0 \times (AADT_{total})^{a_1} \times (AADT_{minor\ road})^{a_2} \times e^{\sum_i^n b_i x_j} \qquad (2.17)$$

In a similar fashion to the Empirical Bayes refinement discussed earlier, the above regression estimates are then combined with the accident history for each site to yield a site-specific refined estimates. Kulmala (1992) has used this GLIM/EB accident prediction model to estimate the changes in the number of accidents due to road measures in a before and after study. Most notably, it was observed that the number of accidents at three-leg intersections was reduced by 44% and 48% by implementing stop signs and lighting respectively.

Similarly, for four-leg intersections, the above countermeasures reduced the accidents by 5% and 15% respectively. It is also worth noting that the implementation of a right turning lane has resulted in a 17% increase in the number of accidents for three-leg junctions and a 17% decrease for four-leg junctions.

2.3.3.2 Accident Classification at Intersections

Hauer et al (1988) produced disaggregate models for 145 four-legged, fixed time, signalized intersections in Metropolitan Toronto. He estimated separate models for the most common accident types, such as rear-end, angle accidents, etc. He used the following model forms:

Rear-end:

$$Accidents = a_0 \times (approach\ traffic)^{a_1} \tag{2.18}$$

Other accident types:

$$Accidents = a_0 \times (V_1)^{a_1} \times (V_2)^{a_2} \tag{2.19}$$

where V_1 and V_2 are the pattern specific traffic volumes.

Categorizing accident in this fashion has many advantages: better accident prediction models for specific types of accidents can be obtained; also from a remedy point of view, one could implement specific countermeasures to target specific types of accidents (such as constructing a left-turn lane for intersections with high predicted frequency of left-turning and straight through accidents).

2.3.3.3 Accident Prediction Models for Road Sections

The traffic exposure for road sections is typically defined in terms of the average annual daily traffic in millions of vehicles per kilometer. The development of accident prediction models for road, especially freeway sections, is of great interest because of the severity and frequency of these, typically high speed, accidents. Persaud (1991) used the GLIM model with the two step EB refinement procedure described earlier to study the unsafety potential of three types of classes of Ontario road sections. The GLIM model suggested for the unsafety estimates as a function of the section length and the annual average daily traffic was given by:

$$Accidents\,/\,day \;=\; a_0 \times (section\ length) \times (AADT)^{a_1} \tag{2.20}$$

where *a, b* are the model parameters estimated by GLIM. The GLIM estimates are then refined using the EB method.

It is worth mentioning that one would expect that the unsafety estimate of a road section would depend on many traffic and road geometric variables other than the section length and the average daily traffic. However, this model does indirectly account for many other variables through using three highway classes. For instance, class 1 road sections consists of freeway sections that are multi- lane, divided, have the same high geometric standards, and are usually similar in other features such as speed limit.

Therefore, it seems unlikely that a model which incorporates some of these road geometric variables would be significantly better than the above traffic volume, road section length model. This is because, these variables are kept practically constant within the class 1 reference group. Indeed, the author has shown that many attempts to add more geometric variables to the model did not result in significant difference in the estimates. This is an important observation, since one can deduce that simple GLIM accident prediction models can be developed if a reference population, in which some traffic and road geometric variables are kept constant, is appropriately chosen.

Several researchers have developed multi-variate models which consider other variables in addition to traffic flow. For example, Zeeger et al (1986) suggested the following multi-variate GLIM model for two-lane rural roads:

$$Accidents/(year\ -\ km) = a_0 \times AADT^{a_1} \times a_2^{\text{lane width}} \times a_3^{\text{average paved shoulder width}} \times a_4^{\text{average unpaved shoulder width}} \times a_5^{\text{median recovery distance edge of shoulder}} \tag{2.21}$$

Example for other multi-variate models can be found in Forkenbrock et al. (1994).

2.4 Summary

Although many studies have addressed the relationship between accidents and the traffic and geometric factors at road segments, only a few studies have addressed this relationship for signalized intersections. This is surprising given that accidents at signalized intersections represent a significant proportion of the total accident population, especially in urban areas. Most of the existing research on accident prediction models for signalized intersections only consider the relationship between accidents and traffic flow.

Several researchers have shown that conventional linear regression models lack the distributional property to adequately describe random, discrete, non-negative, and typically sporadic events which are all characteristics of traffic accidents. The generalized linear modeling (GLIM) approach addresses and overcomes the error structure problems that are associated with the conventional linear regression theory and allows for the use of nonlinear relationships in the model. In addition, the safety predictions obtained from GLIM models can be refined using the Empirical Bayes approach to provide, more accurate, site-specific safety estimates. The use of the complementary Empirical Bayes approach can significantly reduce the regression to the mean bias that is inherent in observed accident counts.

There several applications of accident prediction models and the complementary Empirical Bayes refinement process including: the identification of accident prone locations, before and after safety analysis and safety planning.

3.0 MODEL DEVELOPMENT

3.1 Background

The research made use of sample accident, traffic and intersection design data corresponding to signalized intersections located in the Greater Vancouver Region. The data set contained 67 urban intersections from the City of Richmond and 72 urban intersections from the City of Vancouver giving a total of 139 intersections.

Before the analysis is considered, it is essential to provide a description of the data. The need to describe the data stems from various principles of accident data analysis and prediction model development which include:

- Accident prediction models are often data dependent, in the sense that an empirical relationship that provides a best-fit for accident data for locations in the Greater Vancouver, may not be the best-fit model for accident data for sites in Metro Toronto, for instance. In the absence of a unique and conventional accident prediction model that fits accident data everywhere, the mathematical forms of these empirical relationships are often intuitive and data dependent, and are only meant to describe the original accident data under study.

- The accuracy and reliability of these models also depend on the accuracy, the availability and the collection procedure of the accident data. The success of these accident prediction models in reliably estimating the long term safety potential of the signalized intersections under study is directly proportional to the quality of accident data which are available. Consequently, the reliability of the applications of these models such as identification of accident prone locations and before and after studies is closely related to the accuracy and availability of the accident data.

Therefore, in order to develop reliable accident prediction models for the signalized intersection under study, time and effort was devoted to checking and validating the accident data. In this section, a brief description of accident data, traffic exposure measure and intersections characteristics is provided.

3.1.1 Accident Data

Three years of accident data was available for analysis on each intersection (1993 - 1995). The source of the accident data is the MV104 accident reporting form, the British Columbia's accident police report. The MV104 police report is the principal tool used to collect information giving accidents in British Columbia. In this form, there are about one hundred pieces of information giving the accident circumstances, type and outcome as well as the characteristics of the driver(s), the vehicle involved and the locations of the accident. The data set included a total of 6255 accidents.

Table 1 provides a summary of the variables extracted from the MV104 forms for use in this study. These variables will be considered as the dependent variables in the multiple-regression accident prediction models. The data is included in Appendix A.

Variable	*Description*
TOTAL	Total number of accidents
PDO	Number of Property Damage Only (PDO) accidents
INJ	Number of injury accidents
PEAK(AM)	Number of morning peak hour accidents
PEAK(PM)	Number of afternoon peak hour accidents
NIGHT	Number of night-time accidents
DAY	Number of day-time accidents
LEFT	Number of left-turn accidents
REAR-END	Number of rear-end accidents

Table 3.1. Accident Variables from the MV104

3.1.2 Traffic Exposure Measures

The most commonly used intersection accident exposure is the total number of vehicles entering the intersection (sum of traffic flows). However, recent studies (Hauer et al. (1988) and others) have indicated that the "product-of-traffic-flows-to-power" model is more suitable to represent traffic exposure than the "sum of traffic flows". In these models accident frequency is a function of the product of traffic flows raised to a specific power (usually less than one).

Hauer et al. (1988) considered only traffic flows related to the accident pattern. Others, related the accident frequency to the product of average daily traffic of the major and minor roads. This later approach will be used in this study because of the difficulty in obtaining accident patterns-related traffic flows.

Table 3.2 provides a statistical summary of the ranges of traffic volume and accidents for the intersections used in this study.

Variable	Statistics			
	Minimum	Maximum	Mean	Std. Dev.
Major Road ADT	5700	67840	23790	9670
Minor Road ADT	300	54060	9420	7250
Accident/Year	2	42	15	9

Table 3.2 Statistical Summary for the Accident Traffic Volume Data for the Study Intersections

3.1.3 Intersection Characteristics

The objective of this research is to develop accident prediction models that link accident measures to not only traffic exposure but also to the intersection infrastructure and geometric design characteristics. The aim of accounting for these intersection attributes was to examine the road-related factors which appeared to underlie the occurrence of accidents. The intersection characteristics variables included in the model are summarized in Table 3.3. Some of the data was collected by direct observation of the intersections, and others are recorded from intersection layouts provided by the City of Vancouver and the City of Richmond.

Variable	*Description*
MANL	Number of Major Road lanes
MINL	Number of Minor Road Lanes
LT	Number of Left Turn Lanes
PRO	0, Unprotected; 1, Protected Left Turn Lane
RT	Number of Right Turn Lanes
PC	Existence of Pedestrian Crosswalk

Table 3.3 Intersection characteristics variables used in the study

3.2 The Modeling Technique

As described in Section 2.0, the use of Generalized Linear Regression (GLIM) models overcomes the statistical and practical shortcomings associated with the use of conventional regression models for accident prediction. The GLIM method allows for the flexibility of nonlinear accident-traffic relationship and the specification of a negative binomial or a Poisson error structure for the dependent variable which is more appropriate for accident counts than the traditional normal distribution. In the present study the Poisson distribution has been assumed for the error structure.

3.2.1 Measures of Significance

In conventional regression models, where a normal error structure is assumed, the coefficient of determination (R^2) is usually used to indicate the model significance. However, the use of (R^2) is not appropriate when the error structure is other than normal (Belanger, 1993).

Significance tests for GLIM models are based on a *scaled deviance SD* - a likelihood measure of discrepancy between two models. In GLIM, the *SD* is defined as the likelihood test ratios measuring the difference between the log likelihood of the studied model and the saturated model. Differences between scaled deviance for the two models, one of which is a submodel of the other, is attributed to the extra parameters included in the larger model. This provides a method of assessing the significance of each factor, which is based on the decrease of the scaled deviance SD after the inclusion of a considered factor into the model. Thus the deviance fills the role of the residual sum of squares in the normal model, in providing a significance test for the importance of parameters omitted or added to the model. The difference in scaled deviance between two models is assumed to follow a Chi-Square (χ^2) distribution, and the significance of such difference is assessed based on the degrees of freedom (*n-p*) and the value read from the χ^2 table.

Another measure of significance is the Pearson χ^2 statistic defined as:

$$Pearson\chi^2 = \sum_i \frac{\left[x_i - E(m_i)\right]^2}{VAR(x_i)} \tag{3.2}$$

Miao et al. (1992), Bonneson and McCoy (1993), and Persaud and Dzbik (1993) evaluated their models using the *Pearson* χ^2 method which follows a χ^2 distribution. Both the scaled deviance *SD* and the *Pearson* χ^2 will be used for the indication of the goodness of fit in this research.

In addition, a useful subjective measure of the model goodness of fit is to plot the predicted accident frequency versus the observed accident frequency. A well-fitted model should have all points in the graph clustered symmetrically around the 45° line.

3.2.2 Model Development

The main task of this research is to develop multivariate models to estimate the number of accidents. In this project, modeling was undertaken in three stages: (1) developing models relating the total number of accidents to traffic volume; (2) developing models relating accidents of a specific type to traffic volume; and (3) developing models incorporating intersection layout variables.

Because of the relatively small sample size, results for the development of separate models for the city of Vancouver and the city of Richmond is not reported. However, they are included in Appendix B.

3.2.2.1. Models for the Total Number of Accidents

As described earlier, there are different functional forms which can be used to relate the accident occurring at an intersection with the traffic volume. Two models where selected for use in this study:

(1) The model used by Hauer et. al. (1988) and Bonneson and McCoy (1993) :

$$Accidents\ /\ year\ =\ a_0 \times (AADT_{major\ road})^{a_1} \times (AADT_{minor\ road})^{a_2} \tag{3.3}$$

where AADT is the Average Daily traffic in 1000 veh/day.

The model used by Brude and Larsson (1988)

$$Accidents/year = a_0 \times (AADT_{major\ road} + AADT_{\min or\ road})^{a_1} \times (\frac{AADT_{\min or\ road}}{AADT_{major\ road} + AADT_{\min orroad}})^{a_2} \quad (3.4)$$

where AADT is the Average Daily traffic in 1000 veh/day.

The results of fitting the two models to the accident data are illustrated in Table 3.4.

Model	Model Form	Coefficient Estimates		SD (df)	k	Pearson χ^2
1	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0	2.1813	355.2	9.0	128.40*
		a_1	0.3286	(136)		
		a_2	0.4418			
2	$E(m) = a_0 \times (V_1 + V_2)^{a_1} \times (\frac{V_2}{V_1 + V_2})^{a_2}$	a_0	1.9066	360.7	8.8	128.57*
		a_1	0.7432	(136)		
		a_2	0.3622			

- Denotes significance at a 95-percent confidence level ($\chi^2_{0.05,136} = 163.8$).

Table 3.4 Models for the Estimation of the Total Number of Accidents

Table 3.4 shows that the two models present a relatively good-fit with the first model having a slightly higher goodness of fit (a smaller *SD*). Figures 3.1 and 3.2 show the relationships between the observed number of accidents/year versus the predicted number of accidents/year for the two models. The results are symmetrically clustered around the 45° line to a reasonable extent; which is desirable. The dispersion of the results from the 45° line is also within acceptable limits.

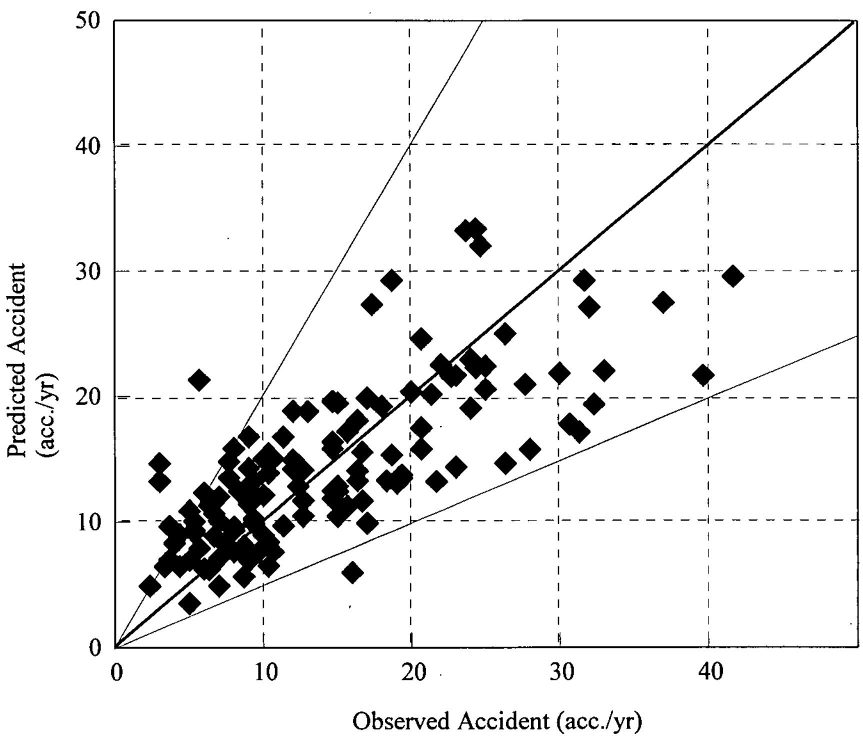

Figure 3.1 Model (1): Observed Versus Predicted Number of Accidents/yr

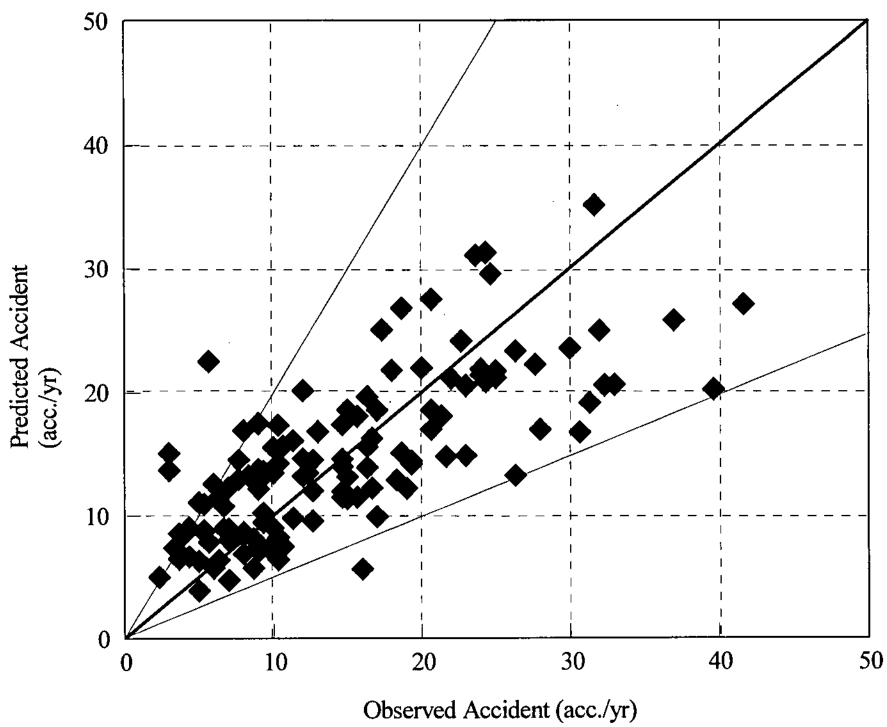

Figure 3.2 Model (2): Observed Versus Predicted Number of Accidents/yr.

3.2.2.2 Models for Specific Accident Types

The "product-of-flow-to-power" model suggested by Hauer et. al. (1988) and Bonneson and McCoy (1993) was used to model the relationships between specific accident types and traffic flow. Table 3.5 shows different accident type models and their goodness of fit. In the case of rear-end accidents the total traffic volume entering the intersection was used instead of the "product-of-flow-to-power" model. All models have a relatively good-fit and the χ^2 values are significant at a 95-percent confidence level.

Accident Type	Model Form	Coefficient Estimates		SD	k	Chi-square
PDO Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0	1.3012	253.7	10.8	128.25
		a_1	0.3409			
		a_2	0.4412			
Injury Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0	0.9049	169.2	14.4	120.54
		a_1	0.2836			
		a_2	0.4490			
Day time Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0	1.5470	259.6	11.2	126.97
		a_1	0.3470			
		a_2	0.4206			
Night time Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0	0.6714	204.4	7.8	137.30
		a_1	0.2547			
		a_2	0.5007			
Rear-end Accidents	$E(m) = a_0 \times V^{a_1}$	a_0	0.0627	159.3	19.5	120.74
		a_1	1.2360			
PM Rush hour Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0	0.3256	131.6	88.85	125.63
		a_1	0.4266			
		a_2	0.3944			
Rush Hour Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0	0.4439	154.3	26.64	154.30
		a_1	0.4466			
		a_2	0.4025			
Non-Rush Hour Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0	1.7400	269.6	10.36	128.45
		a_1	0.2873			
		a_2	0.4599			
Left-Turn Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0	0.5572	278.2	3.33	121.73
		a_1	0.2999			
		a_2	0.4950			

Table 3.5 Models for Specific Accident Types

The relationship between the types of accidents and major traffic volume for various minor road volumes are shown in Figures 3.3-3.12. Figure 3.13 shows the relationship between different accident types and the major traffic flow. Figure 3.14 shows the relationship between time specific accidents and the major road traffic flow. The minor road flow is kept at its mean value in Figures 3.13 and 3.14.

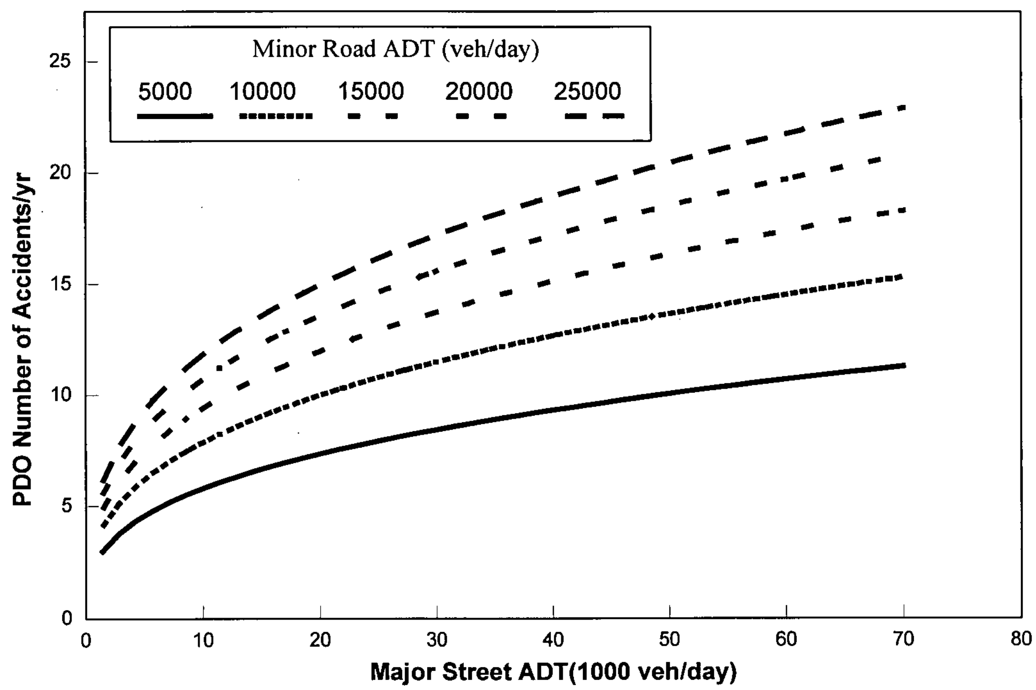

Figure 3.3. Model of PDO accidents

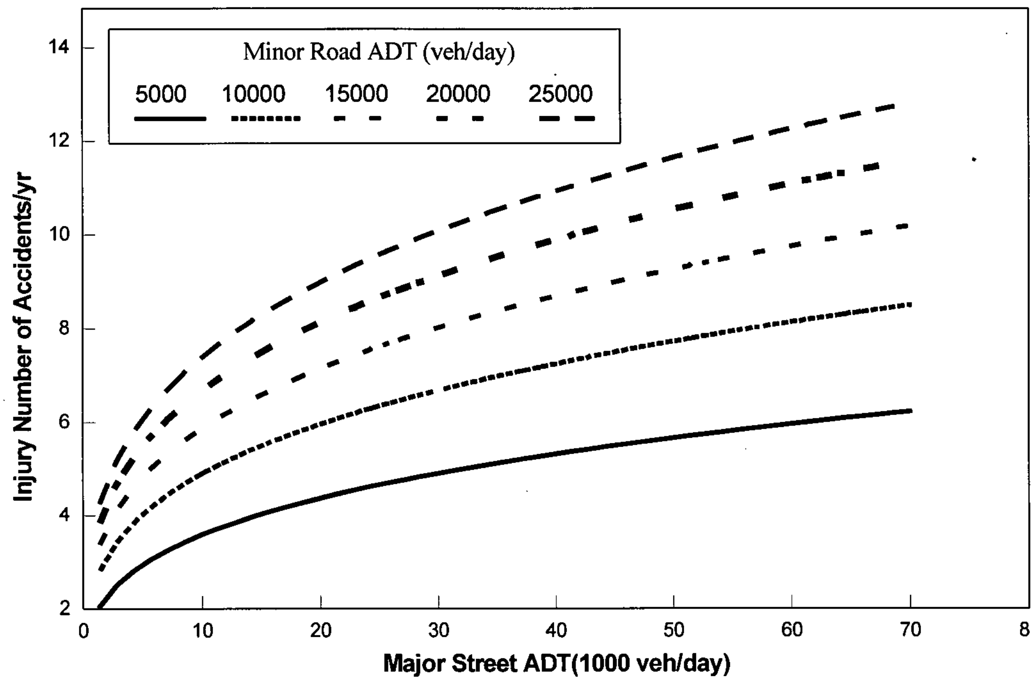

Figure 3.4. Model of injury accidents

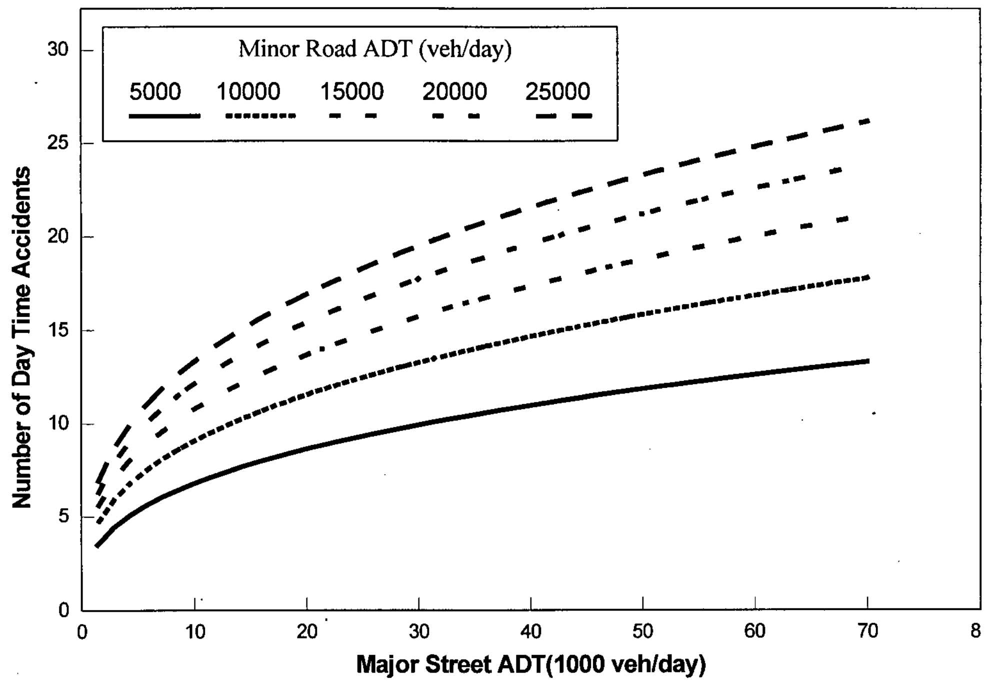

Figure 3.5 Model of day time accidents

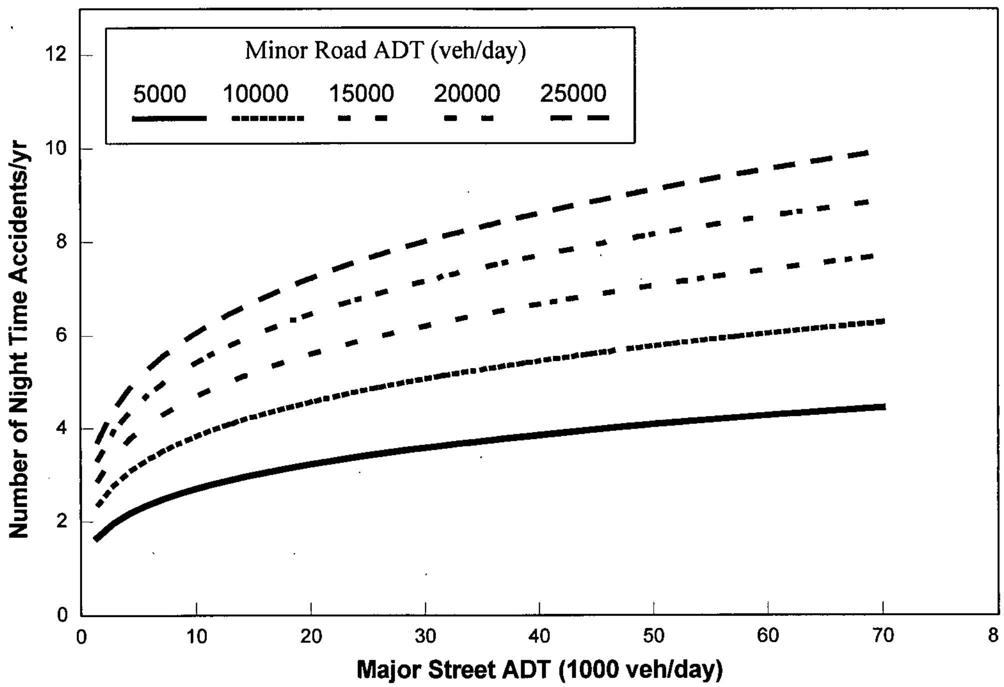

Figure 3.6 Model of night time accidents

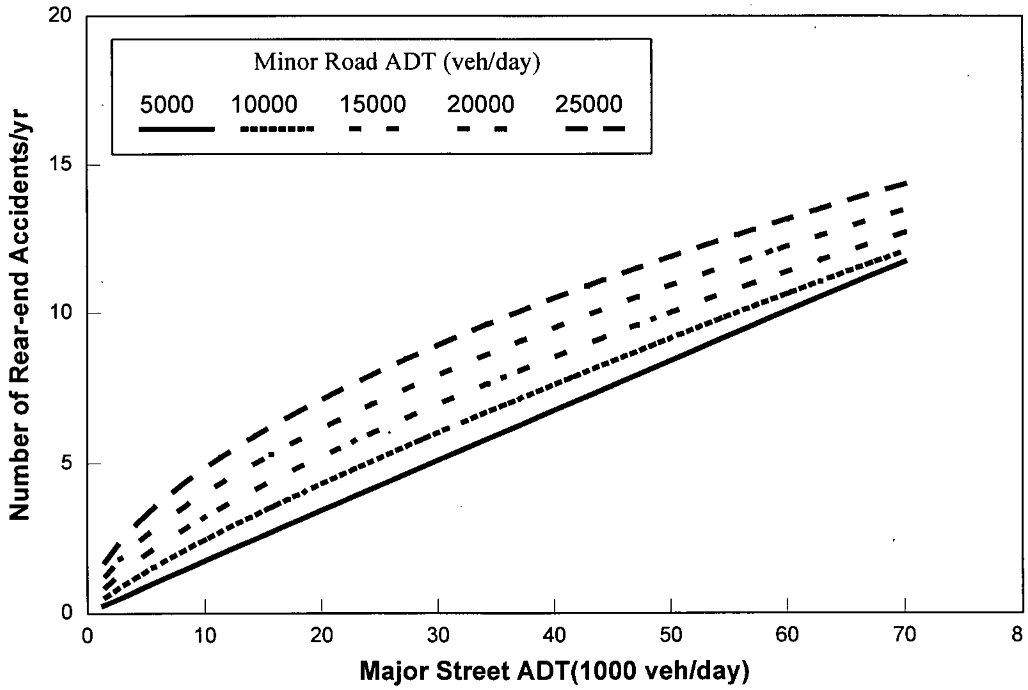

Figure 3.7 Model of Rear-end accidents

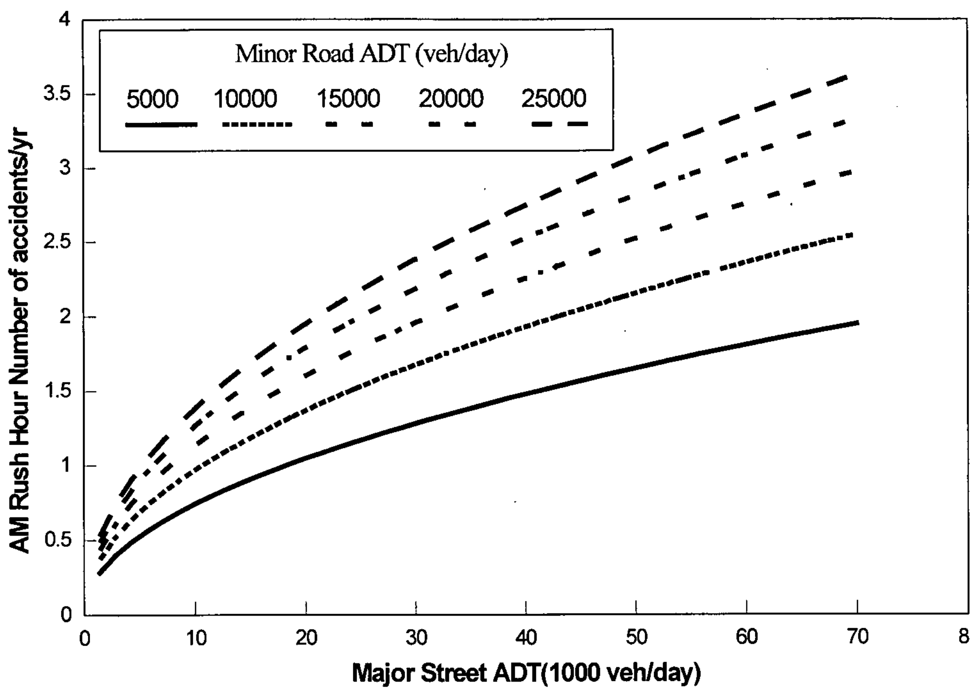

Figure 3.8 Model of AM rush hour accidents

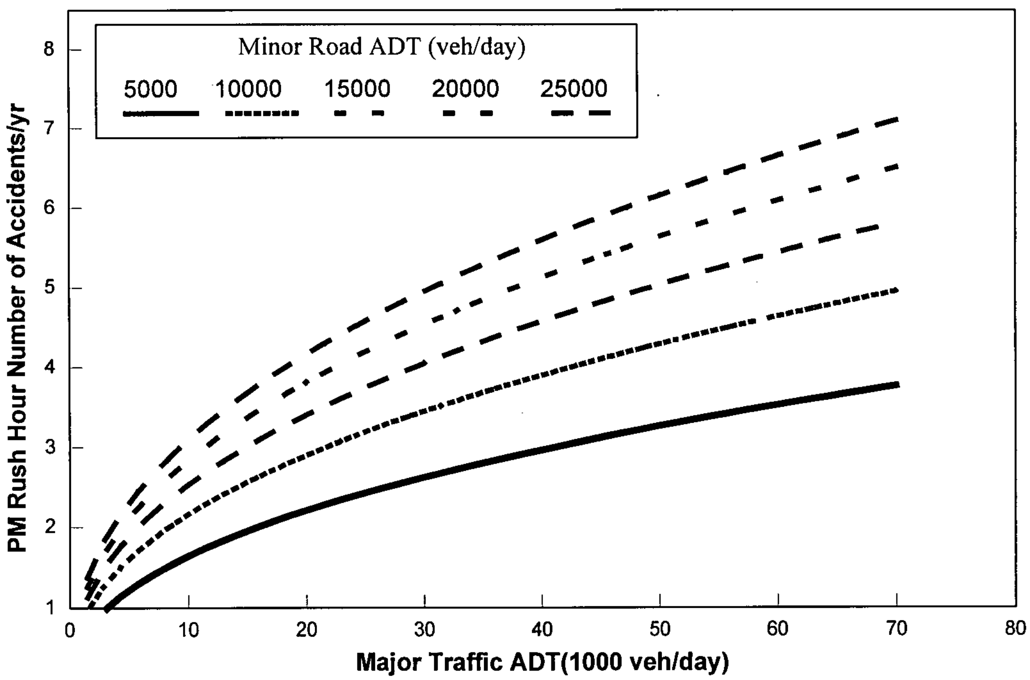

Figure 3.9 Model of PM rush hour accidents

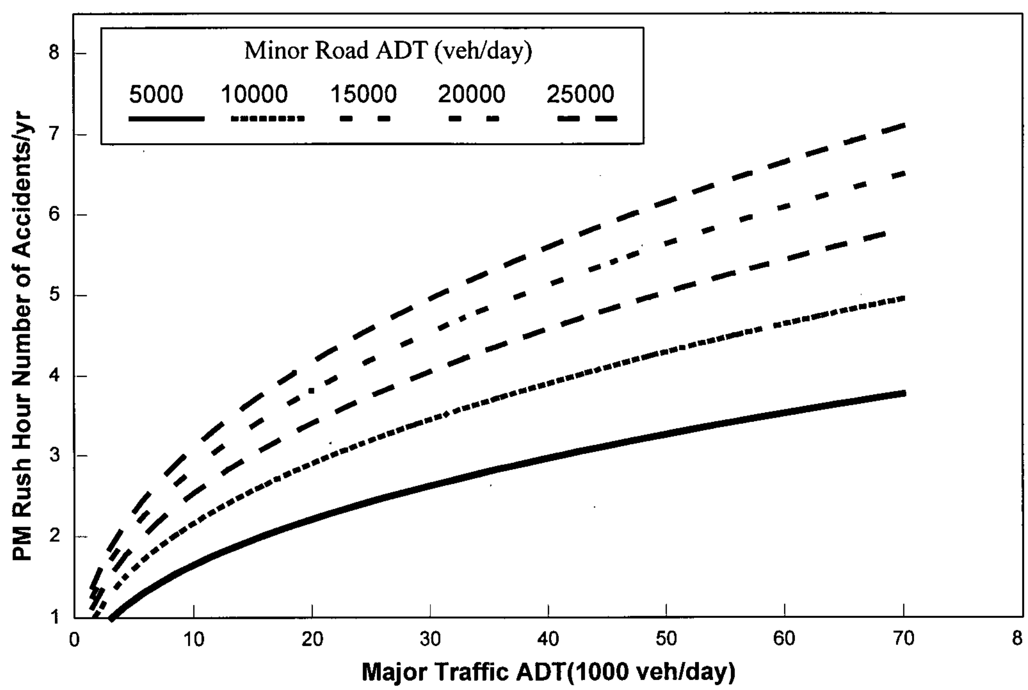

Figure 3.10 Model of rush hour accidents

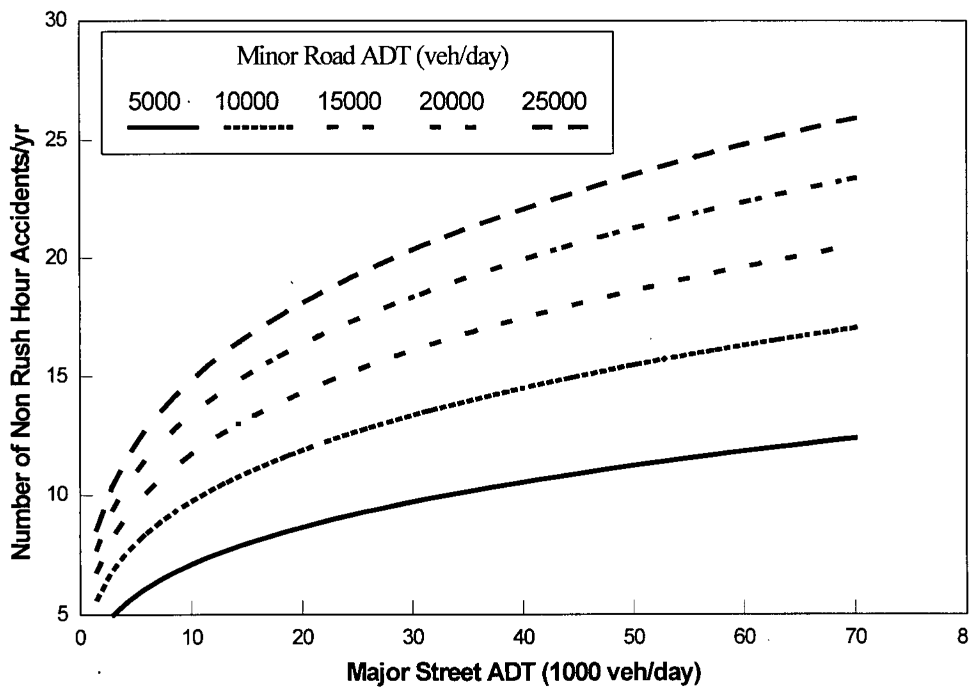

Figure 3.11 Model of Non rush hour accidents

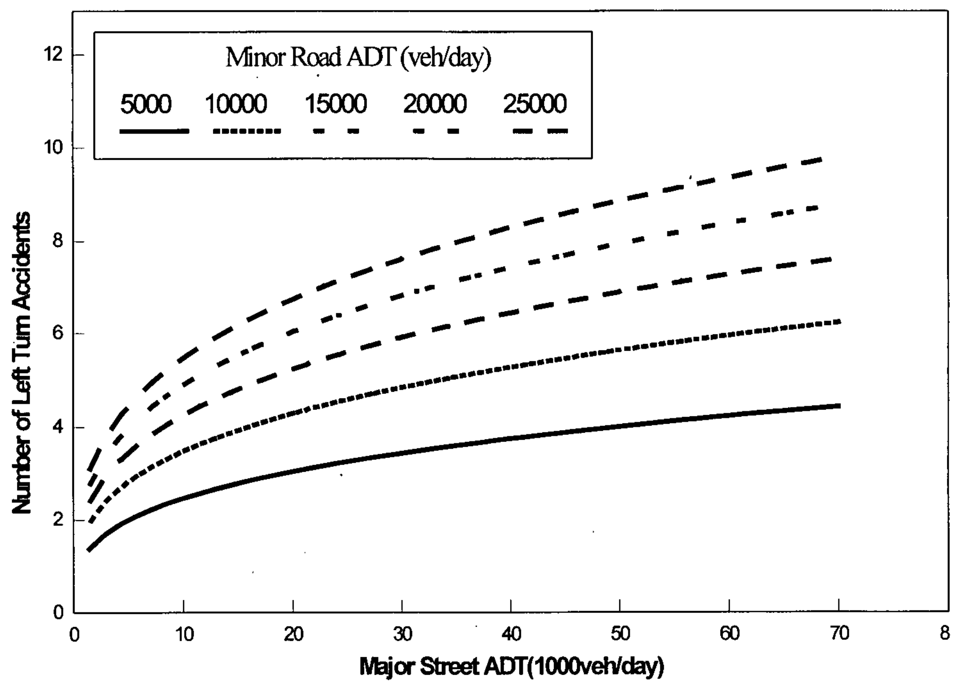

Figure 3.12 Model of left turn accidents

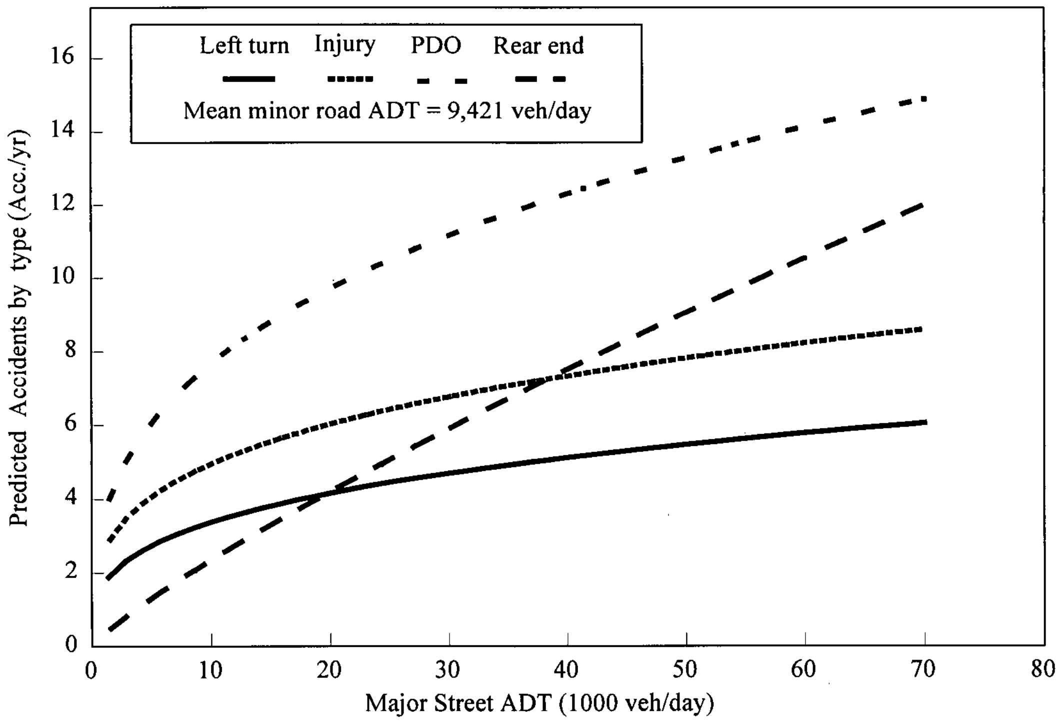

Figure 3.13 Models of different types of Accidents

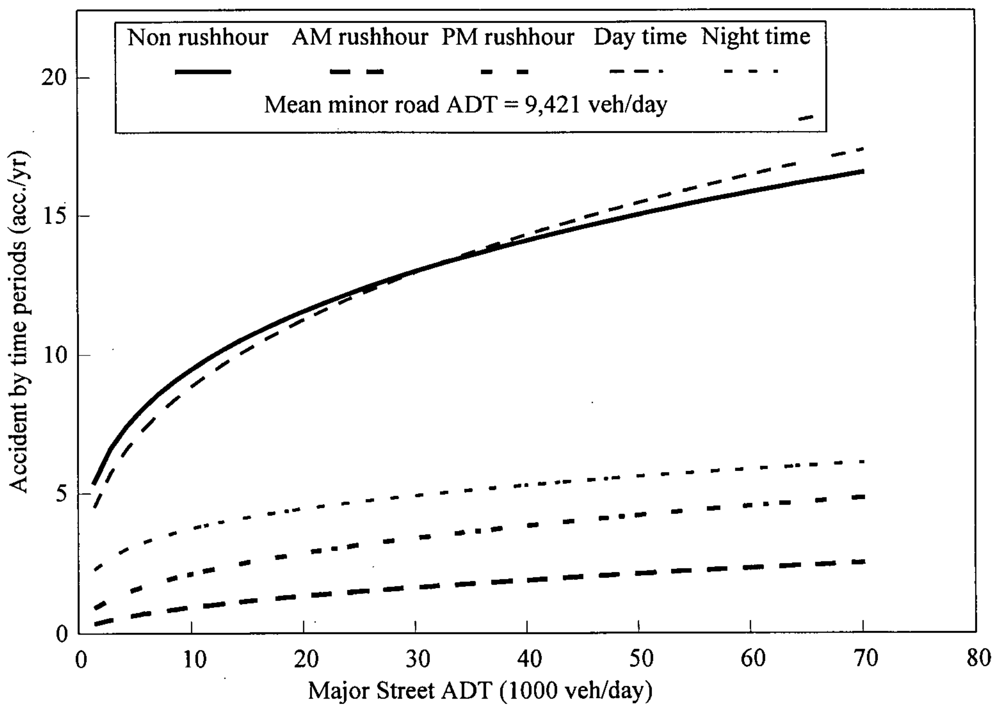

Figure 3.14 Models of time specific accidents

3.2.2.3 Evaluation of the Effect of Intersection Layout Variables

Table 3.6 shows the parameter estimates for a model including intersection layout variables, as well as the significance of the layout variables as given by the t-ratio. The t-ratio is the ratio between the parameter value and its standard error. The critical t-ratio at the 5% significance level is about 1.96. The following layout variables were included in the model:

MANL = Average number of lanes on the major road, both approaches

MINL = Average number of lanes on the minor road, both approaches

LTL = Number of left turn lanes, all approaches

RTL = Number of right turn lanes, all approaches

PC = Number of pedestrian crosswalks

TABLE 3.6 - Parameter Estimates for the Model Incorporating Layout Variables

Model Form	t-ratio		κ	Pearson χ^2
$$Accidents/year = 1.3708 \times \left(\frac{AADT_{\text{major road}}}{1000}\right)^{0.4217} \times \left(\frac{AADT_{\text{minor road}}}{1000}\right)^{0.3857} \times e^{\sum \beta_i x_i}$$ $$\sum \beta_i x_i = 0.0386 MANL + 0.2086 MINL - 0.0264 LTL - 0.0403 RTL - 0.0123 PC$$	*MANL*	0.50	14.42	134.13*
	MINL	3.20		
	LTL	-1.43		
	RTL	-1.09		
	PC	-0.38		

- Denotes significance at a 95-percent confidence level ($\chi^2{}_{0.05,130} = 157.6$)

The variables *PC, MANL, RTL, and LTL* were not significant (t-ratio is less than critical value). A stepwise elimination procedure was used to remove insignificant variables from the model. The procedure involves removing one insignificant variable at a time starting with the one having the least t-ratio. The resulting model is shown in Table 3.7.

TABLE 3.7 Parameter Estimates for the Model Incorporating Layout Variables

Model Form	t-ratio		κ	Pearson χ^2
$Accidents/year = 1.3470 \times \left(\frac{AADT_{\text{major road}}}{1000}\right)^{0.4351} \times \left(\frac{AADT_{\text{minor road}}}{1000}\right)^{0.3720} \times e^{\sum \beta_i x_i}$ $\sum \beta_i x_i = 0.2381\ MINL - 0.03525\ LTL$	*MINL*	4.11	14.16	125.21*
	LTL	-2.04		

* Denotes significance at a 95-percent confidence level ($\chi^2_{0.05;133} = 160.9$)

Figure 3.15 shows the observed versus predicted number of accidents for the model. The dispersion around the 45° line is much less than total accidents model (Figure 3.1) which indicates the significance of including these layout variables.

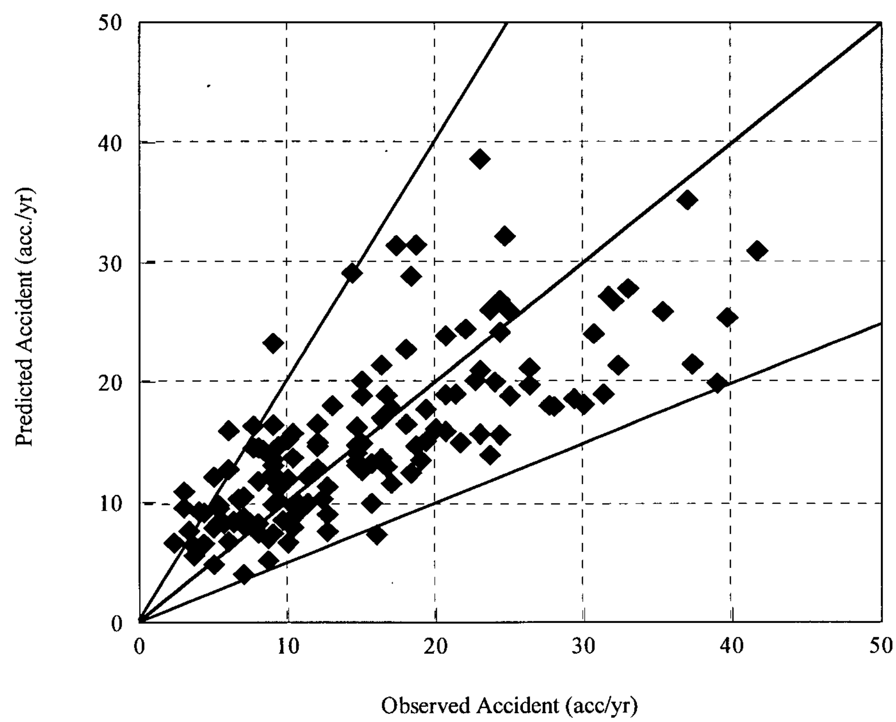

Figure 3.15 Observed number of accidents/yr vs. predicted number of accidents/yr

4.0 APPLICATIONS

4.1 Empirical Bayes Refinements to the Models

As described in section 2.0, the Empirical Bayes (EB) approach can be used to refine the estimate of the expected number of accidents at a location by combining the observed number of accidents at the location with the predicted number of accidents obtained from the GLIM model. The EB estimated number of accidents for any intersection can be calculated using (Hauer et al. ,1988):

$$EB\ safety\ estimate = \alpha . pred + (1-\alpha).count, \tag{4.1}$$

$$\alpha = \frac{1}{1+\dfrac{var(pred)}{pred}} \tag{4.2}$$

where

count = observed number of accidents at the location

pred = predicted number of accidents as estimated from the GLIM model

var(pred) = the variance of the GLIM estimates.

Since $var(pred) = \dfrac{(pred)^2}{k}$ as described earlier, equation 4.1 and 4.2 can be rearranged to yield:

$$EB\ safety\ estimate = \left(\frac{k}{k + pred}\right) pred + \left(\frac{pred}{k + pred}\right) count \qquad (4.3)$$

In addition, the variance of the EB estimate can be calculated using:

$$var\ (EB\ safety\ estimate) = \frac{k\ pred^2}{(k + pred)^2} + \left(\frac{pred}{k + pred}\right)^2 count \qquad (4.4)$$

As can be noted from Equation (4.3), the EB safety estimate lies between the observed number of accidents and the predicted number of accidents, taking into account the individual accident history of the location and the GLIM model prediction which combine data for similar locations. As noted in Section 2, the EB estimate is important since it provides correction for the regression to the mean phenomenon. Figure 4.1 illustrates the EB refinement estimation versus the value predicted from the GLIM model. The following is an example illustrating the use of Equation (4.3):

Assume that an intersection has the following data:

- Major road ADT = 40,000 veh/day
- Minor road ADT = 10,000 veh/day
- Observed accidents/year = 29 acc/yr

Using the first model from Table 3.3, the predicted number of accidents for this intersection is:

$$pred = 2.181 \times (40)^{0.329} \times (10)^{0.442} = 20.27\ acc / yr$$

Using equation 4.3 and 4.4, the empirical safety estimate and its variance can be calculated as:

$$EB\ safety\ estimate = \left(\frac{9}{9 + 20.27}\right) \times 20.27 + \left(\frac{20.27}{9 + 20.27}\right) \times 29 = 26.31\ acc / yr$$

$$var\ (EB\ safety\ estimate) = \frac{9 \times 20.27^2}{(9 + 20.27)^2} + \left(\frac{20.27}{9 + 20.27}\right)^2 \times 29 = 18.22\ (acc / yr)^2$$

In this example the expected number of accidents is reduced from 29 to 26.31 which corresponds to about 9.2 percent regression to the mean correction.

In the following section, two applications for using the GLIM models and the Empirical Bayes' estimates will be discussed.

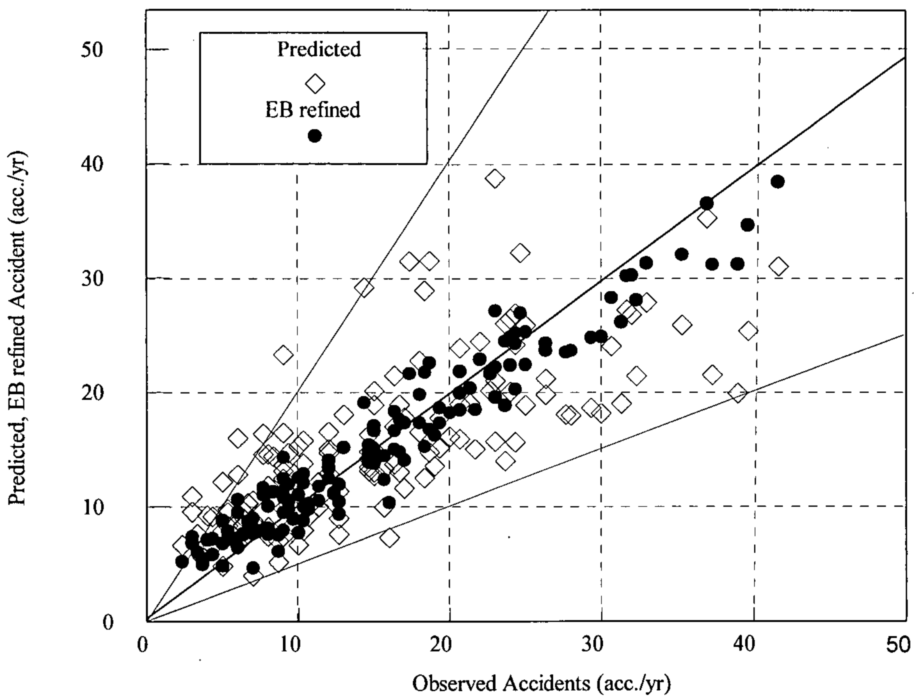

Figure 4.1 Predicted vs. EB Refined number of accidents

4.2 Identification of Accident Prone Location

Accident Prone Locations (APLs) are usually defined as locations which exhibit a significant number of accidents compared to a specific norm. Because of the randomness inherent in accident occurrence, statistical techniques that account for this randomness should be employed when identifying APLs. The EB refinement process discussed in the previous section can be used to identify accident prone locations as follows (Belanger, 1994):

1. Estimate the predicted number of accidents and its variance for the intersection using the appropriate GLIM model and plot the probability density function of the distribution (gamma distribution)

2. Determine the appropriate point of comparison based on the mean and variance values obtained in step (1). (usually the 50^{th} percentile, P_{50} is used as a point of comparison)

3. Calculate the EB safety estimate and its variance from equations 4.3 and 4.4 and plot the distribution.

4. Identify the location as accident prone if there is a significant probability that the intersection's safety estimate exceeds the P_{50} value.

To continue with the previous example, the predicted number of accidents and its variance using model (1) can be calculated as 20.27 acc/yr and 45.65 $(acc/yr)^2$ respectively. The P_{50} value can be estimated from plotting the probability density function of the gamma distribution. From Figure 4.2, P_{50} = 19.53 acc/yr. The EB safety estimate and its variance were calculated as 26.31 acc/yr and 18.22 $(acc/\ yr)^2$ respectively. The updated distribution is also shown in Figure 4.2. From the figure, it can be shown that the probability of having accidents less than P_{50} is only 4.5 percent (the shaded area in the figure). This means that there is a significant probability (95.5%) of exceeding the P_{50} value and the intersection can be considered accident prone.

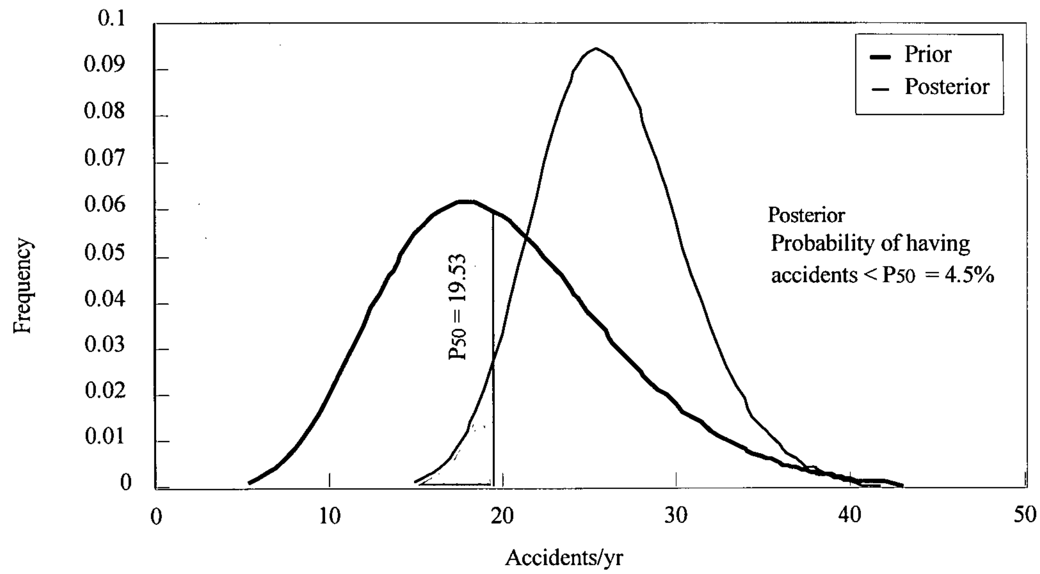

Figure 4.2 Identification of Accident Prone Locations

To facilitate the process of identifying accident prone locations, critical accident frequency curves for different significance levels can be developed. An example of these curves is shown in Figure 4.3 using the κ value for the model relating the total number of accidents to traffic flows ($\kappa = 9.0$). To illustrate, using the same example as before, for a predicted accident value of 30 accidents/ year, the observed number of accidents at the intersection must exceed 43 accidents/ year to be identified as accident prone at the 99 % level, 39 accidents/ year at the 95% level, and 36 accidents/ year at the 90% level of confidence.

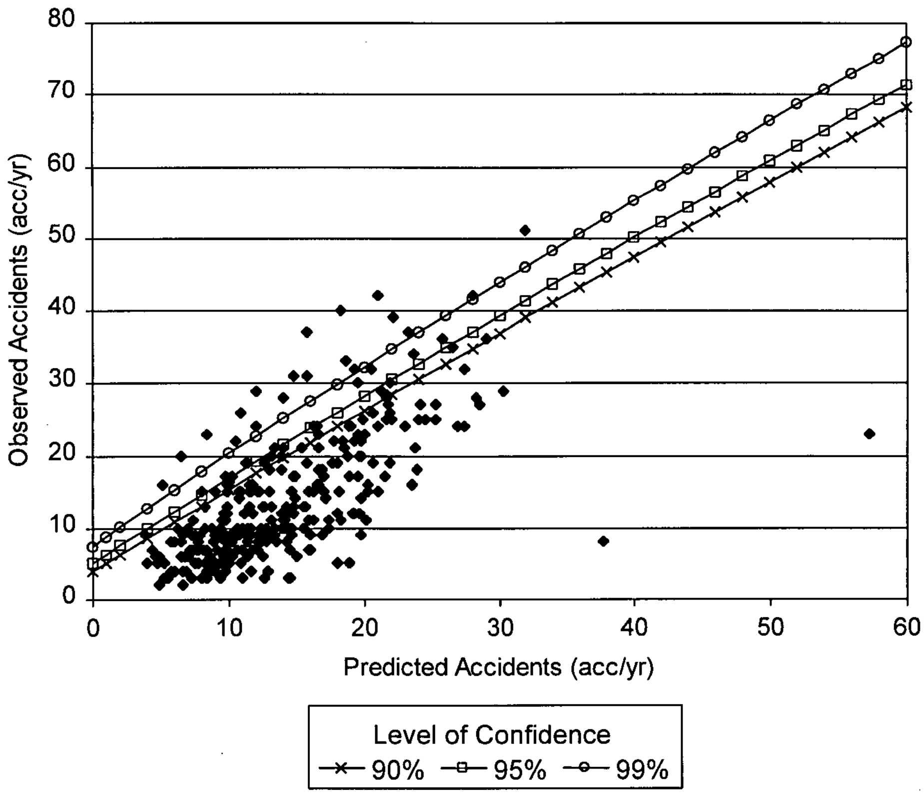

Figure 4.3. Accident Prone Locations for Total Model

4.3 Before and After Safety Evaluation

The effect of a safety measure is often studied by comparing the number of accidents observed after the implementation of the measure to the expected number of accidents had the measure not been implemented. In simple before and after studies, the observed number of accidents in the period before the implementation is used to estimate the latter value. However, because of the random variations in accident occurrence (e.g. the regression to the mean effect), the observed number of accidents before the implementation may not be a good estimate of what would have happened had no measure been implemented. An alternative and more accurate approach is to use the EB refinement process discussed in Section 4.1.

Considering the same example as before, assume that a specific safety measure to reduce the number of accidents at the intersection was implemented. The observed number of accident in one year after the implementation is 20 accidents. The effectiveness of the measure can then be calculated as:

$$Measure\ of\ Effectiveness = 1 - \frac{20}{26.31} = 0.24$$

which indicates a 24 percent reduction in total accidents because of the treatment. If the data on specific accidents types is available, then the measure of effectiveness in reducing specific accident types can be estimated by applying the appropriate GLIM model from Table 3.4.

4.4 Safety Planning

Accident prediction models can be used in safety planning by identifying the traffic and geometric variables that have the most impact on the safety performance of signalized intersections. These variables should be the focus of road authorities attention and investment. As well, the models can be used to estimate the incremental safety benefit associated with the change in any traffic or geometric design variable.

For example, a sensitivity analysis was carried out for the variables included in the model shown in Table 3.6. Figure 4.4 shows the analysis in a non-dimensional form. For the variables examined, accident occurrence is found to be the most sensitive to the number of lanes of minor road followed by the total and minor road traffic volumes and the number of left turn lanes, respectively. An increase in the first three variables will increase the expected number of accidents while an increase in the number of left turn lanes would cause a decrease in the expected number of accidents.

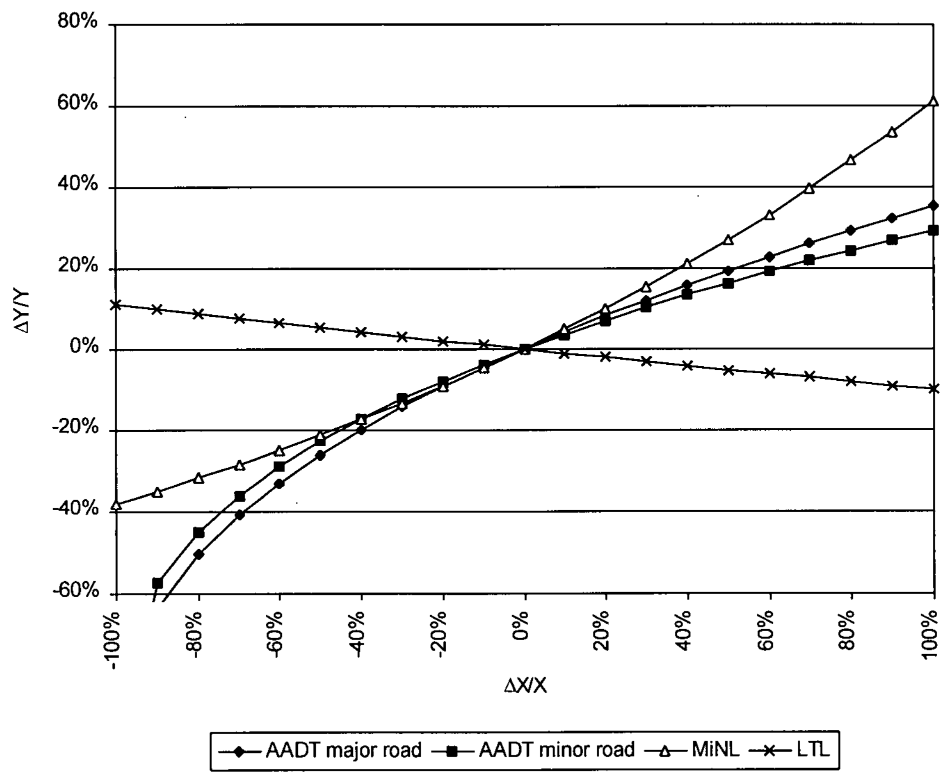

Figure 4.4 Sensitivity Chart

5.0 CONCLUSION

This thesis has documented the results of a study to develop accident prediction models for signalized intersections in the Greater Vancouver Regional District (GVRD). To avoid the shortcomings associated with the conventional linear regression approach, the models were developed using the Generalized Linear Regression Models (GLIM) approach. The GLIM approach addresses the error structure problems that are associated with the conventional linear regression theory and allows for the use of nonlinear relationships in the model.

Three types of models were developed: models for the total number of accidents, models for specific accident types, and models which account for variables other than traffic volumes. The models provided adequate goodness of fit for the accident data used. An Empirical Bayes procedure was used to refine the estimates of the GLIM models to provide more accurate site-specific safety estimates. Application of the Empirical Bayes procedure included the identification of accident prone locations and performing before and after safety analysis.

There are several improvements which can enhance the models developed in this thesis. First, the sample size used in this study is relatively small (139 intersections). Therefore, expanding the sample size to include more intersections and re-calibrating the models is recommended.

Secondly, in developing the models for predicting specific accident types, the average daily traffic volumes on the major and minor roads were used. Hauer (1989) and others recommended the use of traffic flows related to the specific accident type. However, detailed traffic volume data were not available for this research. Therefore, collecting detailed traffic volume data and developing models relating accident types to the traffic flows generating these accidents may be of interest.

Thirdly, all signalized intersections included in this research are in the urban environment. The collection of data for other environments (e.g. rural, suburban) and the development of corresponding models are also suggested.

REFERENCES

1) Al-Bakri, A., Conaway, M. R. and Stoner, J., (1991). "Conditional analysis of accidents at four-approach traffic circles", Transportation Research Record, 1327, Transportation Research Board, National Research Council, Washington, D.C., pp. 42-46

2) Al-Senan, S. H. And Wright, P. H., (1987). "Prediction of head-on accident sites", Transportation Research Record, 1122, Transportation Research Board, National Research Council, Washington, D.C., pp. 79-85

3) Bélanger, C., (1994). "Estimation of safety of four-legged unsignalised intersections", Transportation Research Record, 1467, Transportation Research Board, National Research Council, Washington, D.C., pp. 23-29

4) Bonneson J. A. and McCoy P. T., (1993). "Estimation of safety at two-way stop-controlled intersection on rural highways", Transportation Research Record, 1401, Transportation Research Board, National Research Council, Washington, D.C., pp. 83-89

5) Bowman, B. L. , Vecellio, R. L. and Miao, J., (1995). "Vehicle and Pedestrian accident models for median locations", Journal of Transportation Engineering. /Nov./Dec. , pp. 531-537

6) Brüde, U. and Larsson, J., (1988). "The use of prediction models for eliminating effects due to regression-to-the-mean in road accident data", Accident Analysis and Prevention, Vol. 20, No. 4, pp. 299-310

7) Coleman J., and Stewart, G.R., (1976). "Investigation of accident data for railroad-highway grade crossings", Transportation Research Record 611. Transportation Research Board, pp. 60-67

8) Davis, S. E., Robertson, D. H. and Ellisking, L., (1989). "Pedestrian/Vehicle conflicts: an accident prediction model", Transportation Research Record, 1210, Transportation Research Board, National Research Council, Washington, D.C., pp. 1-11

9) Forkenbrock, D. J., Foster, N. S. J. and Pogue, T.F., (1994). "Safety and highway investment", Midwest Transportation Centre, Iowa State University and University of Iowa,

10) Garder, P., Leden, L. and Thedeen, T., (1994). "Safety implications of bicycle paths at signalised intersections", Accident Analysis and Prevention. Vol. 26. No. 4, pp. 429-439

11) Golias, J. C., (1992). "Establishing relationships between accident and flows at urban priority road junctions", Accident Analysis and Prevention, Vol. 24, No. 6, pp. 689-694

12) Hanna, J. T., Flynn, T. E. and Tyler, W. K., (1976). "Characteristics of intersection accidents in rural municipalities", Transportation Research Record 601, Transportation Research Board, pp. 79-82

13) Hauer, E., (1994). "The effect of resurfacing on the safety of two-lane rural roads in New York State", Paper 940541. Transportation Research Board, 73rd Annual Meeting. January 9-13, 1994, Washington D.C.

14) Hauer, E. and Papaioannou, p. (1991). "Prediction in road safety studies: an empirical inquiry", Accident Analysis and Prevention, Vol. 23, No. 6, pp. 595-607

15) Hauer, E., Ng J. C. N. and Lovell, J., (1989). "Estimation of safety at signalised intersection", Transportation Research Record, 1185, Transportation Research Board, National Research Council, Washington, D.C., pp. 48-61

16) Hauer, E. , (1992). "Empirical Bayes approach to the estimation of 'unsafety': The multivariate regression method", Accident Analysis and Prevention, Vol. 24, No. 5, pp. 457-477

17) Kulmala, R., (1992). "Prediction models for accident at highway junction", ITE 1992 Compendium of Technical Papers

18) Kulmala, R., (1994). "Measuring the safety effect of road measures at junctions", Accident Analysis and Prevention, Vol. 26. No. 6, pp. 781-794

19) Lau, M., May, A. D., and Smith, R. N., (1989). "Applications of accident prediction models". Transportation Research Record, 1238, Transportation Research Board, National Research Council, Washington, D.C., pp. 20-30

20) Lau, M. and May A. D., (1988). "Injury accident prediction models for signalised intersections", Transportation Research Record, 1172, Transportation Research Board, National Research Council, Washington, D.C., pp. 58-67

21) Lin, T. , Jovanis, P. P. and Yang, C., (1994). "Time of day models of motor carrier accident risk", Transportation Research Record, 1467, Transportation Research Board, National Research Council, Washington, D.C., pp. 1-7

22) Maher, M. J., (1991). "A new bivariate negative binomial model for accident frequencies", Traffic Engineering and Control, Vol. 32, pp. 422-423

23) Margiotta, R. and Chatterjee, A., (1995). "Accident on suburban highways-Tennessee's experience", Journal of Transportation Engineering, /May/June, pp. 255-261

24) McCullagh, P., and J. A. Nelder. (1983). "Generalized linear models", The University Press, New York.

25) Miaou, S. and Lum, H., (1993). "Modeling vehicle accident and highway geometric design relationships", Accident Analysis and Prevention, Vol. 25, No. 6, pp. 689-709

26) Miaou, S., Hu, P. S. and Wright, T., (1992). "Relationship between truck accidents and highway geometric design: A Poisson regression approach", Transportation Research Record, 1376 Transportation Research Board, National Research Council, Washington, D.C., pp. 10-18

27) Nembhard, D. A. and Young, M. R., (1995). "Parametric empirical estimates of truck accident rates", Journal of Transportation Engineering, /July/August, pp. 359-363

28) Persaud, B., Hauer, E., Vallurupalli, R. and Mucsi, K., (1995) "Study of the safety effect of removing traffic signals in Philadelphia", Draft Final Report, October 1995

29) Persaud, B. and Dzbik, L., (1993). "Accident prediction models for freeways", Transportation Research Record, 1401, Transportation Research Board, National Research Council, Washington, D.C., pp. 55-60

30) Persaud, B. N. and Mucsi, K., (1994). "Microscopic Accident potential models for two-lane rural roads", Transportation Research Record, 1485, Transportation Research Board, National Research Council, Washington, D.C., pp. 134-139

31) Persaud, B. N., (1991). "Estimating accident potential of Ontario road section", Transportation research record, 1327, Transportation Research Board, National Research Council, Washington, D.C., pp. 47-53

32) Pursula M. and Pastinen V., (1993). "A Bayesian approach to update traffic flows from traffic counts", Transportation and Traffic Theory. /C.F. Daganzo(Editor). pp. 507-522

33) Seneviratne, P. N. and Islam M. N., (1992). "Speed estimates for roadway design and traffic control", Transportation Research Record, 1375, Transportation Research Board, National Research Council, Washington, D.C., pp. 37-43

34) Stark, D. C., (1994). "Relating speed to accident risk at priority junctions", Traffic Management and Road Safety: Proceedings of Seminar J., Held at the PTRC European Transport Forum, University of Warwick, England Planning and Transport Research and Computation Co. Meeting (22nd: 1994: Warwick, England)

35) Tarko, A. and Tracz, M., (1995). "Accident prediction models for signalised crosswalks", Safety Science, pp. 109-118

36) Webb, G. M., (1955). "The relation between accidents and traffic volumes at signalised intersections", ITE Proceedings, , pp. 149-167

37) Zegeer, C. V., Hummer, J., Reinfurt, D., Herf, L., and Hunter, W., (1986). "Safety effects of cross section design for two lane roads", Volume I, Final Report, FHWA-RD-87/008. Federal Highway Administration, Transportation Research Board, Dec.1986.

APPENDIX A

Accident data

Vancouver Data

LOCATIONS	total volume	major volume	minor volume	major number of lanes	minor number of lanes	# of ltl	# of rtl	# of ped lanes	TOTAL
Abbott Street/Hastings	14188	11214	2974	2.5	1	0	0	4	24
Boundary Road/Hastings	38420	27709	10711	2	1.5	2	1	4	51
Cambie Street/Hastings	19723	16179	3544	2.5	2	0	0	4	24
Campbell Avenue/Hastings	24023	22083	1940	3	1	0	0	0	32
Carrall Street/Hastings	15627	14024	1603	2.5	1.5	0	1	4	48
Cassiar Street/Hastings	45251	34086	11165	3	2	5	4	4	36
Clark Drive/Hastings	36174	26672	9502	3	2.5	2	0	4	112
Columbia Street/Hastings	19855	15328	4527	3	0	0	0	4	38
Gore Avenue/Hastings	21556	19438	2118	3	1	0	0	4	29
Granville Street/Hastings	15444	13120	2324	2	0.5	3	1	4	11
Homer Street/Hastings	13532	11004	2528	2.5	1.5	0	0	4	15
Hornby Street/Hastings	15148	10414	4734	2	0	1	1	4	11
Howe Street/Hastings	19723	16179	3544	2	1	3	0	4	16
Kootenay Street/Hastings	31838	30396	1442	3	1	0	0	2	27
Main Street/Hastings	24419	16425	7994	3	3	0	0	4	79
Nanaimo Street/Hastings	41940	28268	13672	3	3	3	0	4	72
Penticton Street/Hastings	27744	26644	1100	3	1	0	0	4	19
Renfrew Street/Hastings	36950	28267	8683	3	3	2	0	4	92
Richards Street/Hastings	13669	8587	5082	2.5	1	0	0	4	24
Seymour Street/Hastings	20054	10477	9577	3	2	0	0	4	23
Skeena Street/Hastings	33869	31273	2596	3	1	0	0	2	51
Victoria Street/Hastings	30957	26347	4610	3	1.5	0	0	2	44
Arbutus Street/broadway	24249	14022	10227	3	2	2	0	4	44
Ash Street/broadway	22023	20557	1466	3	1	0	0	4	31
Blenheim Street/broadway	12694	10558	2136	2	1	0	0	2	18
Burrard Street/broadway	43565	19358	24207	3	3	2	0	4	52
Clark Drive/broadway	46336	24468	21868	3	3	4	0	4	125

LOCATIONS	total volume	major volume	minor volume	major number of lanes	minor number of lanes	# of ltl	# of rtl	# of ped lanes	TOTAL
Commercial Drive/broadway	40603	28229	12374	3	3	3	0	4	119
Fir Street/broadway	28096	15838	12258	3	2	1	0	4	39
Fraser Street/broadway	27594	21978	5616	3	2	3	0	4	57
Glen Drive/broadway	26334	24118	2216	3	1	0	0	2	31
Granville Street/broadway	43344	23779	19565	3	3	4	0	4	43
Heather Street/broadway	24956	21641	3315	3	1	0	0	4	28
Kingsway/broadway	45686	24017	21669	3	3	2	0	4	56
MacDonald Street/broadway	22217	12014	10203	2	2	0	0	4	44
Nanaimo Street/broadway	51056	33848	17208	3	3	0	0	4	111
Oak Street/broadway	21460	12213	9247	3	2	5	0	4	36
Renfrew Street/broadway	42255	29487	12768	3	2	0	0	0	73
Rupert Street/broadway	43025	30167	12858	3	2	0	0	0	66
Willow Street/broadway	24582	21738	2844	3	1	0	0	4	16
Yukon Street/broadway	26604	21428	5176	3	2	1	0	0	45
Commercial Dr/Powell St	29165	28009	1156	2	1	0	0	4	13
Commercial Dr/Venables St	21553	12817	8736	2	2	0	0	0	18
Commercial Dr /E Hasting St	35334	30572	4762	3	1	4	0	0	55
Commercial Dr./N Grandview Hwy	16269	11408	4861	2	2	2	1	4	21
Kingsway / Victoria Dr.	49778	35765	14013	2.5	1	2	0	4	117
E 12TH Ave / Commercial Dr.	45164	32940	12224	2	2	0	0	0	75
E 1ST Ave/ Commercial Dr.	44580	33080	11500	2	2	4	0	4	69
E 33RD Ave / Victoria Dr	29776	16138	13638	2	2	4	0	0	64
E 41ST Ave / Victoria Dr	44694	26057	18637	2.5	2	4	0	4	96
E 54TH Ave / Victoria Dr	27495	13837	13658	2	1	1	3	2	44
E Broadway St/Commercial Dr.	41492	28888	12604	3	3	2	0	4	99
SE Marine Dr. / Victoria Dr.	56707	52861	3846	3	1	2	0	4	69
Cambie Dr. / Pacific Bl.	15616	14571	1045	3	0	3	1	4	21
Cambie Dr. / Smithe St.	27005	20058	6947	2	1	0	0	4	27
Cambie Dr. /W King Edward Ave	58173	34770	23403	3	2	4	2	4	71
W 12TH Ave / Cambie St.	59284	36028	23256	3	2	5	0	4	73
W 16TH Ave / Cambie St	43029	33684	9345	3	2	4	1	4	45

LOCATIONS	total volume	major volume	minor volume	major number of lanes	minor number of lanes	# of ltl	# of rtl	# of ped lanes	TOTAL
W 29TH Ave / Cambie St	39765	32578	7187	3	1	5	2	4	34
W 33RD Ave / Cambie St	32260	28274	3986	3	1.5	1	0	4	44
W 41ST Ave/ Cambie St	51884	30997	20887	3.5	3	8	1	4	55
W 43RD Ave / Cambie St.	28981	28678	303	3.5	0.5	2	1	4	15
W 45TH Ave / Cambie St.	29525	28206	1319	3	1	2	2	4	27
W 49TH Ave / Cambie St.	45415	29997	15418	3	2	5	0	4	79
W 57TH Ave / Cambie St.	34223	31457	2766	2.5	1	1	0	4	28
W 7TH Ave / Cambie St.	43715	38229	5486	3	1	2	0	4	56
W Broadway St. / Cambie St.	52356	28810	23546	3.5	3	2	1	4	74

Richmond Data

LOCATIONS	total volume	major volume	minor volume	major number of lanes	minor number of lanes	# of ltl	# of rtl	# of ped lanes	TOTAL
1050	25526	14154	11372	2	1.5	2	1	4	49
1100	23030	17034	5996	2	1	6	0	4	47
1130	20150	19274	876	2	1	2	0	4	7
1200	26332	18134	8198	2	2	0	0	4	49
1300	21783	12733	9050	2	1	0	2	4	28
1550	31804	21146	10658	2	1	3	0	4	24
1600	16992	9329	7663	1	1	3	0	4	16
1650	20445	12478	7967	2	1.5	0	1	4	27
1700	26608	14862	11746	2	1.5	0	1	4	73
1805	27102	24811	2291	2	1	2	1	4	12
2050	36323	23128	13195	2	2	2	0	4	49
2100	25316	15837	9479	2	2	4	0	4	58
2150	29247	21596	7651	2	2	0	0	4	58
2250	51891	31929	19962	2	2	7	2	4	54
2300	57552	35566	21986	2	2	8	1	4	106
2305	21871	18382	3489	2	0	4	1	4	13
2374	41653	37107	4546	2.5	1	4	3	4	9
2550	30263	26345	3918	2	1.5	3	0	4	15
2600	23533	12531	11002	2	2	4	0	4	65
2650	25321	15758	9563	2	2	2	0	4	23
2700	36488	19368	17120	2	2	6	0	4	72
2750	44602	23828	20774	2	2	7	3	4	45
2756	43795	29920	13875	2	1.5	7	3	4	60
2800	43873	22627	21246	2	2	8	0	4	27
2806	51587	31536	20051	2	2	8	1	4	62
2821	37094	20463	16631	2	1.5	7	1	4	54
2896	25900	19762	6138	1.5	0	4	0	4	21
2922	55670	54986	684	2	0	3	1	4	10

LOCATIONS	total volume	major volume	minor volume	major number of lanes	minor number of lanes	# of ltl	# of rtl	# of ped lanes	TOTAL
2933	52877	51988	889	2.5	0	4	0	4	21
2954	121892	67836	54056	3	1.5	8	2	4	69
3050	33532	26560	6972	2	1.5	2	1	4	36
3100	35769	22618	13151	2	1	6	1	4	71
3150	33553	24438	9115	2	2	2	0	4	50
3205	21800	14311	7489	2	1.5	1	0	4	50
3250	60585	30335	30250	2	2	8	2	4	95
3270	34705	26937	7768	2	1.5	3	1	4	30
3280	37150	25991	11159	2	1.5	4	1	4	62
3285	23919	14959	8960	2	2	0	0	4	27
3305	43762	26183	17579	2	2	8	2	4	68
3325	28474	15475	12999	2	1.5	2	0	4	31
3335	21203	11349	9854	2	1.5	4	0	4	30
3423	36288	30350	5938	2	1	3	0	4	31
3600	18681	15277	3404	1	1	2	1	4	22
3650	17907	9990	7917	2	2	0	1	4	45
3700	26660	14697	11963	2	2	0	0	4	88
3780	31384	27247	4137	2	0	3	1	4	19
3830	35697	28709	6988	2	0	4	2	4	9
3950	25240	22330	2910	2	0	2	2	4	30
4050	33027	27141	5886	2	1.5	3	0	4	28
4150	23105	19944	3161	2	1.5	2	1	4	20
4200	31100	19980	11120	2	2	0	1	4	62
4250	24677	19986	4691	2	1	0	2	4	20
4350	48675	34171	14504	2	1.5	5	2	4	90
4400	30932	19782	11150	2	2	2	0	4	84
4450	32237	23536	8701	2	0	4	1	4	32
4550	33623	28867	4756	2	1.5	3	1	4	18
4800	37708	34006	3702	2	0	4	1	4	38
4850	50856	39576	11280	2	2	4	4	4	75
4900	39957	27440	12517	2	1	8	1	4	36

LOCATIONS	total volume	major volume	minor volume	major number of lanes	minor number of lanes	# of ltl	# of rtl	# of ped lanes	TOTAL
4950	29992	23899	6093	2	1.5	2	1	4	45
4953	26389	22993	3396	2	1	0	2	4	34
5050	43559	31737	11822	2	2	3	0	4	97
5200	22918	14286	8632	2	2	4	0	4	25
5300	49938	37731	12207	2	2	6	4	4	83
5400	32677	18623	14054	2	2	4	0	4	94
5458	35975	29709	6266	2	1	4	3	4	38
5800	40003	31212	8791	2	1	3	2	4	27
5805	38162	28077	10085	1.5	1	4	3	4	47
5900	8953	5702	3251	2	1	2	2	4	26
5909	15804	11828	3976	2	1	2	2	4	26
6300	24940	16890	8050	2	1	6	1	4	37
6400	17988	14872	3116	2	1.5	3	1	4	17

Appendix B MODELS FOR VANCOUVER AND RICHMOND INTERSECTIONS

Total of 67 intersections in Vancouver were obtained for analysis.

Model	Model Form	Coefficient Estimates		SD	k	Chi-square
1. Bonneson and McCoy	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	0.9 0.6894 0.2738	101.8	14.65	59.74
2.Brude and Larsson	$E(m) = a_0 \times (V_1+V_2)^{a_1} \times (\frac{V_2}{V_1+V_2})^{a_2}$	a_0 a_1 a_2	0.9 0.9630 0.2738	101.8	14.56	59.74
3. Kulmala	$E(m) = a_0 \times V^{a_1} \times V_2^{a_2} \times e^{\sum \beta_i x_i}$	a_0 a_1 a_2 β_1 β_2 β_3 β_4 β_5 β_6 β_7 β_8 β_9 β_1	0.7 0.5617 0.1779 0.1532 0.07 0.04 0.1638 0.08 0.18 0.14 -0.06 -0.08 -0.08	62	11298	53.98
4. PDO Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	0.7 0.6117 0.4117	59.7	16.43	--
5.Injury Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	0.6 0.4958 0.408	72.1	--	54.41
6. Day time Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	0.7 0.6681 0.377	69.7	12.34	69.3
7. Night time Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	0.7 0.3302 0.5007	73	--	66.43
8. Rear-end Accidents	$E(m) = a_0 \times V^{a_1}$	a_0 a_1	0.1 1.296	67.3	16.66	123.17
9. Angle Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	1.3 0.04 0.3697	44.6	65.52	--
10 AM Rush hour accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	0.1 0.9178 0.3373	24.1	--	--

Model	Model Form	Coefficient Estimates		SD	k	Chi-square
11 PM Rush hour accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	0.1 0.8566 0.3318	41.3	--	--
12 Rush Hour Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	0.33 0.4995 0.4461	41.3	--	--
13 Non Rush Hour Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	1.2 0.4605 0.4399	82.1	23.6	57.34
14 Left turn accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	0.1 1.048 0.496	71.9	4.08	46.07

* "--" means the indicated value is a minus figure which shows a not proper form for the corresponding model.

Total of 72 intersections in Richmond city were obtained for analysis.

Model	Model Form	Coefficient Estimates		SD	k	Chi-square
1.Bonneson and McCoy	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	2.0 0.1537 0.6826	127.8	14.65	64.06
2. Brude and Larsson	$E(m) = a_0 \times (V_1 + V_2)^{a_1} \times (\frac{V_2}{V_1 + V_2})^{a_2}$	a_0 a_1 a_2	1.7 0.8412 0.6336	127.6	14.56	64.06
3. Kulmala	$E(m) = a_0 \times V^{a_1} \times V_2^{a_2} \times e^{\sum_i^n \beta_i x_i}$	a_0 a_1 a_2 β_1 β_2 β_3 β_4 β_5 β_6 β_7 β_8 β_9 β_{10}	0.3 0.4033 0.2915 0.1536 0.018 0.093 0.1696 0.062 -0.156 0.1233 -0.017 -0.1 -0.026	73.6	11298	70.94
4. PDO Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	1.0 0.3869 0.5141	100.7	16.43	65.15
5. Injury Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	0.7 0.3341 0.5148	54.48	--	--
6. Day time Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	1.2 0.3898 0.4771	115.6	12.34	62.37
7. Night time Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	0.5 0.2985 0.5743	50.24	--	--
8. Rear-end Accidents	$E(m) = a_0 \times V^{a_1}$	a_0 a_1	0.03 1.358	74.9	16.66	52.05
9. Angle Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	2.0 -0.235 0.5018	65.87	65.52	123.81
10. AM Rush hour accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	0.1 0.4843 0.4664	43.42	--	--

Model	Model Form	Coefficient Estimates		SD	k	Chi-square
11 PM Rush hour accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	0.2 0.5095 0.4272	43.42	--	--
12 Rush Hour Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	0.3 0.4995 0.4461	60	--	--
13 Non Rush Hour Accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	1.3 0.3368 0.5256	91.55	23.6	61.02
14 Left turn accidents	$E(m) = a_0 \times V_1^{a_1} \times V_2^{a_2}$	a_0 a_1 a_2	0.3 0.4518 0.577	112.2	4.08	4.08