Accident Prediction Models

Transfund New Zealand Research Report No.192

Accident Prediction Models

Shane Turner

Beca Carter Hollings & Ferner

Transfund New Zealand Research Report No. 192

ISBN 0-478–25050-9
ISSN 1174–0574

PO Box 2331, Lambton Quay, Wellington, New Zealand
Telephone 64-4-473-0220; Facsimile 64-4-499-0733

Turner, S. A., 2001. Accident prediction models. *Transfund New Zealand Research Report 192.* 80 pp.

Keywords: accidents, economic evaluation, prediction models, test statistics, rural and urban, intersections, mid-block sites, Bayesian methods

An Important Note For The Reader

The research detailed in this report was commissioned by Transfund New Zealand.

Transfund New Zealand is a Crown entity established under the Transit New Zealand Act 1989. Its principal objective is to allocate resources to achieve a safe and efficient roading system. Each year Transfund New Zealand invests a portion of its funds on research that contributes to this objective.

Contents

cont.

Executive Summary

The prediction of accident rates on road links and at intersections is an important input to road safety improvement. Without some knowledge of the relationships between road, traffic, and environmental factors and accident causation, it is not possible to gauge the effects of remedial measures.

The main objective of the study was to develop accident prediction models that can be used for economic evaluation of roading projects, and more specifically for network wide studies.

While models already exist in the Transfund Project Evaluation Manual, these models are rather simplistic in form, not relating well to accident causing mechanisms, and have been developed using normal linear regression, which is not appropriate for accident data. The new models detailed in this report use the latest statistical analysis techniques (generalised linear regression) and relate particular accident collision types to their associated conflicting flow movements.

Models have been developed for a number of rural and urban intersection and 'mid-block' sites including; traffic signal, roundabout, priority and uncontrolled intersections and strategic (arterials, motorways and highways) and local (collector and local) routes. National sample sets were collected for each intersection types, as summarised in Table 1.

Table 1 - Model Sample Sizes

Site Type	Sample Size	Site Type	Sample Size
Signalised X-Roads	109	Rural T-Junctions	50
4-Arm Roundabouts	55	Urban Arterials	165
Priority X-Roads	76	Urban Collectors	149
Signalised T-junctions	30	Urban Local	119
Priority T-junctions	89	Rural Highways	76
Uncontrolled T-Junctions	51	Motorways	100

The data was collected from a variety of sources, including local authorities, Transit New Zealand and consultants, and from different parts of New Zealand.

The models have either one or two predictor variables. Two variable models have been used for accidents involving two or more vehicles, where the two vehicles come

from different approaches. For single-vehicle, mid-block and 'same approach' (eg. rear-end) accidents a single predictor variable has been used. Models have the following form:

$$A = b_0 \cdot q_1^{b1} \cdot (q_2^{b2})$$

For intersection models 'A' is the number of accidents, of each type, in five years on each approach, and q_1 and q_2 are the 24-hour turning volume movements (right turn, straight through or left turn), the approach or circulating flow (the latter for roundabouts only). On links, the traffic flow variable is the one or two-way link volume (AADT). A full description of the models and the parameter values can be found in the main body of the text.

Prediction models have also been produced using the link flows, for situations where turning volume counts are unavailable. In these models (product-of-link-flow models) the predictor variables are the two-way traffic volumes on each of the links that intersection. For crossroads the first variable (Q_{major}) is the highest of the link flows and Q_{minor} is the flow on the other intersection route. For T-junctions the first predictor variable is always the through road and second is the stem volume, irrespective of the magnitude of the flows. The parameter values for each of these models can be found in the main body of the text.

Goodness-of-fit statistics (scaled deviance) have been calculated for all the models produced (Table 5.24). The majority of the models were found to be statistically significant to the 95% level of confidence. Confidence intervals were also produced to show the flow ranges over which the predictions were most accurate. This varies according to the intersection type, with signalised junction models generally being more accurate at high flows and uncontrolled junctions being more accurate at lower flows.

A proposed change to Appendix A6, the accident analysis section, of the Transfund PEM has been drafted, and appears in Appendix C. Further discussion, on and development of, this procedure is required, particularly in terms of the trade-off between level of complexity and level of accuracy of the predictions. The proposed procedure adds considerable complexity to the accident analysis procedure, but also provides a lot more information on expected accident occurrence (by accident type) to analysts and Transfund, therefore allowing more accurate predictions of accident costs.

It has been shown that the empirical Bayes method, which makes use of both accident prediction models and historical accident data has merit, and can be used to take into account local factors at particular crash sites. This is an important consideration at sites where factors such as geometric deficiencies (steep grades and acute angle approaches) are prevalent, and models based only on flows are available. Further development and testing (case studies) of this procedure is required prior to it being introduced into the Transfund PEM.

Abstract

Accident prediction models have been developed for rural and urban intersections and links. This includes models for traffic signals, roundabouts, priority and uncontrolled intersections, rural highways, motorways and urban arterial, collector and local streets. Accident (1995 to 1999) and flow data at over 1000 sites throughout New Zealand were used to develop the models. Test statistics have been prepared showing the goodness-of-fit and confidence intervals of the models. Application of the models in economic evaluation has been discussed along with the changes that would need to be made to the Project Evaluation Manual to incorporate the models.

Acknowledgements

The author wishes to acknowledge the valuable assistance received from:

Professor Graham Wood, Massey University
Associate Professor Alan Nicholson, Canterbury University
Mr Ian Bone, Beca Carter Hollings & Ferner, Auckland
Mr Ian Melsom, Transfund
Mr Calum Nicholson, Research Assistant

1 Introduction

1.1 Background

During 1997 there were 13,378 reported injuries resulting from accidents in New Zealand and 540 deaths (LTSA, 1998). Since studies of hospital records have shown that approximately only half of all injury road accidents are reported to the New Zealand Police (LTSA, 1994), the true number of injuries is much higher than this.

A comparison with other OECD countries shows that New Zealand injury accident and death rates are higher than most. For example, in deaths per 100,000 population, out of 24 OECD countries, New Zealand ranks fifth highest (LTSA, 1994), with over twice the number of deaths of Sweden, which has the lowest rate. In light of this comparison it is understandable that there has been, and will continue to be, a lot of emphasis placed on road safety and measures, be they engineering, behavioural or enforcement, to reduce crash occurrence.

Many different organisations and occupations are involved in this endeavour. The role of traffic planners and road safety engineers is to create as safe a road environment as possible, within financial constraints. This is achieved by applying accident remedial measures to unsafe parts of the road network, investigating and improving road standards and ensuring that new routes and other network capacity improvement don't compromise safety standards. In the latter the sometimes competing needs of capacity and safety improvements needs to be balanced. Ideally capacity improvements do not produce a deterioration in road safety, or if they do this is kept within bounds. The effects on safety can be managed only if an estimate of the effects of a network change can be calculated.

Accident prediction models and rates can be used to calculate the expected change in accident occurrence at a intersections, routes (i.e. mid-block), sites (e.g. narrow bridge) or combinations of these, before and after engineering improvements. Accident prediction models and rates have been available in the Transfund (formerly Transit) Project Evaluation Manual (PEM) for intersections, mid-block urban routes and rural routes since the early 1990's. The accident prediction models and rates in the Transfund PEM are becoming more and more important as the need for more refined economic analysis of roading projects increases.

Since the early 1990s there has been a downward trend in accident occurrence, and there have been a number of advancements in the development of accident prediction models. Many of the developments in statistical analysis are detailed in Turner (1995), and subsequently in work by this author (refer to Nicholson & Turner, 1996, Turner & Nicholson 1998a & 1998b). This work also contains accident prediction models developed using these procedures. The main purpose of this study in to develop new accident prediction models for the Transfund PEM for rural and urban intersections and mid-block sections/routes using the model forms in Turner (1995). This study also takes a closer look at the accuracy of the prediction models and includes new developments that have occurred in statistical analysis.

1.2 Objectives

The project was split into two stages. The objectives of stage 1 are:

- To develop goodness-of-fit statistics for the Turner (1995) accident prediction equations for urban intersections;
- To develop confidence intervals for the Turner (1995) accident prediction equations, and to establish the size of the statistical errors;
- To compare the project benefits and benefit-cost ratios derived using both the Turner (1995) and Gabites Porter (current Transfund PEM) prediction models (referred to as a zero-based approach in the project brief).
- To develop accident prediction procedures (that are suitable for insertion into the PEM) for practical application.
- To contact local authorities to determine the amount of data available for developing accident prediction equations in stage 2.

The objectives of stage 2 are:

- To review international literature and current N.Z practice on predicting accidents along routes and at intersections.
- To develop accident prediction models for major accident types for urban and rural links and intersections using the macros developed by Turner (1995).
- To develop goodness-of-fit statistics/graphs and confidence intervals for the accident prediction models.
- To assess the feasibility of using historical accident records and the empirical Bayes method to improve accident predictions at individual intersections and links.

In addition, the report introduces the topic of accident prediction, discusses the current Transfund PEM procedures, and why changes to this procedure are warranted. Such information is considered important to demonstrate why this research has been commissioned.

1.3 Report Layout

The report has been divided into six sections (chapters 2–7), excluding the introduction and references. This report includes some of the contents of the stage 1 report.

Chapter 2 introduces the topic of accident prediction and the research undertaken by Turner (1995) on accident prediction models. It reviews the procedures currently recommended by Transfund New Zealand for calculation of typical crash rates at urban intersections and for mid-block crashes on urban and rural roads. It suggests where there is scope for improvement to these procedures.

Chapters 3 details the review of local and international literature. The review concentrates on the period 1995 to 2000, as Turner (1995) contains a review of literature prior to 1995. This section discusses the merits of new procedures and methods introduced in the papers reviewed and those methods which have been included in this research.

Chapter 4 discusses the data collected from different road controlling authorities (traffic volume and layout data) and the Land Transport Safety Authority (accident data). It contains tables and graphs showing the sample sizes, range in flows and major accident types.

Chapter 5 contains the accident prediction models produced for urban and rural intersections and links. The goodness-of-fit and confidence intervals are also presented for each model.

Chapter 6 describes the changes that would be required to include the new accident prediction models in Appendix A6 of the Transfund PEM. Appendix C shows the new pages and changes to text that would be required to the Appendix A6.

Chapter 7 discusses the empirical Bayes method, that allows an analyst to update the accident prediction from a model (or base model) using site accident history data.

2 Accident Prediction

2.1 Introduction

Traffic accidents, although always too many in human cost, are still relatively rare events for which there is a host of contributory factors related to the road layout, road vehicles, drivers and the environmental conditions. Data quality is also variable, relying on data about officially reported and recorded accidents that are variable in quality and completeness.

As accident severity reduces, the proportion of accidents that are reported also reduces. Reporting is also thought to reduce with remoteness, so that there is a greater likelihood of a minor accident being reported on an urban motorway than on a remote low volume rural road. Reporting rates are also susceptible to the resources available for traffic enforcement, and these vary geographically and over time. There are few data available to independently verify accident numbers, so quantifying the rate of under-reporting is difficult and becomes more uncertain with decreasing injury severity. Accident prediction is therefore an area with much uncertainty, not all of which can be controlled or quantified.

The objective of accident prediction in the context of road investment planning is to forecast the accident performance of a new or substantially altered section of the road network, for which the historic accident record is either not relevant or does not exist.

This new section of network may be an individual intersection, a route realignment, or a new route. Another area of application is in urban traffic modelling, where the effects of network changes and additions are to be evaluated. In this case the effect of the redistribution of traffic through the network on crashes at intersections and routes needs to be predicted.

2.2 Data Availability

A large number of explanatory variables can potentially be introduced to accident prediction models. The greater the number of such variables, the more demanding the resulting models become on data, both in their development and application. Increasing the number of variables and the complexity of the model must be balanced against the improvements in forecasting accuracy that this additional complexity provides.

Traffic volume is a primary explanatory variable in accident prediction. For intersection accident models, this may be disaggregated at various levels for explanatory variables:

- Total inflow to the intersection – referred to here as Type 1 models
- Inflows on the major and minor approaches – Type 2 models
- Conflicting turning flows – Type 3 models

Other variables, such as geometric and land use variables can also be considered in crash prediction models. However, previous research has not clearly indicated which of these non-flow variables are the most important.

The form and complexity of accident prediction models must take account of the limitations in data availability. For strategic network modelling, traffic data will often be limited to approach flows and data on the physical layout of the intersections will be limited. For local area network models turning volume flows and some geometric data are often available. For projects focused on individual road links and intersections, detailed layout and traffic flow data will often be available.

In general, previous research indicates that the more information that is available for an intersection or mid-block route, and given that appropriate crash prediction models are available, the better the crash predictions. When data are limited, as is the case for the strategic network models, the crash predictions will not be accurate at individual intersections, but over a medium sized network (say 15 intersections plus) the overall prediction within the network is usually reasonably good (Turner, 1995). To predict crashes at the intersection or mid-block link level, much more detailed information and models are required to provide accurate estimates.

2.3 Theoretical and Empirical Findings from Earlier Research

Research into road safety and accident occurrence has a long history and a very wide technical literature. A number of general findings have been made over time which are an important background to any new work. This section details some of the important findings in the technical literature up to 1995, which is the period covered by Turner (1995). Developments in the post 1995 period, and any modifications that are proposed to the Turner accident prediction models and procedures, are discussed in the next section; literature review.

2.3.1 Accident relationship to traffic volume

Intuitively, accident numbers would be expected to increase with traffic volume and the empirical evidence generally supports this hypothesis. However total accident numbers do not necessarily increase in linear proportion to traffic volume.

2.3.1.1 *Mid-Block accident rates*

For accidents on routes and mid-block, some findings show the single vehicle accident rate falls with increasing traffic volume, while the multiple vehicle accident rate increases with traffic volume but levels off and possibly declines at high volumes. This levelling off is attributed to the effects of congestion, and reductions in vehicle speeds and speed differences at higher traffic densities.

The levelling off and, in some cases, decline in the crash rate is difficult to measure, as the drop in crash severity at high traffic volumes due to congestion means that the error associated with under-reporting of crashes is likely to be greater than in free flow conditions. It may be that some of this levelling off or decline is due to under-reporting of crashes, rather than an overall reduction in crash occurrence.

The overall graph of accidents versus traffic volume is a combination of these effects. Over a mid-range of volumes, accidents can be expected to increase with traffic volume, in a relationship that may be close to linear. At low volumes the shape of the graph is unclear. At high volumes, accident numbers level off and may even decline at high traffic flows.

2.3.1.2 *Intersection accident rates*

For intersections, combined empirical and theoretical research has shown accident numbers to increase in proportion to the product of the intersecting or conflicting flows raised to a power:

$$A = b_0 \,.\, Q_1^{b1}.Q_2^{b2}$$

Researchers have made various findings on b_1 and b_2, including $b_1 = b_2 = 1$ (accidents proportional to the product of flows), and $b_1 = b_2 = 0.5$ (accidents proportional to the square root of the product of flows). However, in some models the b_1 and b_2 factors are not fixed to 1 or 0.5, but are instead allowed to vary. Typically, parameters fall within the range of 0.2 and 1.5. All parameters below 1 indicate the classical trend of a levelling off in the crash rate at high traffic volumes. The higher parameters values (greater than 1) are typically associated with crashes that occur more frequently in congested conditions, such as rear-end crashes, due to the increase in vehicle queuing and drivers' limited ability to cope in such conditions. In addition the impacts in terms of non-recurring congestion, would mean that such crashes would warrant a Police response, even if the consequences (injuries) are relatively minor.

2.3.2 Disaggregation of accident types

In 2.3.1, possible reasons for separating single vehicle and multiple vehicle accidents on road links were noted. Other disaggregation may also be useful. At intersections, as well as single versus multi-vehicle accidents, the majority of accidents can be grouped by type of movement, in order of frequency, as:

- Crossing, no turning (Type H)
- Crossing, right turn (Type L)
- Rear End (Type F)
- Lost Control (Types C,D)
- Crossing/Turning (JA)
- Other (pedestrian, cyclist)

The conflicting volumes that correspond to these types of collision will vary with the type of intersection and the relative proportions of crossing and turning traffic.

2.3.3 Disaggregation by time period

For intersection accidents, the accident rate does not demonstrate any marked variation between day and night, or by time of day, with high crash periods generally being associated with high traffic volume periods. Exceptions are loss-of-control (C & D) accidents, which have a higher night-time frequency when traffic volumes are lowest. During the day the occurrence of such crashes is lower, which is probably due to the fact that errant vehicles are more likely to hit another vehicle during the day, when there is more traffic on the road. Higher vehicle speeds at night, due to less congested flow conditions, also contribute to the higher crash rates, as do other factors that vary between daytime and night-time, such as alcohol consumption.

Turner (1995) attempted to develop accident prediction models by time period and found that the model parameters were not always converging to within the typical range of 0.2 to 1.5. Some parameters were negative and other were 2 and above. This is thought to be a result of the low number of accident observations, particularly in the 2-hour morning and afternoon peak periods. Much larger sample sizes, probably in excess of 300 intersections, would be required to develop realistic models.

2.4 The Statistical Distribution of Accident Occurrence

The development of accident prediction models involves multivariate statistical analysis between the dependent variable of accidents and independent variables such as traffic flows. For metric variables (as opposed to ordinal or nominal variables), simple (one independent variable) or multiple (more than one independent variable) linear regression (sometimes referred to as normal linear regression) is frequently used to determine a best fit relationship between the dependent and independent variables.

In normal linear regression it must be assumed that the probability distribution of the dependent variable is either normally distributed or sufficiently close to being normally distributed, that the regression models and other statistics (such as goodness-of-fit) have a low error. In normal regression the independent and dependent variables should not only be metric, but also continuous. In addition, the variability in the dependent variable should be constant for all values of the independent variable.

Even when some of these assumptions are not strictly met, analysts often choose normal linear regression over other available techniques because it has considerable advantages in ease of computation, due to the mathematical properties of the normal distribution. There has also been a lot more development of test statistics, such as the goodness-of-fit statistic (Pearson's R^2), the standard deviation and confidence intervals. These statistics are easily understood as they are based on the theory of least squares, which can be explained within a graphical framework. The more complicated alternative of the theory of maximum likelihood, which is the basis for generalised linear regression models, is more difficult to explain and much more difficult to calculate.

However, in the case of accident prediction, there is no reason to expect that the dependent variable will be normally distributed. In addition, the accident frequency variable is discrete, rather than continuous, and the variability in the dependent variable generally increases as accident frequency increases. In each of these cases the Poisson or Negative Binomial distribution are more appropriate. The difference between these two distributions is that in the case of the Poisson distribution the variance is equal to the mean, which takes into account the site variability (that is the variability one might expect from year to year at a particular site), be it an intersection or route. The Negative Binomial distribution has a variance greater than the mean, which allows for the between-site variability due to features at each site that are not considered in the predictions, such as a dangerous geometric layout.

To undertake regression with a Poisson or Negative Binomial dependent variable, one needs to use the theory of maximum likelihood and to develop generalised linear models. These models require much more computation power than the normal

regression models. Additionally, the test statistics for these models are not as well developed, because of the complex mathematics involved.

2.5 The Existing PEM Procedures

2.5.1 Introduction

The accident prediction relationships for urban intersections currently set out in the Transfund Project Evaluation Manual (PEM), derive from work carried out by Gabites Porter Partners (1991). For mid-block accidents, the analysis is based on work by the Land Transport Safety Authority (Jackett, 1992 & 1993), which is regularly updated, Beca Carter Hollings & Ferner Ltd (1994) and Opus International Consultants (1997) on multilane median-divided highways and motorways.

2.6 Urban Intersection Accident Prediction

The accident prediction equations for urban intersections currently used in the PEM are based on work by Gabites Porter (1991) and are of the linear form:

$$A_T = b_0.Q_T$$

where:

A_T = total number of reported injury accidents in a five year period
Q_T = total flow entering the intersection

These superseded a quadratic relationship of the form:

$$A_T = b_0.Q_T^2$$

which was used from September 1993 to March 1995, based on the same research but with modified coefficients b_0. Prior to September 1992, the linear form was used (note that from September 1992 to August 1993 the $A_T = b_0.Q_T$ model form was used but with incorrect coefficients).

Prior to 1991, the National Roads Board Economic Appraisal Manual gave limited advice on typical accident frequencies and accident prediction methods, relying on work by Kitto (1980) on accident rates for urban right angled X and T-intersections under various forms of control.

The current equation assumes a linear relationship between the dependent variable (accidents) and the independent traffic volume (the sum of the approach flows). The foregoing discussion, based on the bulk of research in this area, concluded that a linear relationship is not appropriate, and therefore the model exponent should generally not be one, and particularly for total crashes should be less than one. There are a number of other problems with the existing models, one being that the models were developed using normal linear regression.

These models also use the sum of the approach flows as the independent variable. This is not considered a good predictor of crashes as it does not take into account the fact that most multiple vehicle collisions occur due to the interaction of vehicle streams, a factor not taken into account by a combined traffic volume variable.

For example, crash prediction models based on the sum of the approach flows will return the same crash rate at a cross-roads if 90% of the traffic is on one road, or 50% is on each road. In the second case there is more chance that a vehicle will conflict with another vehicle and one would therefore expect a higher crash rate.

Another problem with the models is that the traffic flows were derived from transportation network models, rather than traffic count data. Even with traffic count data taken only in three one-hour count periods, there is likely to be considerable error. The use of modelled rather than actual traffic counts will add modelling error on top of this. Given that both normal linear regression and generalised linear regression assume that the independent variables contain no error, to minimise the error that this assumption causes in the models, raw counts rather than model counts should be used for developing the accident prediction models.

These models do not predict individual crash types, which can be important when a network change alters the traffic flow pattern so that fewer vehicles make higher crash rate manoeuvres, such as right turns.

2.7 Mid-Block Accident Prediction

The accident prediction model currently used in the PEM for urban and rural mid-block crashes is:

$$A_T = b_0.X_T$$

where X_T is the exposure in 100 million vehicle kilometres over the mid-block section length. Values of b_0 have been fitted to this model from analysis of the crash accident database, distinguishing between speed limit zones, presence or absence of a solid median, and by traffic volume band for rural roads. A separate value of b_0 has been determined for motorways and other multilane median divided roads.

The models presented in the PEM are based on models developed using Poisson regression by the Land Transport Safety Authority (Jackett, 1992). The initial models considered a number of variables, with the more significant variable classifications included in the PEM tables. Again the models assume a linear relationship between the traffic flow variable and the accident frequency, and that the accident frequency can be related to the sum of the traffic volumes travelling in each direction. While this may be the case for some mid-block crash types, such as loss-of-control, the product of the flows in each direction is more likely to be correct for head-on and driveway/minor intersection type crashes.

Again, these models do not predict individual crash types, which can be important when calculating the crash benefits of roading improvements because some types of accidents are "more expensive" than others.

2.8 Crash Procedures

The PEM provides for two types of crash procedures, accident-by-accident analysis and accident rate analysis.

2.8.1 Accident-by-accident analysis

An accident-by-accident analysis is used when the intersection or route modifications do not significantly alter the layout or traffic flow at an intersection or along a route, so that one can either apply a global crash reduction/increase percentage to either all crashes or each crash type. Obviously to do this type of analysis there must also be an existing record of crashes for a route or intersection. This crash record generally needs to be of five years duration and there should not have been any major alterations or traffic growth within the five year period that may have significantly changed the crash rate.

It is not intended that the crash prediction models being developed in this research project would result in any modification to this type of analysis. However, there needs to be more guidance within the PEM on when an analyst should use accident-by-accident analysis and when he/she should use accident rate analysis. This guidance could be in the form of examples or specific criteria. This has to some extent been addressed by the new Addendum (B6) in the Transfund project evaluation manual, but requires further attention.

2.8.2 Accident rate analysis

Accident rate analysis is used when the intersection, route or network modification is expected to significantly effect the layout or traffic flow pattern, and for new routes and intersections. This type of analysis makes use of the crash prediction models and crash rates to establish the expected crash rate for an intersection or route. For an existing intersection or route the existing crash rate of a site is compared with this typical crash rate, and it is determined whether the proposed modifications will reduce the crash rate to this typical rate. If so, the crash benefit is derived from this reduction in the crash rate, otherwise some fraction of this difference is used depending on the expected effectiveness of the change.

For new intersections or routes it is expected that the layout will be to the appropriate Transit or Austroads layout standards, and therefore the crash rate will be equal to or less than the typical crash rate of existing intersections or routes of that type.

When considering network changes, as might occur when a new link is added, it is often more realistic to calculate future crash frequencies for routes and intersections from existing crash rates on the network. For example, a new link may reduce the traffic volumes on a parallel route by 50%. If it is assumed that the crash rate on this existing route remains the same, then the crash frequency on this existing route will halve. Other routes such as those feeding into the new route will carry more traffic than previously and the crash rates on these routes will therefore increase. For the new routes there are no existing crash records and crash frequencies can only be calculated from typical crash rates.

This method is probably accurate enough for small changes in traffic volumes, as it can be assumed for small changes in volume, particularly for links and intersections with mid-range traffic volumes, that the relationship between traffic volumes and accidents is linear. However, for larger traffic flow changes, such as the 50% considered in the example above, the non-linear relationship between flow and crashes may result in large errors in the crash predictions, and therefore crash costs. In such cases it is necessary to make use of both existing crash rates, to get the right level of magnitude, and the crash prediction models, to work out the relative percentage reduction in crashes for a given reduction in traffic volumes. This issue will be examined in more detail in stage 2 of this research, where the Empirical Bayes method will be examined. This method makes use of both accident rates and the existing accident data.

2.9 Improvements to the Existing Accident Prediction Equations

2.9.1 Introduction

The previous section has identified a number of theoretical problems with the existing crash prediction models and crash rates in the PEM. New crash prediction models are required, preferably those which can predict individual crash types. These new models should be based on state-of-art statistical techniques; in this case generalised linear models with a Poisson or Negative Binomial distribution. A major advantage of these new models will be that analysts, Transit NZ and Transfund NZ will be able to have more confidence in the estimates these models produce, due to the better theoretical basis for the models and the support of international research findings.

This confidence will also flow through to the crash cost estimates, produced by multiplying the crash rates by the cost per crash. Not only will there be more confidence in the accuracy of the crash predictions, but it will be possible to use crash costs for different crash types, rather than one overall average crash cost. Given that this area of road user benefits has been identified as one that has high error, more accurate crash predictions is considered well overdue.

A number of crash prediction models of the type recommended have already been developed by Turner (1995) for major crash types and most intersection and control types. These equations are discussed in the next section. However, a number of new models will need to be produced, particularly for routes (or mid-block locations), to provide a comprehensive set of models for the PEM.

2.9.2 The Turner accident prediction equations

Turner (1995) developed accident prediction models for urban intersections based on conflicting flow volumes. The models were developed using generalised regression techniques for over 360 T- and X-intersections including traffic signals, roundabouts, priority control and uncontrolled. The models were then tested in predictive models for three road networks in Christchurch and Lower Hutt and yielded promising results.

The Turner models are either of a Negative Binomial or Poisson error structure, the latter being adopted for situations in which the variance is equal to or less than the mean:

$$A_k = b_0 . q_{j1}^{b1} q_{j2}^{b2}$$

where:

A_k= conflicting flow crash type between flows q_{j1} and q_{j2}
q_{j1}, q_{j2} = conflicting turning volume counts.
and b_0, b_1 and b_2 are regression coefficients

Models were developed for the major crash type for each form of control. Models were not developed for the more minor crash types as there were either an insufficient number of crashes observed at the sample intersections or the crashes involved travel modes for which counts were not available (cyclists and pedestrians). While equations are not available for these minor crash types, the average proportion of these minor crash types at a typical intersection is known. Having calculated the total number of major crashes, the number of minor crashes can be calculated using this proportion. The only missing information at this stage is a crash cost for these minor crashes, which will be different from the overall crash costs given in the PEM.

While this method will satisfy in the interim, additional models should be developed, particularly for cyclists and pedestrians and for centres, such as Christchurch, where there is a high proportion of cyclists. As for the route equations, the model structures and analysis methods are already available, and therefore it is just a matter of collecting new data and processing them, to generate the models.

Ideally, separate intersection models should also be produced for rural intersections, to take into account the different speed limits in the rural road environment. As a first step it would be useful to examine the crash patterns occurring at rural intersections, so that the differences between the urban environment (as explored in Turner's thesis) and the rural environment can be compared. It may be possible to use a factor to take into account the difference in speed environment, although it is expected that separate models will need to be developed.

2.9.3 Models for routes

It is proposed that the route accident prediction models should also be updated using the model structure and techniques developed by Turner, as the LTSA (Jackett, 1992) models/rates do not take into consideration the non-linear relationship between traffic flows and accidents. Models will need to be developed for both rural and urban routes.

This will require a data collection exercise, although it is expected that much of the information used by Jackett, which is updated annually, will be available, although more information will probably be required for urban routes. It is proposed that a similar model structure to that for intersections, with either one flow variable (total two-way flow) or two flow variables (flow in each direction), would be used.

2.10 Goodness-of-fit Statistics

Goodness-of-fit statistics are used in linear (and non-linear) regression to test how much of the variability in the dependent variable, in this case the accident frequency, can be explained by the independent variables, the traffic flows (or other non-flow variables). It can be used to assess whether a new independent variable should be added to the prediction equation, by comparing the value of the test statistic with and without the variable in the regression equation. If the change in the test statistic is large, this usually indicates that the new variable has explained a considerable amount of variability in the observed values of the dependent variable, and should be added to the list of important predictor variables and the regression equation. If the change is low then the independent variable is usually unimportant and is discarded.

The goodness-of-fit statistic is also used to assess the accuracy of prediction models, by quantifying the proportion of variability in the dependent variable the regression equation predicts. When this proportion is low then there is likely to be considerable error in the crash predictions as the model being used does not have sufficient variables to explain enough of the dependent variables variability. When the proportion is high an analyst can have some confidence that the predictions will be reasonably accurate.

Unfortunately, most accident prediction models do not explain a high proportion of the variability in the accident frequency, as deriving a set of adequate variables is not easy for the road environment, due to the number of factors which impact on this environment.

In normal linear regression the Pearson's r^2 statistic is usually used to express the goodness-of-fit, as it has mathematical properties that make it easy to explain and calculate. However, in generalised linear models, the Pearson's r^2 statistic can no longer be used. While there are a number of statistics that have been proposed to test the goodness-of-fit of generalised linear models, the most commonly used statistic being the scaled deviance (SD). See Wood (2000b) as to how the scaled deviance can be calculated.

2.11 Confidence Intervals

The method used to develop confidence interval estimates in multiple or normal linear regression, can not be directly applied to prediction models developed using generalised linear models theory. To develop confidence intervals for the Turner prediction models it was necessary to go back to first principles and to develop, using the statistical theory of maximum likelihood, new confidence interval equations. Details of the development of the confidence interval equations and the equations themselves are given in Appendix E.

3 Literature Review

Following Turner (1995), there have been a number of advances in the field of accident prediction modelling. This section outlines some of the most relevant papers produced during the last five years.

The first paper, by Maher & Summersgill (1996), provides a summary of a number of technical (statistical) problems that have been identified with the development of generalised linear models (GLMs) for accident prediction. This paper provides some solutions to these problems. The remaining papers in this section, except that by Trinadha & Rengaraju (1998) on modelling conflicts, expand on the problems and solutions raised by Maher & Summersgill.

3.1 GLM Technical Problems

The technical problems identified by Maher & Summersgill are applicable to the development of GLMs for accident prediction. The technical problems may or may not be a problem in other applications of GLMs.

The 'low-mean-value' problem, which as the name applies occurs when the mean of the dependent variable (accidents) is low (generally <0.5), as commonly occurs when accidents are broken down by accident type, results in poor accuracy of the scaled deviance goodness-of-fit statistic. At low mean values the scaled deviance statistic can indicate that the model has a good fit to the data, when compared with the x^2 distribution, when the fit is in fact not particularly good. Wood (2000a and/or 2000b) expands on this discussion, and shows that the accuracy of the scaled deviance test deteriorates as the mean value reduces. Wood details a method for addressing problems with the goodness-of fit testing (see later section on low-mean –value problem).

Overdispersion is another problem associated with accident data. Accident data sets often have a variance significantly greater than the mean (overdispersion), so that it is not satisfactory to assume that the data is Poisson distributed (variance equal to or close to the mean). The overdispersion is thought to occur because of the between-intersection variability and because of factors, such as geometric variables, that are not represented in the prediction models.

Overdispersion can be addressed in a number of ways. The most common approach in accident prediction is to assume a negative binomial distribution for the accident data set. The negative binomial distribution consisting of the within-intersection variability, represented by the Poisson distribution, and the inter-intersection variability, being represented by a Gamma distribution. This is the approach favoured by most overseas researches and this research team, because of the convenient statistical properties associated with using the Gamma distribution.

Other distributions can be used to represent the overdisperion, including the quasi-Poisson distribution and other members of the negative binomial family of distributions, which involve different inter-intersection distributions to the Gamma. Discussion on these other options can be found in Maher & Summersgill (1996).

Dissaggregation of data can also be a problem, because there can be a scarcity of accident data, and because of correlation between data values. Accident data and sample sets can be dissaggregated by accident type, time-period, year and intersection approach. Further discussion on these matters is contained in Poch & Mannering (1996) and Lord & Persaud (1999), whose finding are discussed in more detail latter.

Leading on from the issue of whether or not to dissaggregate data, or to continue with aggregated data is that of identifying trends in accident occurrence over a number of years. In most studies the accident data over several years, often five years, is used to produce models without factoring in whether there is a trend in the accidents over the time period. In New Zealand there is evidence that there has been a major downward trend in accident occurrence, particularly during the 1990s. The accident rates at sites have been dropping, even though many sites have not received engineering remedial treatments, because of increased levels of enforcement, changes in driver behaviour and 'safer' vehicles. Lord & Persaud (1999) discusses how the trend in accident occurrence can be factored into the development of GLMs (see latter).

The issue of variability in the traffic flow estimates, when an assumption of the standard GLM procedures is that independent variables are deterministic, has been discussed in Maher (1989). A method for dealing with this error is presented in Maher, and summarised in Turner (1995). No attempt has been made to incorporate the Maher procedure in this work, as there is insufficient traffic data, for intersections at least, to improve the GLMs. Lord (1999) also discusses method that can be used to estimate missing traffic counts and improve the reliability of the traffic flow variables.

Maher and Summersgill briefly discuss the combination of model predictions with site observations, or accident history data. They discuss how much weight should be placed on site observations and the models. This topic is explored in more detail in Chapter 7, on the Empirical Bayes method.

3.2 Low-Mean-Value Solution

Wood (2000b) has developed a procedure that deals with the error in the goodness-of-fit statistic, scaled deviance, when the mean of the accident data is low. The method involves grouping the data into traffic volume bands, and using the average flow and accident frequency for each band in the goodness-of-fit test. The lower the mean of the raw accident data the more sites/approaches that have to be grouped before testing.

While grouping does reduce the degree-of-freedom of the test, which is not desirable, it does bring added confidence to the value of the scaled deviance. Wood (2000a) details the procedure used, including all the statistical theory behind his technique, while Wood (2000b) provides a guide on how to use the technique, for transport modellers.

The method developed by Wood has been used in the goodness-of-fit testing. The results of the goodness-of-fit testing (and the groupings used) are presented in Section 5.5 of chapter 5.

3.3 Dissagreggation of Accident Data by Time and/or Approach

A number of researchers, including Maher & Summersgill (1996) have raised concerns with the dissagreggation of accident data by time, accident type and approach. In terms of time, accident data can be broken down by year and by weekday and weekend time periods (e.g, weekday morning peak). Two problems surface when this is done; (1) there can be a scarcity of data if the sample size is not sufficient and (2) there is likely to be correlation between accidents occurring at the same site.

The first issue can be countered by increasing the sample size. Dissagreggation should only occur to a level where there is still sufficient accident data to produce 'reasonable' models. A check of the model parameters can usually determine whether the models are 'reasonable', as such parameters should be in the range of 0.3 to about 1.5. In Turner (1995) the accident prediction models by weekday and weekend time period involved a number of unreasonable parameters, that were either negative, or above 2. It was concluded that scarcity of data was a problem and that the sample sizes were insufficient to dissaggregate the accident data to this level of detail. To the author's knowledge a statistical method that indicates how much dissaggregation can occur for a particular sample size and mean accident rate is not available.

The second consequence of dissaggregation, that of correlation between data from the same intersection, has been examined by Poch & Mannering (1996). It was thought that such correlation would result in a loss of estimation efficiency, which would effect the reliability of the prediction models. In their paper they examined, using likelihood ratio tests, whether models results are significantly effected by such violations of independence between data values. They concluded that "independence violations are not significantly affecting model results".

3.4 Annual Trends in Accident Data

There is evidence in New Zealand, as overseas, that there is a downward trend in accident occurrence. While car ownership levels and vehicle-km of travel continues to rise, accident numbers in general, and fatality rates, continue to drop. This downward trend has occurred because of the efforts of numerous government organisations, including road controlling agencies who undertake physical engineering safety works.

This downward trend in accident occurrence is generally not factored into accident prediction models. Lord and Persaud (1999) discuss a number of models that take into account trend in accident data by year. This includes marginal models (MM), transition models (TM) and random effect models (REM). The paper details one of the MM methods and produces models using the method for 868 four-legged signalised intersections in Toronto, Canada. The MM, which included the trend in accident occurrence, were compared with standard Negative Binomial models, without trend. They conclude that the models incorporating trend are superior to the models that do not.

3.5 Conflict Modelling

Trinada & Rengaraju (1998) detail a method developed for modelling traffic conflicts using simulation. The paper details a model that has been developed for a X-road and can simulate the effects of vehicle composition (2, 3 & 4 wheeled motorised vehicles, pedal cycles and heavy vehicles), gap acceptance (critical gaps and move-up times) and degrees of priority (give-way, stop and signals) on conflict rates, particularly in congested conditions. Field data (conflict surveys) were collected to validate the conflict rates predicted by the model. The authors did not relate the conflict rates (which measure the close misses at an intersection) to the actual accident occurrence, but this would be the next step in this research.

This paper has been included in the literature review as it shows a promising avenue for the development of accident prediction modelling, particular in regard to developing a better understanding of the mechanisms and causal factors that influence accident occurrence. Such understanding is required to develop future models which capture more of the variation in accident occurrence by selecting a set of appropriate predictor variables. The area of relating conflict data to accident occurrence is also an area of much promise, as conflict data at a site can be collected over short time periods, whereas a researcher must wait a longer time for accident trends to become evident. By finding such a relationship we might better be able to predict crash occurrence.

3 Data Collection

3.1 Introduction

This section details the traffic, site layout and accident data collected from road controlling authorities and the Land Transport Safety Authority (LTSA). Tables have been produced showing the number of intersections and links that have been selected from different parts of the country (Appendix A) shows the actual road controlling authorities who provided data, by intersection and link type). The accident data for each intersection and link type has been broken down by accident type and presented in pie graphs. However, before we move onto the data collected we will first discuss the results of the survey undertaken at the end of stage 1 to assess the availability of data for the prediction models.

3.2 Survey of Local Authorities)

This section details the results of a survey questionnaire sent to local authorities, Transit NZ and consultants. The main purpose of the survey was to assess the availability of data, particularly turning volume counts for intersections. Such data can be difficult to come by, particularly for low-volume intersections. With the large sample size required (50 to 100 intersections) to produce good quality prediction models, it was important to assess whether sufficient data could be sourced. This survey was undertaken prior to stage 2, so that the scope for stage 2 could be refined, if necessary, to allow for the collection of additional traffic counts.

Survey Questionnaires were sent out to 58 organisations, including Councils, Transit offices and consultants. Appendix B lists the organisations approached and a copy of the questionnaire. As expected the initial response rate was low. However, with follow-up phone calls the final response rate was 66%. Tables 7.1, 7.2, 7.3 and 7.4 summarise the data that are available. Further details of the questionnaire responses can be found in the table in Appendix A

Table 4.1 City Council Traffic Count Availability

City Council	Number of Links	Number of Intersections	Others (eg. railway crossings)
Manukau	546	0	0
Invercargill	10	18	0
Christchurch	200+	200+	0
Dunedin	77	42	0
Napier	173	12	0
Wellington	31	15	0
Palmerston North	205	201	24
Porirua	0	2	0
Waitakere	50	18	1
TOTAL	**1292**	**508**	**25**

Table 4.2 District Council Traffic Count Availability

District Council	Number of Links	Number of Intersections	Others (eg. railway crossings)
Queenstown-Lakes	360	0	0
Tasman	15	10	0
Westland	0	0	0
Whangarei	60	34	0
Hastings	66	22	0
Banks Peninsula	0	0	0
Waimakariri	800	0	0
Timaru	330	22	0
Ashburton	1	3	0
Buller	0	0	0
Waimate	100	0	0
Rodney	500	0	0
Rotorua	0	3	0
Hurunui	100	0	0
McKenzie	68	0	0
TOTAL	**2400**	**94**	**0**

Table 4.3 Transit NZ Regional Offices Traffic Count Availability

Transit NZ Regional Office	Number of Links	Number of Intersections	Others (eg. railway crossings)
Hamilton	124	0	0
Dunedin	77	0	0
Auckland	95	30	0
Wellington	75	20	0
Christchurch	200+	2	0
Wanganui	102	0	0
Napier	27	0	0
TOTAL	**700+**	**52**	**0**

Table 4.4 Consultants and Others Traffic Count Availability

Organisation	Number of Links	Number of Intersections	Others (eg. railway crossings)
Opus (Hamilton)	0	0	0
Opus (Lower Hutt)	2	0	0
Montgomery Watson	0	23	0
Montgomery Watson (Chch)	0	6	0
Traffic Design Group	0	49	0
Beca Carter	10	30	0
Canterbury Regional Council	0	0	0
TOTAL	**12**	**108**	**0**

As expected the district and city councils have both link and intersection count data, with the latter having the higher number of intersection counts.

The Transit NZ regional offices mainly have route counts, and most of these are on state highways. Consultants tend to have few traffic counts, with those that are available being intersection counts.

The Land Transport Safety Authority provided their database on link data, including traffic counts and layout variables. This information was used to develop the urban mid-block prediction models in the Transfund PEM. This data was used, along with information collected from the questionnaire participants, to calculate accident prediction models for the urban mid-blocks and rural routes.

A more detailed analysis of the questionnaire results shows that there are few turning volume counts available for rural intersections. Subsequent to this finding a series of traffic counts (at 50 sites) were collected at rural T-junctions in Canterbury, Wellington, Wanganui and Tauranga. There are also plans to collect counts at rural cross-roads (X-roads).

Small sample sizes were available for signalised T-junctions and 3-arm roundabouts. During data collection (see next section) a number of additional signalised T-junctions, with traffic count information, were identified, and the sample size was increased. We were unable to increase the sample size of 3-arm roundabouts to a desirable level. It was decided that given the layout of roundabouts, being equivalent to a series of give-way controlled T-junctions on a one-way main road, that we would compare the small sample of 3-arm roundabouts with their 4-arm counterparts. It was proposed that a factor be developed to take into account the greater distance between intersection arms on 3 –arm roundabouts (on at least one approach).

4.3 Data Collected

During stage 2 most of the local authorities, Transit offices and consultants that responded to the stage 1 questionnaire were approached for traffic count and other data. The exceptions were those who indicated on their questionnaires that were

unable to provide data, usually because they did not have adequate staff resource to extract the information from their records.

For intersections we requested 1-hour turning volume counts for the morning and evening peaks and a non-peak period. We also asked for information on intersection layout and to be notified whether there had been a change in control, or some other major change that would have effected the underlying true accident rate. For links traffic volume information was collected along with information on land-use.

Table 4.2 summarises the number of intersection and links for which data was collected from survey participants, from different regions in the country. A full list of sites by road controlling authority is contained in Appendix A

Our target sample size for each intersection and link type was 100 sites. Most of the intersection types have sample sizes less than 100, but greater than 50, which while not ideal is considered sufficient to develop prediction models. Again the problem seems to be in the availability of turning volume counts. We would suggest that when the models are next updated that additional intersections be added to the samples.

The sample size for 3-arm roundabouts is too small, and hence we have not produced models for this intersection type. Instead we have calculated an adjustment factor, using this small sample, for the 4-arm roundabout prediction models to allow for the missing arm.

3.4 Major Accident Types

The following pie graphs show the major accident types at each intersection and link type. Prediction models have been developed for each of the major accident types, which ranges from 3 to 5 types for each sample (refer to Section 5 for prediction models). The remaining accident types have been group under 'other'. Models for the other category have been developed for completeness, even though such a model does not account for the causal factors that lead to each particular accident type.

All intersection accidents, except rear-end accidents, have been removed from the link sample sets. There are a number of factors that can cause rear-end accidents, and often it is a combination of factors. The presence of intersections along a route is a major causal factor, but other factors such as presence of kerbside parking and high driveway volumes, are also factors. While the LTSA coding system does attempt to categorise rear-end factors by type, the categories are not sufficiently fine, nor is there supplementary data that allows a researcher to confidently determine what rear-end accidents are associated with intersections and which are associated with mid-block sections. Jackett (1993) dealt with this problem by using a variable for the number of intersections per km in his models, and this should be considered in subsequent research.

Table 4.5 Intersection/Links by Region

Intersection & Link Types	Auckl and Regio	Waik ato/B ay of Plent	Welli ngton & Wang	Chris tchur ch & Dune	Trans it NZ	Total
Urban Intersections						
Signalised X-	20	11	17	52	9	109
Roundabouts (4-	7	-	15	32	1	55
Priority X-Roads	11	2	15	38	10	76
Signalised T-	1	-	3	26	-	30
Roundabouts (3-						
Priority T-	17	2	18	44	8	89
Uncontrolled T-	11	-	18	22	-	51
Urban Mid-Blocks						
Arterials	67	49	12	37	-	165
Collectors	6	44	74	25	-	149
Local Roads	52	67	-	-	-	119
Rural Intersections *						
Priority X-Roads						
Priority T-	-	20	10	20	-	50
100kph Routes						
Rural Highways	5	25	29	41	-	100
Rural Local Roads						
Motorways	65	3	8	-	-	76

* Additional traffic count data was collected because data was not readily available for rural intersections.

With intersection accidents (except rear-end accidents) being removed from the analysis, separate predictions should be made for intersection using the intersection models. However, rear-end accidents should not be predicted, as this would result in double-counting.

4.4.1 Urban Intersections

Figure 4.1 to 4.6 show the major accident types for the urban intersection samples:

1. Signalised X-Roads
2. Roundabouts (4-arm)
3. Priority X-Roads
4. Signalised T-Junctions
5. Priority T-junctions
6. Uncontrolled T-junctions

Figure 4.1: Signalised X-Roads

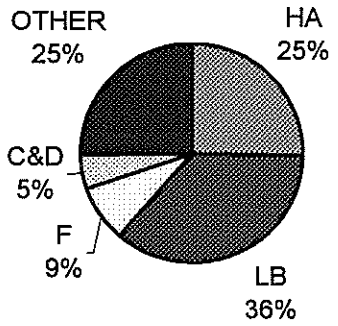


Figure 4.2: Roundabouts (4-arm)

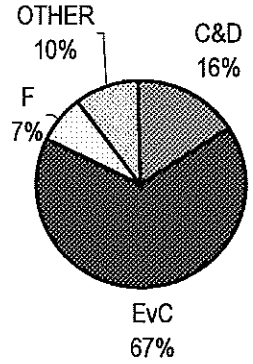


Figure 4.3: Priority X-Roads

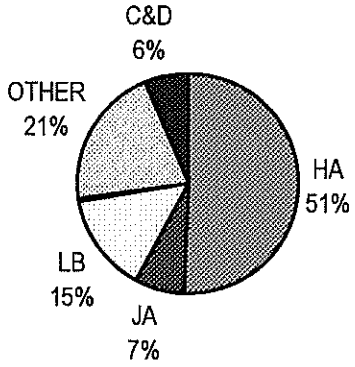


Figure 4.4: Signalised T-Junctions

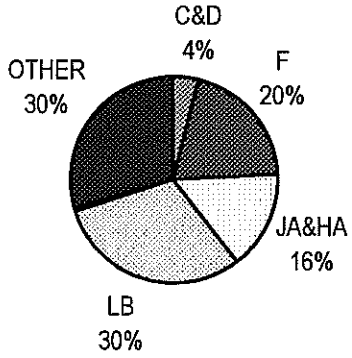


Figure 4.5: Priority T-Junctions

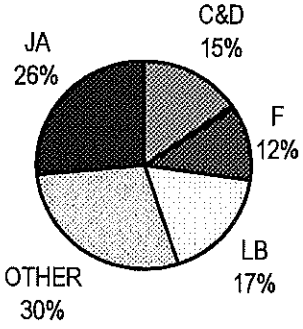
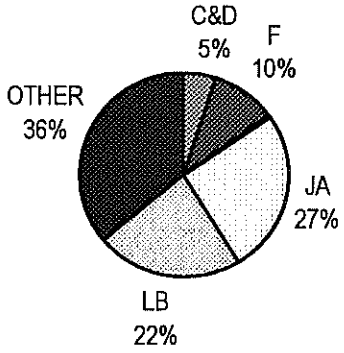


Figure 4.6: Uncontrolled T-Junctions



* refer to Appendix D for the crash coding definitions

At signalised cross-road the major accident types are right-turn against (LB), red-light running (HA, crossing, both straight), rear-end (F) and loss-of-control (C & D). All other accident types, including those involving pedestrians are grouped under 'other'. The four major accident types make up 75% of all the reported injury accidents at the study sites.

At 4-arm roundabout a variety of accident types are grouped under the heading of entering versus circulating accidents. While individual accident types, such as right-turn against, are coded at roundabouts, there is evidence of mis-coding of many of the

accidents because Police officers do not always ascertain the direction each vehicle was intending to travel. Hence a crossing-no-turning accidents might actually be a right-turn-against accident. Given that all such accidents share common conflicting flow pairs (the entering and circulating flow), it was decided to group them together. 67% of all reported injury crashes are of this type. A further 23% are loss-of-control or rearend accidents, making a total 90%.

The major accident type at priority cross-roads is associated with vehicles not giving way; the crossing-no-turning accident type (HA). 51% of accidents are of this type. A further 28% are loss-of-control, crossing-right-turn (JA) and right turn against. This gives a total of 79% of reported injury accidents.

Like signalised cross-roads, a high proportion of T-junction traffic signal accidents, in this case 30%, are in the 'other' category. This is because of the generally higher occurrence of pedestrian related accidents at traffic signals. The remaining 70% of accidents, at the sample sites, were divided over four accident types; right-turn-against (highest), rearend, loss-of-control and right-turn-crossing.

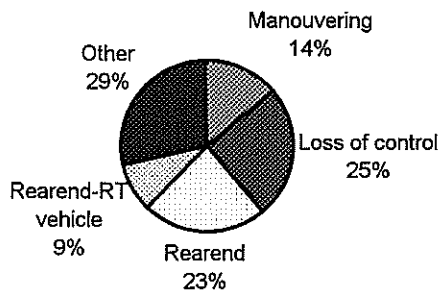
The same major accident types occur at priority and uncontrolled T-junctions, with the proportion of 'other' accidents increasing as the level-of-control and traffic volumes reduce. At uncontrolled T-junctions many of the accidents are links accidents that just happened to occur within the analysis section, and don't necessarily related to the intersections themselves.

4.4.2 Urban Mid-block Sections

Figures 4.7 to 4.9 show the major accident types for the urban mid-block samples:

1. Urban Arterials
2. Urban Collectors
3. Urban Local Roads

Figure 4.7: Urban Arterials



All intersection accidents have been removed from the data-sets. The major accident types are manoeuvring (including hit-object), loss-of-control and rearend, which is split into those involving straight through vehicles and those involving one vehicle turning right. 29% of accident were of 'other' accident types. The other links types show a similar pattern of accident occurrence.

Figure 4.8: Urban Collectors

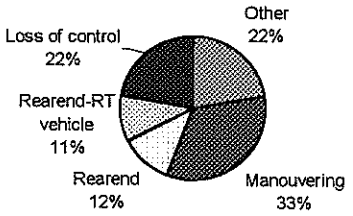
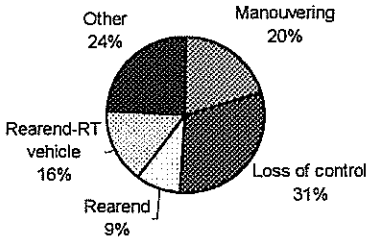


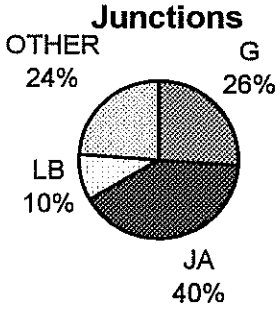
Figure 4.9: Urban Local Roads



4.4.3 Rural Intersections

Figure 4.10 shows the major accident types for the rural T-junction sample set:

Figure 4.10: Rural T-Junctions



* refer to Appendix D for the crash coding definitions.

There were few accidents (less than 50) at the 50 rural T-junctions used in this study. Ideally a larger sample set should be used in update studies. Given the small sample size we could only separate out three of the accident types; right-turn-against, right-turning-same-direction (G) and right-turn-crossing. These types made up 76% of all accidents observed at the 50 sites. Interestingly each of these accident types involve either the right turn movement from the side-road (40% of accidents) or the main road (36% of accidents).

4.4.4 100 km/h Routes

Figures 4.11 and 4.12 show the major accident types for the rural routes and motorway samples:

1. Rural Highways (excluding motorways)
2. Urban & Rural Motorways

Figure 4.11: Rural Highways

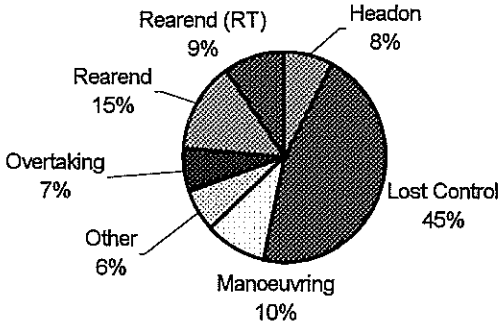
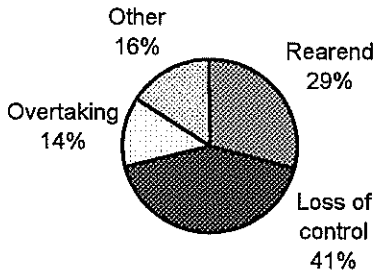


Figure 4.12: Urban & Rural Motorways



Again intersection accidents have been removed from the data-set. The predominate accident type on rural highways is loss-of-control accidents, with 45% of all reported injury accidents. The other major accident types are rear-end (again split into two categories), head-on, overtaking and manoeuvring. These six major accident types include 94% of 'mid-block' accidents occurring on rural highways.

The type of accidents that can occur on motorways is limited by the 'limited access' nature of motorways. The three major accident types are loss-of-control, rearend and overtaking (or weaving).

5 Accident Prediction Models

5.1 Prediction Models

The following subsections present the accident prediction models produced for urban and rural intersections and links. These generalised linear models (GLMs) were produced using the method used by Turner (1995), which is in turn based on work by Hauer (1989) and other researchers. Goodness-of-fit statistics and confidence intervals for these models are presented in subsequent sections.

5.1.1 Urban Intersections, 50 km/h and 70 km/h Speed Limit Areas

For urban intersections the typical reported injury accident rates are determined from a series of prediction equations. Where turning movement counts are available, the accident rate should be predicted by accident type and approach using the prediction models and parameters in Tables 5.1 to 5.12. The total accident rate can be predicted by summing the predictions by accident type and approach. If only approach flows are available, then the total accident rate can be predicted using the parameters in Table 5.13. The predictions are for a five-year period, being the length of the accident period used to produce the models.

5.1.1.1 *Signalised Cross-roads*

The accident rates at signalised cross-roads are predicted by accident type and approach using the equations in Table 5.1 and the parameters in Table 5.2. Figure 5.1 illustrates the different conflicting and approach flows at crossroads.

Table 5.1 Signalised Cross-road Accident Prediction Equations

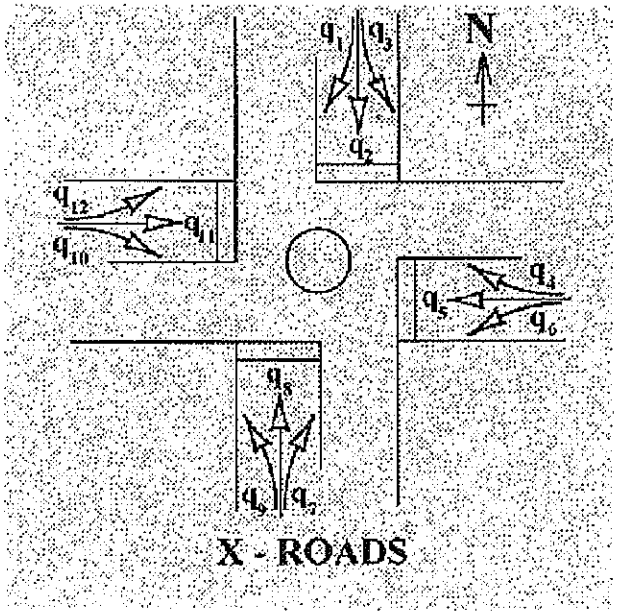
Accident Type	LTSA	Equation (accidents per approach)
	HA	$A = b_0 * q_2^{b1} * q_{11}^{b2}$
Right Turn Against	LB	$A = b_0 * q_2^{b1} * q_7^{b2}$
Rear-end	FA to FE	$A = b_0 * Q_e^{b1}$
Loss-of –control	C & D	$A = b_0 * Q_e^{b1}$
Others		$A = b_0 * Q_e^{b1}$

Figure 5.2 Conflicting and Approach Flow Types (Cross-roads)

Accident Type	LTSA Codes	b_0	b_1	b_2	K
Crossing (No	HA	$1.00E^{-3}$	0.34	0.37	1.1
Right Turn Against	LB	$4.85E^{-4}$	0.49	0.41	1.9
Rear-end	FA to FD	$8.52E^{-6}$	1.07	-	1.7
Loss-of –control	C & D	$1.56E^{-5}$	0.94	-	0.8
Others		$6.11E^{-3}$	0.46	-	1.5

* is the Gamma Shape Parameter. This is required when using Empirical Bayes Method (see section 7).

Table 5.2 Signalised Cross-roads – Prediction Model Parameters



5.1.1.2 *Roundabouts*

The accident rates at 4-arm roundabouts are predicted by accident type and approach using the equations in Table 5.3, and the parameters in Table 5.4. The circulating flow (Q_c) is the traffic that the entering flow (Q_e) at each roundabout approach must give-way to.

Table 5.3 Roundabout Accident Prediction Equations

Accident Type	LTSA Codes	Equation (accidents per approach)
Entering vs Circulating	HA, LB, JA,	$A = b_0 * Q_e^{b1} * Q_c^{b2}$
Rear-end	FA to FD	$A = b_0 * Q_e^{b1}$
Loss-of–control	C & D	$A = b_0 * Q_e^{b1}$
Others		$A = b_0 * Q_e^{b1}$

Table 5.4 Roundabouts – Prediction Model Parameters

Accident Type	LTSA Codes	b_0	b_1	b_2	K values
Entering vs Circulating	HA, LB, JA,	$4.46E^{-4}$	0.42	0.45	1.2
Rear-end	FA to FD	$2.88E^{-6}$	1.19	-	0.6
Loss-of –control	C & D	$1.51E^{-3}$	0.55	-	0.8
Others		$1.14E^{-2}$	0.26	-	0.4

5.1.1.3 *Priority cross-roads (give-way and stop Control)*

The accident rates at priority cross-roads are predicted by accident type and approach using the equations in Table 5.5 and the parameters in Table 5.6. At priority cross-roads the straight through flows are differentiated into those with priority (q_p) and

those which have to give-way ($q_{g/w}$), or stop. For the crossing (no turns) accidents, both the q_2 and q_{11} flows are used as predictors, but their order in the equation depends on their priority.

Table 5.5 Priority Cross-road Accident Prediction Equations

Accident Type	LTSA Codes	Equation (accidents per approach)
Crossing (No Turns)	HA	$A = b_0 * q_{g/w}^{b1} * q_p^{b2}$
Right Turn Against	LB	$A = b_0 * q_2^{b1} * q_7^{b2}$
Crossing (Vehicle Turning)	JA	$A = b_0 * q_2^{b1} * q_4^{b2}$
Loss-of –control	C & D	$A = b_0 * Q_e^{b1}$
Others		$A = b_0 * Q_e^{b1}$

Table 5.6 Priority Cross-roads – Prediction Model Parameters

Accident Type	LTSA Codes	b_0	b_1	b_2	K values
Crossing (No Turns)	HA	$1.95E^{-3}$	0.38	0.37	1.2
Right Turn Against	LB	$3.75E^{-3}$	0.05	0.53	0.5
Crossing (Vehicle Turning)	JA	$5.40E^{-7}$	1.13	0.44	3.0
Loss-of –control	C & D	$5.22E^{-3}$	0.30	-	0.3
Others		$1.87E^{-3}$	0.57	-	2.1

5.1.1.4 *Signalised T-junctions*

The accident rates at signalised T-junctions are predicted by accident type and approach using the equations in Table 5.7 and the parameters in Table 5.8. Figure 5.2 illustrates the different conflicting and approach flows at T-junctions.

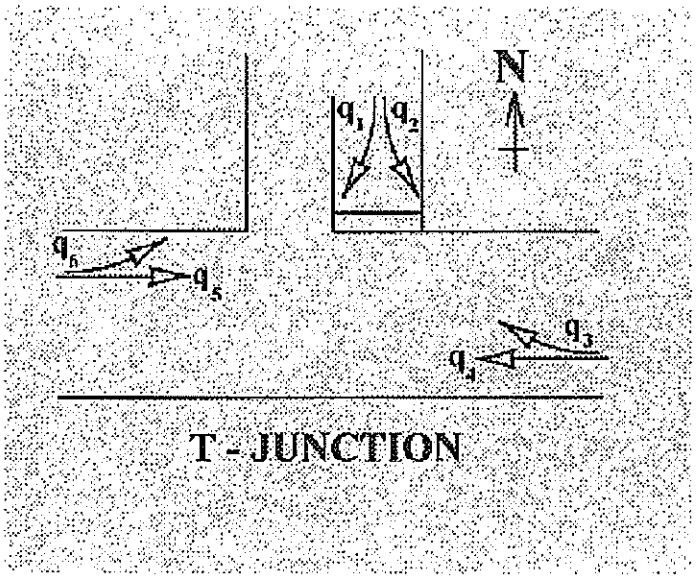
Table 5.7 Signalised T-junction Accident Prediction Equations

Accident Type	LTSA Codes	Equation (accidents per approach)
Right Turn Against	LB	$A = b_0 * q_5^{b1} * q_3^{b2}$
Rear-end	FA to FD	$A = b_0 * Q_e^{b1}$
Crossing (vehicles turning)	JA	$A = b_0 * q_5^{b1} * q_1^{b2}$
Loss of control	C & D	$A = b_0 * Q_e^{b1}$
Others		$A = b_0 * Q_e^{b1}$

Table 5.8 Signalised T-junctions – Prediction Model Parameters

Accident Type	LTSA Codes	b_0	b_1	b_2	K values
Right Turn Against	LB	0.583	-0.43	0.60	3.0
Rear-end	FA to FD	$3.83E^{-7}$	1.45	-	0.5
Crossing (Vehicles Turning)	JA	0.161	-0.34	0.51	1.2
Loss-of –control	C & D	$9.37E^{-3}$	0.17	-	3.0
Others		$8.47E^{-2}$	0.15	-	2.4

Figure 5.2 - Conflicting and Approach Flow Types (T-junction)



The large negative parameters in the 'LB' and 'JA' models indicate that intersections with higher flows have fewer accidents. This is an unexpected result. It is speculated that at high right turning flows the installation of right turn bays and exclusive right turn phases reduces accident occurrence. Further research examining these other variables is required.

5.1.1.5 *Priority T-junctions*

The accident rates at priority T-junctions are predicted by accident type and approach using the equations in Table 5.9 and the parameters in Table 5.10.

Table 5.9 Priority T-junction Accident Prediction Equations

Accident Type	LTSA Codes	Equation (accidents per approach)
Right Turn Against	LB	$A = b_0 * q_5^{b1} * q_3^{b2}$
Rear-end	FA to FD	$A = b_0 * Q_e^{b1}$
Crossing (Vehicles Turning)	JA	$A = b_0 * q_5^{b1} * q_1^{b2}$
Loss-of –control	C & D	$A = b_0 * Q_e^{b1}$
Others		$A = b_0 * Q_e^{b1}$

Table 5.10 Priority T-junctions Prediction Model Parameters

Accident Type	LTSA Codes	b_0	b_1	b_2	K values
Right Turn Against	LB	$3.33E^{-6}$	0.48	0.42	1.5
Rear-end	FA to FD	$1.45E^{-6}$	1.18	-	0.5
Crossing (Vehicles Turning)	JA	$3.60E^{-5}$	0.93	0.22	1.0
Loss-of –control	C & D	$8.22E^{-3}$	0.30	-	3.0
Others		$2.49E^{-3}$	0.51	-	3.0

5.1.1.6 *Uncontrolled T-junctions*

The accident rates at uncontrolled T-junctions (that is T-junctions that have no give-way, stop or signal controls) are predicted by accident type and approach using the equations in Table 5.11 and the parameters in Table 5.12.

Table 5.11 Uncontrolled T-junction Accident Prediction Equations

Accident Type	LTSA Codes	Equation (accidents per approach)
Right Turn Against	LB	$A = b_0 * q_5^{b1} * q_3^{b2}$
Rear-end	FA to FD	$A = b_0 * Q_e^{b1}$
Crossing (Vehicles Turning)	JA	$A = b_0 * q_5^{b1} * q_1^{b2}$
Loss-of –control	C & D	$A = b_0 * Q_e^{b1}$
Others		$A = b_0 * Q_e^{b1}$

Table 5.12 Uncontrolled T-junctions Prediction Model Parameters

Accident Type	LTSA Codes	b_0	b_1	b_2	K values
Right Turn Against	LB	$1.49E^{-3}$	0.31	0.42	3.0
Rear-end	FA to FD	$8.69E^{-8}$	1.50	-	0.7
Crossing (Vehicles Turning)	JA	$3.62E^{-4}$	0.22	0.81	3.0
Loss-of –control	C & D	$2.51E^{-3}$	0.31	-	4.0
Others		$6.27E^{-3}$	0.41	-	0.4

5.1.1.7 *Product-of-Link-Flow Models*

The models in this section predict the accident rate at an intersection from the link (two-way) flows on each of the intersecting roads. These models should be used only when turning movement counts are not available, or can not be predicted using transport models.

These models should not be used when the volume of traffic on opposite arms of an intersection differs by more than 25% of the higher flow. If the majority of traffic on a link turns left or right at a cross-roads intersection, so that the opposing arm has a lot less traffic, then this type of model is inappropriate. Where volumes on both approaches of a link are available then the two approach flows should be summed to calculate the link volume.

The total reported accident rate for each intersection types is determined using the equation:

$$A_T = b_0 * Q_{minor}^{b1} * Q_{major}^{b2}$$

where Q_{minor} is the lowest of the two-way link volumes for cross-roads, and the stem flow for T-junctions.

Table 5.13 Product-of-Link-Flow Models

Intersection Type	b_0	b_1	b_2	K values
Signalised Cross-roads	$2.04E^{-2}$	0.14	0.45	3.0
4 –arm Roundabout	$1.81E^{-3}$	0.48	0.37	3.0
Priority Cross-roads	$7.09E^{-3}$	0.51	0.21	2.3
Signalised T-junctions	0.778	0.13	0.04	3.0
Priority T-junction	$3.70E^{-4}$	0.19	0.75	3.0
Uncontrolled T-junction	$1.44E^{-2}$	0.19	0.36	2.6

5.2 Urban Mid-block Sections, 50 km/h and 70 km/h Speed Limit Areas

For urban arterial, collector and local mid-block accidents, average injury accident rates can be associated with speed limit, roadside development and for arterials the presence of a solid median. The accident types predicted for urban mid-blocks sections, and the model types, are given in Table 5.14. The flow variable used in all models is the two-way traffic volume per day (Q_T).

Table 5.14 Urban Mid-block Accident Prediction Equations

Accident Type	LTSA Codes	Equation (accidents per approach)
Rear-end (both straight)	FA to FF	$A = b_0 * Q_T^{\ b1}$
Rear-end (one turning right)	GC to GE	$A = b_0 * Q_T^{\ b1}$
Loss-of–control	C & D	$A = b_0 * Q_T^{\ b1}$
Manoeuvring & Hit Object	M & E	$A = b_0 * Q_T^{\ b1}$
Other		$A = b_0 * Q_T^{\ b1}$

Accident prediction models and parameters for the major accident types are given for arterials, collectors and local streets in Tables 5.15 to 5.17.

Table 5.15 Urban Arterials, 50 and 60 km/h Areas

	Commercial/Industrial		Residential		K values
Accident Type	b_0	b_1	b_0	b_1	
Rear-end (both straight)	$6.93E^{-7}$	1.59	$6.03E^{-7}$	1.59	1.3
Rear-end (one turning right)	$3.21E^{-3}$	0.64	$2.25E^{-3}$	0.64	0.8
Loss-of–control	$4.07E^{-4}$	0.90	$5.88E^{-4}$	0.90	1.5
Manoeuvring & Hit Object	$2.92E^{-2}$	0.45	$2.13E^{-2}$	0.45	0.8
Other	$1.07E^{-5}$	1.34	$8.68E^{-6}$	1.34	1.2

Table 5.16 Urban Collectors, 50 km/h Areas

	Commercial/Industrial		Residential		K values
Accident Type	b_0	b_1	b_1	b_0	
Rear-end (both straight)	$4.32E^{-8}$	1.96	$2.42E^{-8}$	1.96	1.5
Rear-end (one turning right)	$1.78E^{-3}$	0.70	$1.36E^{-3}$	0.70	1.8
Loss-of–control	$8.94E^{-2}$	0.25	0.145	0.25	3.0
Manoeuvring & Hit Object	$6.22E^{-4}$	0.98	$3.82E^{-4}$	0.98	2.9
Other	$3.10E^{-4}$	0.93	$4.46E^{-4}$	0.93	3.0

Table 5.17 Urban Local Streets, 50 km/h Areas

Accident Type	b_0	b_1	K values
Rear-end (both straight)	$1.46E^{-4}$	1.13	0.1
Rear-end (one turning right)	$1.25E^{-3}$	0.90	0.1
Loss-of–control	$1.99E^{-2}$	0.61	0.1
Manoeuvring & Hit Object	$2.73E^{-4}$	1.12	0.2
Other	$8.56E^{-5}$	1.33	0.2

5.3 Rural Intersections, 80 km/h and 100 km/h Speed Limit Areas

The typical reported injury accident rates (per year) for rural T-junction intersections are calculated by using the urban intersection prediction equations with 'rural' parameters (Tables 5.18). Where turning movement counts are available, the accident rate should be predicted by accident type and approach. The total accident rate can then be predicted by summing the predictions by accident type and approach.

Where only approach flows are available, then the total accident rate can be predicted using the parameters at the bottom of each table, and the 'product-of-link-flow' equation (see Section 5.1.1 [g]).

Table 5.18 Rural Priority & Uncontrolled T-junctions Prediction Model Parameters

Accident Type	LTSA Codes	b_0	b_1	b_2	K
Right Turn Against	LB	$1.21E^{-7}$	0.54	1.63	**3.0**
Crossing (Vehicles Turning)	JA	$1.98E^{-4}$	0.34	0.93	3.0
Turning verses same direction	G	$1.13E^{-3}$	0.58	-	0.4
Others		$6.23E^{-3}$	0.34	-	3.0
Total (Product-of-link-flows)	All	$1.23E^{-3}$	0.53	0.42	3.0

Models for Rural X-roads were not developed, as traffic volume data was not readily availble.

5.4 Rural Mid-block Sections, 80 km/h and 100 km/h Speed Limit Areas

For rural highways (both Transit NZ and district) and local streets (all other 80 and 100 km/h streets), the average injury accident rates can be associated with the terrain type (flat, rolling and mountainous). The accident types predicted for rural mid-blocks sections, and the model types, are given in Table 5.19. The flow variable used in all models is the two-way traffic volume per day (Q_T). In the head-on model it is assumed that the traffic vehicle split by direction is approximately 50:50 over 24-hours.

Table 5.19 Rural Mid-block Accident Prediction Equations

Accident Type	LTSA Codes	Equation (accidents per approach)
Head-on	B	$A = b_0*((Q_T/2)^2)^{b1}$
Overtaking	A	$A = b_0*((Q_T/2)^2)^{b1}$
Rear-end (both straight)	FA to FF	$A = b_0*Q_T^{\ b1}$
Rear-end (one turning right)	GC to GE	$A = b_0*Q_T^{\ b1}$
Loss-of–control	C & D	$A = b_0*Q_T^{\ b1}$
Manoeuvring & Hit Object	M & E	$A = b_0*Q_T^{\ b1}$
Other		$A = b_0*Q_T^{\ b1}$

The accident prediction model parameters for the major accident types are given for rural highways and motorways/expressways in Tables 5.20 and 5.21. Insufficient data was available for rural local roads.

Table 5.20 Rural Highway Accident Prediction Equations

	Level		Rolling		K values
Accident Type	b_0	b_1	b_0	b_1	
Head-on	$9.54E^{-4}$	0.33	$1.24E^{-3}$	0.33	3.0
Overtaking	$5.15E^{-6}$	0.65	$1.58E^{-6}$	0.65	2.2
Rear-end (both straight)	$1.09E^{-7}$	1.72	$9.89E^{-9}$	1.72	1.7
Rear-end (one turning right)	$4.25E^{-4}$	0.78	$6.27E^{-5}$	0.78	1.4
Loss-of–control	$2.83E^{-2}$	0.48	$1.82E^{-2}$	0.48	1.2
Manoeuvring & Hit Object	$4.00E^{-3}$	0.52	$1.07E^{-3}$	0.52	3.0
Other	$1.68E^{-4}$	0.84	$1.22E^{-5}$	0.84	3.0

Table 5.21 Motorways and Expressways

Accident Type	b_0	b_1	K values
Rear-end	$5.91E^{-9}$	1.88	3.0
Loss-of–control	$2.81E^{-5}$	1.11	3.0
Overtaking	$1.25E^{-5}$	1.10	1.6
Other	$2.65E^{-2}$	0.41	0.5

5.5 Goodness-of-fit Statistics

Goodness-of-fit statistics have been developed for all the crash prediction models. Table 5.22 to 5.24 summarises the results of the analysis. It has been necessary to group data into flow bands to overcome the problem associated with a low mean value. The appropriate group size is dependent on the mean of the accident data and the value of k (the Gamma shape parameter).

All models that have a scaled deviance below the chi-squared value are considered to fit at a 5% significance level. This means that the model explains a significant amount of the variability in the data. The lower the scaled deviance in relation to the critical chi-squared value the better the model explains the variability in the data.

These tables indicate that in all but one case the prediction models are significant at the 95% level of confidence. The only exception is right-turn-against accidents at priority intersections. Research by Turner (1995) shows that at priority and uncontrolled T-junctions there is confusion caused by the New Zealand left-turn give-way rule. In such circumstances the volume of other turning movements has been shown to be significant. The theory being those drivers have to be aware of whether they need to give-way to such movements. So, for this accident type, a larger model, including three or more flow variables should be considered in future research.

5.6 Confidence Intervals

Confidence intervals (95%) have been derived for the accident predictions, using the Minitab macros and worksheets developed by Wood (Transit NZ, 1993) and Turner (1995).

Confidence Interval Graphs

Confidence intervals graphs have been produced for both the one and two predictor variable models. The confident intervals were produced using the following formulas, which are derived in Appendix A of the stage 1 report.
An approximate 95% confidence interval for $b_{0\,x}{}^{b1}$ is:

Lower Limit =

$$\exp\{b_0' + b_1 x - 1.96\sqrt{(X'WX)_{11}^{-1} + 2\log x(X'\ WX)_{12}^{-1} + (\log x)^2 (X'WX)_{22}^{-1}}\}$$

Upper Limit =

$$\exp\{b_0' + b_1 x + 1.96\sqrt{(X'WX)_{11}^{-1} + 2\log x(X'WX)_{12}^{-1} + (\log x)^2 (X'WX)_{22}^{-1}}\}$$

where,

$$(X'WX)^{-1} = \begin{bmatrix} (X'WX)_{11}^{-1} & (X'WX)_{12}^{-1} \\ (X'WX)_{21}^{-1} & (X'WX)_{22}^{-1} \end{bmatrix}$$

and (X'WX)-1 is the inverse of the Fisher Information Matrix (refer to generalised linear modelling theory in Turner (1995)).

For the one variable models, the confidence intervals can be plotted, along with the fitted regression curve, on an accident versus traffic volume graph. For the two variable models the confidence intervals are areas, rather than a single line. The confidence areas can be plotted on a 3-D graph.

Interpretation

The confidence intervals/areas show, for each approach of every intersection type, the range of the 'underlying true accident rates' that would be expected 95% of the time. When compared with the raw accident observations, the confidence intervals don't appear to contain 95% of the data points, especially when the number of zero observations are considered. It must be remembered that while the accident count at a site might be zero the 'underlying true accident rate' is not zero, as there is always some risk if there is traffic travelling through the intersection, that an accident might occur. A better appreciation of the value of the 'underlying true accident rates' and their relationship to the confidence interval can be seen when the data are grouped into different flow bands and plotted on the confidence interval graphs. For convenience we have used the grouped data derived using the goodness-of-fit testing, and plotted it on the confidence interval graphs.

The confidence intervals are wider at high traffic volumes (and sometimes low traffic volumes as well). The variable width of these bands confirm what has previously been speculated: that accident predictions from the models are not particularly accurate at high or low traffic volumes, and that the most accurate estimates are at mid-range traffic flows, as the majority of intersection have flows in the mid-range. Users of the models should be aware of the ranges over which the model predictions are most accurate.

Table 5.22 Urban Intersections: Goodness-of-Fit Statistics

Intersection Type	Crash Type	Model Type P or NB#	'k' value	Mean	Group Size*	Degree-of-Freedom	Chi Squared	Scaled Deviance	Significant Model
Signalised X-road	Crossing (no turns)	NB	1.1	0.37	5	85	108	41	Yes
	Right Turn Against	NB	1.5	0.52	4	107	132	52	Yes
	Rear-end	NB	1.7	0.13	13	32	46	16	Yes
	Loss-of-Control	NB	0.7	0.07	30	14	22	10	Yes
	Other	NB	1.5	0.37	5	85	108	37	Yes
	Crash Type TOTAL	NB	3.0	6.19	1	78	100	37	Yes
4-arm Roundabout	Entering Vs Circulating	NB	1.2	0.61	4	53	71	30	Yes
	Rear-end	NB	0.6	0.07	27	7	14	4	Yes
	Loss-of-Control	NB	0.8	0.15	13	15	25	8	Yes
	Other	NB	0.4	0.10	19	10	18	3	Yes
	Crash Type TOTAL	NB	3.0	4.28	1	34	49	15	Yes
Priority X-road	Crossing (no turns)	NB	1.2	0.44	5	59	78	64	Yes
	Right Turn Against	NB	0.5	0.13	15	19	30	11	Yes
	Crossing (veh. Turning)	NB	3.0	0.06	25	11	20	3	Yes
	Loss-of-Control	NB	0.3	0.06	30	9	17	6	Yes
	Other	NB	2.1	0.19	9	32	46	18	Yes
	Crash type TOTAL	NB	2.3	3.15	1	38	53	21	Yes

Signalised T-Junction	Right Turn Against	NB	3.0	1.03	2	13	22	9	Yes
	Rear-end	NB	0.7	0.22	9	8	16	10	Yes
	Crossing (veh. Turning)	NB	3.0	0.40	4	6	13	5	Yes
	Loss-of-Control	NB	4.0	0.04	30	1	3.8	3.7	Marginal
	Other	NB	0.4	0.33	8	10	18	9	Yes
	Crash type TOTAL	NB	3.0	3.43	1	28	41	30	Yes
Priority T-Junction	Right Turn Against	NB	1.5	0.37	5	16	26	103	No
	Rear-end	NB	0.5	0.09	20	12	21	13	Yes
	Crossing (veh. Turning)	NB	1.0	0.51	4	21	33	10	Yes
	Loss-of-Control	NB	3.0	0.10	16	15	25	6	Yes
	Other	NB	3.0	0.20	9	28	41	14	Yes
	Crash Type TOTAL	NB	3.0	2.12	1	87	110	40	Yes
Uncontrolled T-Junctions	Right Turn Against	NB	3.0	0.39	5	9	17	4	Yes
	Rear-end	NB	0.7	0.06	24	5	11	1	Yes
	Crossing (veh. Turning)	NB	3.0	0.47	4	11	20	8	Yes
	Loss-of-Control	NB	4.0	0.03	30	4	10	2	Yes
	Other	NB	0.4	0.20	11	12	21	5	Yes
	Crash Type TOTAL	NB	2.6	1.84	2	24	36	16	Yes

where P is Poisson and NB is Negative Binomial

Table 5.23 Urban Mid-blocks: Goodness-of-Fit Statistics

Mid-block Type	Crash Type	Model Type P or NB#	'k' value	Mean	Group Sample Size*	Degree-of-Freedom	Chi Squared 95% level	Scaled Deviance	Significant Model Yes/N
Urban Arterials	Rear-End (both straight)	NB	1.3	2.53	2	81	103	37	Yes
	Rear-End (one turning right)	NB	0.8	0.98	3	53	71	29	Yes
	Loss-of-Control	NB	1.5	2.62	2	81	103	37	Yes
	Manoeuvring	NB	0.8	1.52	2	81	103	42	Yes
	Other	NB	1.2	3.14	2	81	103	42	Yes
Urban Collectors	Rear-End (both straight)	NB	1.5	0.62	3	48	65	20	Yes
	Rear-End (one turning right)	NB	1.8	0.52	4	36	51	14	Yes
	Loss-of-Control	NB	3.0	1.11	2	73	94	41	Yes
	Manoeuvring	NB	2.9	1.67	2	73	94	37	Yes
	Other	NB	3.0	1.09	2	73	94	37	Yes
Urban Local Roads	Rear-End (both straight)	NB	0.1	0.52	11	9	17	6	Yes
	Rear-End (one turning right)	NB	0.1	0.82	10	10	18	3	Yes
	Loss-of-Control	NB	0.1	1.46	8	13	22	4	Yes
	Manoeuvring	NB	0.2	0.87	7	15	25	14	Yes
	Other	NB	0.2	1.34	5	22	34	14	Yes

Table 5.24 Rural Intersections and Routes: Goodness-of-Fit Statistics

Intersection Type	Crash Type	Model Type P or NB#	'k' value	Mean	Group Sample Size*	Degree-of-Freedom	Chi Squared 95% level	Scaled Deviance	Significant Model Yes/N
Rural T-Junctions	Right Turn Against	NB	3.0	0.08	18	1	4	1	Yes
	Turning verses same direction	NB	0.4	0.07	27	4	10	3	Yes
	Crossing (veh. Turning)	NB	3.0	0.34	5	8	16	3	Yes
	Other	NB	3.0	0.07	20	6	13	1	Yes
	Crash Type TOTAL	NB	3.0	0.84	2	23	35	12	Yes
Rural Highways	Head-on	NB	3.0	0.17	10	8	16	5	Yes
	Overtaking	NB	2.2	0.15	12	7	14	4	Yes
	Rear-End (both straight)	NB	1.7	0.37	5	18	29	5	Yes
	Rear-End (one turning right)	NB	1.5	0.22	10	8	16	6	Yes
	Loss-of-Control	NB	1.5	0.97	3	32	46	21	Yes
	Manoeuvring	NB	3.0	0.22	10	8	16	4	Yes
	Other	NB	3.0	0.15	12	7	14	4	Yes
Motorways	Rear-End	NB	3.0	3.15	1	74	95	35	Yes
	Loss-of-Control	NB	3.0	3.83	1	74	95	41	Yes
	Overtaking	NB	1.6	1.49	2	36	51	17	Yes
	Other	NB	0.5	1.92	3	24	36	31	Yes

6 Proposed Changes to PEM Accident Analysis Procedures

A proposed update to sections A6.4 of the Transfund PEM has been prepared, and included in Appendix C. Most of the existing models and rates, except those for curves, have been removed and replaced with new accident prediction models. While the size of this section of the manual will need to be increased significantly, this is thought to be worthwhile, because of the greater precision of the generalised linear models and the additional information, in particular accidents by types, which can be calculated. The UK and many of the European countries have switched to generalised models, based on the conflicting flows for accident prediction, because of this greater precision.

Intersection prediction models are to be included in the PEM for both individual accident types (conflicting flow models) and total accidents (product of route flow models). While our preference is for the first type of model, it is acknowledged that in the absence of turning volumes, the second type of model should be available.

The parameters for these new models are based on accident data from the period 1995 to 1999. To produce annual prediction models from the models presented in Section 5, we have divided the coefficient (or b_0) by five.

7 Empirical Bayesian Method

The field of statistics can be divided into two streams, 'conventional' methods and 'Bayesian' methods. Bayesian, like conventional methods, assume that any parameter in a problem (such as the true accident rate at a blackspot in our case) can be regarded as the value of a random variable having a probability distribution. The difference in the two methods being that in Bayesian statistics the probability of the parameter(s) is normally estimated before any data becomes available, with this being known as the *prior* distribution of the parameter. When data becomes available the Bayes theorem is used to convert the prior distribution of the parameter into a *posterior* distribution. If further data becomes available, this posterior distribution becomes the prior and a new posterior distribution is produced. Generally the effect of the prior distribution diminishes quickly as it is updated with more and more observation data.

Most applications of Bayesian methods are considered controversial because the prior distribution is based on an investigator's personal beliefs (or judgement) about the possible values of the parameter, with preferred values having a higher degree of belief.

The empirical Bayes method is a 'hybrid' approach that uses elements of both the conventional and Bayesian methods. It is not considered controversial, and in a sense is not a true Bayesian method, as it does not rely completely on a subjective evaluation of the prior distribution.

In accident prediction modelling (where accidents have a Negative Binomial distribution) the prior distribution is generated using data from a number of typical sites, the conventional accident prediction models (or base models). The prior is then converted to the posterior distribution using Bayes theorem and accident data from a particular site. In this way factors that are not captured in the prediction models can be factored into the accident predictions for particular sites. This method allows both sources of information, typical accident rates for the population of sites and the local accident history, to be utilised in generating accident predictions.

The empirical Bayes method detailed in this section is reliant on a number of assumptions, which include:

- That the annual accident counts at a particular site are Poisson distributed about a constant true accident rate (m) over the accident period.
- That the accident count in each year of the accident period is independent.
- That the true accident rate 'm' varies from site to site.
- The prior distribution of 'm' is described by a gamma probability density function $f_o(m)$.

The combined accident estimate (Z) is given by the following equation:

$$Z = a\,E(m) + (1-a)\,X \quad \text{Equation 7.1}$$

$$\text{where, } a = (1 + Var(m) / E(m))^{-1}$$

E(m) is the accident prediction produced from the base model and 'X' is the accident count from historical accident records. 'a' is the weighting placed on the base model and accident history.

Now E (m) = y, the accident prediction from the base model
and Var (m) = y^2/k, where k is the shape parameter of the gamma distribution

Substituting in the values of E(m) & Var(m) Equation 7.1 becomes:

$$Z = y * (1 + y/k)^{-1} + X * (1 - (1 + y/k)^{-1}) \qquad \text{Equation 7.2}$$

When the accident prediction period (i-years) differs from the accident history period (j-years), Equation 7.3 should be used.

$$Z = y * (j/i) * (1 + (j/i)\, y/k)^{-1} + X * (1 - (1 + (j/i)\, y/k)^{-1}) \qquad \text{Equation 7.3}$$

Example 1 (theoretical)
Proposed changes at the intersection of Colombo Street and Brougham Street (in Christchurch) are expected to change the right turning intersection volumes. The expected flow changes have been predicted using a traffic network model, as indicated in Table 7.1.

Table 7.1 – Right Turn Volumes Before and After Network Change (24hr)

Approach	Right Turn Flow Before	Straight Flow Before	Right Turn Flow After	Straight Flow After
North	747	4784	500	4784
East	577	14759	400	14759
South	830	4075	600	4075
West	2440	13971	700	13971

What is the expected change in the right-turn-against accident rate at this intersection? The accident history shows that there have been five right-turn against accidents observed during the last five years (1995 to 1999), with 2 on the east and west approaches and one on the north approach.

Table 7.2 (next page) shows the predicted accident rate both before and after the traffic flow change with and without the inclusion of accident history data. The base predictions were generated using the model in Table 5.1.

The results in Table 7.2 show that there can be an appreciable difference in the accident predictions calculated using the accident history, compared with those calculated only using the accident prediction models. For the example site the use of the latter method would have resulted in underestimating the benefits resulting from a reduction in the right turning movements.

For additional information on this application of the empirical Bayes method we would refer the reader to Abbess, Jarrett & Wright (1981), Mountain & Fawaz (1989 & 1991) and Lau & May (1989).

Table 7.2 Accident Predictions Before and After Network Changes (in 5 years)

	Observed Accidents	Accident Prediction without accident history data		Accident Prediction with accident history data	
Approach	Before	Before	After	Before	After *
North	1	0.40	0.33	0.51	0.42
East	2	0.56	0.47	0.89	0.74
South	0	0.40	0.34	0.33	0.28
West	2	1.11	0.60	1.44	0.75
Total	5	2.47	3.16	1.74	2.22
Difference		0.73		0.94	

* In the absence of 'after' accident data, this rate has been calculated based on the percentage difference for the 'before' period between predictions with and without accident history. Further research is required to establish whether this is an appropriate method.

References

Abbess. C, Jarett D and Wright C. C. (1981) Accidents at Blackspots: Estimating the effectiveness of remedial measures, with special reference to the 'regression-to-mean' effect, *Traffic Engineering and Control 22 (10)*: 535-542.

Beca Carter Hollings & Ferner Ltd (1994) Motorway Accidents and Costs, unpublished report to Transit NZ

Bowman B. L, Vecellio R. L and Miao J. (1995) Vehicle and Pedestrian Accident Prediction Models for Median Locations, *Journal of Transportation Engineering,* Nov 1995: 531 - 537.

Gabites Porter Partners (1991) System Wide Accident Analysis: A Study of Five Major Cities in New Zealand (unpublished) Transit NZ

Hauer, E, Ng J. C. N, Lovell, J (1989) Estimation of Safety at Signalised Intersections, *Transportation Research* Record 1185: 48 – 61.

Jackett M. (1992) Accident Rates for Urban Routes, IPENZ Conference Proceedings, pp 133-144

Jackett M. (1993) Accident Rates on Urban Routes – 1992 Update, Institute of Professional Engineers of NZ, Transactions V20(1) pp 10-16

Kitto H.J. (1980) Accident Rates at Urban Right Angled Intersections, *Road Research Unit Bulletin No. 48.*

Land Transport Safety Authority (1994) Motor Accidents in New Zealand – 1993, Research and Statistics Strategic Support, Land Transport Safety Authority.

Land Transport Safety Authority (1997) Motor Accidents in New Zealand – 1996, Research and Statistics Strategy Division, Land Transport Safety Authority.

Lau M and May A (1989) Accident prediction model development for unsignalised intersections: final report, Research Report, Institute of Transportation Studies, University of California, Berkeley

Lord D and Persaud B. N. (1999) Accident Prediction Models With and Without Trend: Application of the Generalised Estimating Equations (GEE) Procedures, Transportation Research Board 79th Annual Meeting, Washington Paper N° 00-0496.

Lord, D (1999) Procedure to Estimate Missing Traffic Counts at Intersections. Unpublished Paper, Safety Studies Group, Department of Civil Engineering, University of Toronto, Toronto, Canada.

Maher, M. J. (1989) Fitting Predictive Equations to Accident Data when there is uncertainty in the traffic flow estimates, Proc. of 17th annual PTRC Summer Meeting, Seminar H (Transport Planning), University of Sussex, England.

Maher M. J and I Summersgill (1996) A Comprehensive Methodology for the Fitting of Prediction Accident Models, *Accident Analysis and Prevention 28*: 281 – 296.

Mountain L and Fawaz (1989) The area-wide effects of engineering measures on road accident occurrence, *Traffic Engineering and Control, July/August 1989.*

Mountain L and Fawaz (1991) The accuracy of estimates of expected accident frequencies obtained using an Empirical Bayes approach, *Traffic Engineering and Control, May 1991.*

Nicholson A. J and Turner S. A.(1996) Estimating Accidents in a Road Network, Proceedings ARRB Conference, Part 5, pp.51 - 66, Christchurch, NZ.

Opus International Consultants Ltd (1997) Typical Accident Rates for Rural Passing Lanes and Unsealed Roads, Transit NZ Research Report 89.

Poch and Mannering (1996) Negative Binomial Analysis of Intersection – Accident Frequencies, *Journal of Transportation Engineering, March/April 1996.*

Transit NZ (1993) Development of Models Relating Frequency of Collision to Traffic flows – Stage II, Transit NZ Technical Report, pp. 1-19.

Trinada and Rengaraju (1998) Modelling conflicts of Heterogeneous traffic at Urban Uncontrolled Intersections, *Journal of Transportation Engineering, Jan/Feb 1998*

Turner S. A. (1995) Estimating Accidents in a Road Network, PhD Thesis, Dept of Civil Engineering, University of Canterbury.

Turner S. A and Nicholson A. J. (1998a) Intersection Accident Estimation: The Role of Intersection Location & Non-Collision Flows, *Accident Analysis and Prevention 30*: 505 – 517.

Turner S. A and Nicholson A. J. (1998b) Using Accident Prediction models in area wide Crash Reduction Studies, Proceedings of the 9th REAAA Conference, Wellington.

Wood G. R. (2000a). Assessing goodness-of-fit for Poisson and negative binomial models with low mean (submitted for publication) Institute of Information Sciences and Technology, Massey University, Palmerston North.

Wood G. R. (2000b) Generalised linear accident models and goodness-of-fit testing (submitted for publication), *Accident Analysis and Prevention*, in.press.

Appendix A
Intersection and Link Locations

INTERSECTIONS

CROSSROADS

Signalised X	
City	No.
Christchurch	44
Auckland City	17
Wellington	8
Dunedin	8
Lower Hutt	3
Hamilton	11
North Shore City	2
Waitakere City	1
Other	9
Palmerston North	6
TOTAL	**109**

Roundabout X	
City	No.
Christchurch	32
Auckland City	3
Lower Hutt	5
North Shore City	2
Waitakere City	2
Other	1
Palmerston North	10
TOTAL	**55**

Priority X	
City	No.
Christchurch	29
Auckland City	10
Wellington	5
Dunedin	9
Lower Hutt	9
Hamilton	2
North Shore Ci	1
Other	10
Palmerston No	1
TOTAL	**76**

T-JUNCTIONS

Signalised T	
City	No.
Christchurch	26
Auckland City	1
Wellington	1
Palmerston North	2
TOTAL	**30**

Priority T	
City	No.
Christchurch	39
Auckland City	7
Wellington	6
Dunedin	5
Lower Hutt	8
Hamilton	2
North Shore City	3
Waitakere City	7
Other	8
Palmerston North	4
TOTAL	**89**

T Uncontrolled	
City	No.
Christchurch	22
Auckland City	8
Wellington	4
Lower Hutt	14
Waitakere City	3
TOTAL	**51**

Rural T	
City	No.
Christchurch	20
Wellington	2
Wanganui	8
Tauranga	20
TOTAL	**50**

LINKS

Urban Arterials	
Local Body	No.
ASHBURTON	3
AUCKLAND	23
CHRISTCHURCH	17
DUNEDIN	3
GISBORNE	2
GORE	4
HAMILTON	10
HASTINGS	5
HAURAKI WAIHI	1
HAURAKIPAEROA	2
HOROWHENUA	1
INVERCARGILL	10
MANUKAU	11
MASTERTON	2
NAPIER	5
NEW PLYMOUTH	3
NTH SHORE	9
PAPAKURA	3
PORIRUA	1
RODNEY	2
ROTORUA	4
TAUPO	6
TAURANGA	5
WAIPA CMBGE	4
WAIPA TEAWAMU	2
WAITAKERE	11
WANGANUI	2
WELLINGTON	3
WHAKATANE	3
WHANGAREI	8
TOTAL	**165**

Urban Collectors	
Local Body	No.
ASHBURTON	4
BLENHEIM	4
CENTRAL HAWKES BAY	2
CHRISTCHURCH	9
DUNEDIN	2
GISBORNE	5
HASTINGS	11
HAURAKI WAIHI	2
HOROWHENUA	6
LOWER HUTT	8
MASTERTON	9
NAPIER	5
NELSON	3
NEW PLYMOUTH	3
PALMERSTON N	24
PAPAKURA	4
RODNEY	2
S.WAIKATO PUT	2
S.WAIKATO TKR	8
TIMARU	3
UPPER HUTT	13
WAIPA TEAWAMU	3
WANGANUI	8
WELLINGTON	3
WHAKATANE	6
TOTAL	**149**

Urban Local Roads	
Local Body	No.
HASTINGS	3
HASTINGS URBAN	55
HAV. NTH RURAL	1
HAV. NTH URBAN	8
MANUKAU	52
TOTAL	**119**

Rural Highways	
Local Body	No.
CENTRAL HAWKES BAY	3
CENTRAL OTAGO	10
CHRISTCHURCH	1
CLUTHA	3
DUNEDIN	7
GISBORNE	8
GORE	2
HAMILTON	5
HASTINGS	6
HAURAKIPAEROA	1
HOROWHENUA	5
INVERCARGILL	5
MACKENZIE	1
MANAWATU	10
MANUKAU	5
NEW PLYMOUTH	8
PALMERSTON N	3
QUEENSTOWN LAKES	10
RANGITIKEI	3
SOUTHLAND	1
TAURANGA	2
WAITAKI	1
TOTAL	**100**

Motorways	
Local Body	No.
AUCKLAND	25
FRANKLIN	6
MANUKAU	18
NORTH SHORE	6
OTAHUHU	2
PAPAKURA	4
WAIKATO	3
WAITAKERE	4
WELLINGTON	8
TOTAL	**76**

Appendix B
Local Authorities Questionnaire and Responses

DEVELOPMENT OF NEW ACCIDENT RATES FOR THE TRANSFUND PROJECT EVALUATION MANUAL: QUESTIONNAIRE 1

1. Have you, or one of your colleagues, or your network consultants used the Transfund Project Evaluation Manual (PEM) for a Safety Improvement Project. Yes ☐ No ☐

IF NO GO TO 3

2. Please indicate which type of improvement projects have been assessed using the Transfund PEM

(a)	Route Improvements	Motorway	☐
		Urban Arterial	☐
		Urban Collector	☐
		Urban Local Road	☐
		Rural Highway (State or Local)	☐
		Rural Local Road	☐
		Other (Please Specify)	☐
(b)	Intersection Improvements	Motorway Interchange	☐
		Signal	☐
		Roundabout	☐
		Urban Priority	☐
		Rural Priority	☐
		Rural Priority/Uncontrolled	☐
		Other (Please Specify)	☐
(c)	Other		☐

3. Do You have electronic or manual motor vehicle, cyclist or pedestrian counts for urban or rural routes. Yes ☐ No ☐

4. Do you have electronic or manual motor vehicle, cyclist or pedestrian turning volume counts for intersections (for at least the morning and evening peak hours) Yes ☐ No ☐

IF NO TO 3 & 4 GO TO 7

5. Are you willing to provide electronic/manual count data for this research project (we can provice assistance if necessary) Yes ☐ No ☐

IF YES please indicate the approximate number of sites for which count data (route counts or turning volume counts) can be provided. A site should have remained unmodified during a six year period in the 1990's. Please indicate whether data is from electronic or manual sources, and in what format.

	Site Classification	**Number**	**Type* & Format**
(a) Routes	- Motorway		
	- Urban Arterial		
	- Urban Collector		
	- Urban Local Road		
	- Rural Highway (State or Local)		
	- Rural Local Road		
	- Other (please specify)		

		Number	**Type & Format**
(b) Intersections	- Motorway Interchange		
	- Traffic Signals		
	- Roundabout		
	- Urban Priority (Stop or Giveway)		
	- Urban Un-controlled		
	- Rural Priority (Not State Highway)		
	- Rural Un-controlled (Not State Highway)		
	- Rural (State Highway)		
	- Others (please specify)		

		Number	**Type & Format**
(c) Others	- Railway Crossing		

- Narrow Bridge (one way or two way)		

If this table is too small for your comments, please provide them on an appended sheet. Examples of the count data formats would also be appreciated.

* Type is motor vehicle, cyclist or pedestrian.

6. Can you provide the following information for each site.

Site Information		All Sites	Most Sites	Some Sites	No Sites
(a) Routes	Road side development (predominate type)	☐	☐	☐	☐
	Road cross-section (e.g. median types and lane widths)	☐	☐	☐	☐
(b) Intersections	Intersection layout (layout plan or sketch)	☐	☐	☐	☐
	Approach Grades (0%, 0-3%, 3-6% & 6% plus)	☐	☐	☐	☐
	Signal phasing and controlled legs (stop or give-way)	☐	☐	☐	☐
	Lane and median widths	☐	☐	☐	☐
(c) Railway Crossing	Number of tracks	☐	☐	☐	☐
	Control type (stop or barrier arms)	☐	☐	☐	☐
	Safety equipment (signs, etc)	☐	☐	☐	☐
(d) Bridges	Approach & bridge seal widths	☐	☐	☐	☐
	Length	☐	☐	☐	☐
	Abutment & rail protection measures (sketch plan)	☐	☐	☐	☐

7. Please provide contact details below:

Name:

Position:

Phone:

Fax:

Email:

Thank you for your co-operation. Please return in prepaid envelope by:

TUESDAY 12 OCTOBER.

Appendix C
Proposed PEM Procedures

A6.4 Accident Rates

A6.4.1 Introduction

Change second bullet point to:

- typical injury accident rates and prediction models (A6.4.3 to A6.4.8)

A6.4.2 General Accident Trends

No changes required.

A6.4.3 Typical Injury Accident Rates

Changes suggested below

The typical rates **and prediction models** of reported injury accidents in Sections A6.4.4 to **A6.4.8** are the result of studies carried out for TNZ, **Transfund NZ** and LTSA. **(remove next sentence, as no longer relevant).** While there is a wide spread of results between different geographic areas, **rates are currently not available for each region (this is the subject of on-going research).** Consequently the injury accident rates presented apply nationwide.

When undertaking accident rate studies non-injury accidents shall not be included due to the inconsistency in non-injury reporting rates from district to district.

Replace sections A6.4.4 to A6.4.7 with following.

A6.4.4 Urban Intersections, 50 km/h and 70 km/h Speed Limit Areas

For urban intersections the typical reported injury accident rates (per year) are determined from a series of prediction equations. Where turning movement counts are available, the accident rate should be predicted by accident type and approach using the prediction models and parameters in Tables A6.1 to A6.12. The total accident rate can be predicted by summing the predictions by accident type and approach. If only approach flows are available, then the total accident rate can be predicted using parameters in Table A6.13 (see Section A6.4.4.7).

A6.4.4.1 Signalised Cross-roads

The accident rates at signalised cross-roads are predicted by accident type and approach using the equations in Table A6.1 and the parameters in Table A6.2. Figure A6.2 illustrates the different conflicting and approach flows at crossroads.

		1)Kaikoura District Council	2)Manukau City Council	3)Queenstown-Lakes District Council	4)Tasman District Council	5)Hamilton City Council	6)Papakura District Council	7)Waitaki District Council	8)Westland District Council	9)Opus (Hamilton)	10)Whangarei District Council	11)Selwyn District Council	12)Grey District Council	13)Hastings District Council	14)Opus (Auckland)	15)Transit NZ (Hamilton)	16)Invercargill City Council	17)Land Transport Safety Authority (HQ)	18)Clutha District Council	19)Gisborne District Council	20)Lower Hutt City Council	21)Tauranga District Council	22)Banks Peninsula District Council	23)Waimakariri District Council	24)Manukau Consultants	25)Timaru District Council	26)Opus (Lower Hutt)	27)Nelson City Council	28)Ashburton District Council
1)	Used Transfund PEM? (Y=1/N=0)		1	0	1				0	1	1			1		1	1	1					1	1		1	0		1
2a)	**Route Improvements**																												
	Motorway		-	-	-				-	-	-			-		1	-	-					-	-		-	-		-
	Urban Arterial		1	-	-				-	-	1			1		-	1	-					-	1		1	-		-
	Urban Collector		1	-	1				-	-	1			1		-	1	-					-	-		-	-		-
	Urban Local Road		1	-	-				-	-	1			1		-	1	-					-	-		-	-		-
	Rural Highway		1	-	-				-	-	1			-		1	1	-					-	-		-	1		-
	Rural Local Road		-	-	-				-	-	1			1		-	1	-					1	-		1	-		1
	Other		-	-	-				-	1	-			-		-	-	-					-	1		-	-		-
2b)	**Intersection Improvements**																												
	Motorway Interchange		-	-	-				-	-	-			-		-	-	-					-	-		-	-		-
	Signal		1	-	-				-	-	1			1		1	1	-					-	-		-	-		1
	Roundabout		1	-	1				-	-	1			1		1	1	-					-	1		1	-		1
	Urban Priority		1	-	-				-	-	1			1		1	1	-					-	-		1	-		-
	Rural Priority		-	-	-				-	-	-			1		1	1	-					-	1		-	-		-
	Rural Priority/Uncontrolled		-	-	-				-	-	-			1		-	1	-					-	-		-	-		-
	Other		-	-	-				-	-	-			-		-	-	-					-	-		-	-		-
2c)	**Other**		-	-	-				-	-				-		-		1					-	-		1	-		-
3)	Route Counts (Y=1/N=0)		1	1	1				1	1	1			1		1	1	0					1	1		1	1		1
4)	Intersection Counts (Y=1/N=0)		1	0	1				1	1	1			1		1	1	0					0	0		1	0		0
5)	Provide data? (Y=1/N=0)		1	1	1				1	1	1			1		1	1	-					0	1		1	1		1
5a)	**Routes**																												
	Motorway		-	-	-				?	?	-			-		-	-	-					-	-		-	-		-
	Urban Arterial		146	-	-				?	?	10			27		-	2	-					-	28		30	-		-
	Urban Collector		200	-	5				?	?	10			18		-	2	-					-	52		100	-		-
	Urban Local Road		200	-	5				?	?	20			8		-	2	-					-	156		100	-		-
	Rural Highway (State or Local)		-	100	-				?	?	20			11		124	2	-					-	136		-	2		-
	Rural Local Road		-	260	5				?	?	-			2		-	2	-					-	426		100	-		1
	Other		-	-	-				?	?	-			-		-	-	-					-	-		-	-		-
5b)	**Intersections**																												
	Motorway Interchange		-	-	-				?	?	-			-		-	-	-					-	-		-	-		-
	Traffic Signals		-	-	-				?	?	20			16		-	12	-					-	-		12	-		-
	Roundabout		-	-	2				?	?	4			-		-	-	-					-	-		-	-		-
	Urban Priority (Stop or Giveway)		-	-	2				?	?	10			6		-	6	-					-	-		10	-		3
	Urban Un-controlled		-	-	2				?	?	-			-		-	-	-					-	-		-	-		-
	Rural Priority (Not SH)		-	-	2				?	?	-			-		-	-	-					-	-		-	-		-
	Rural Un-controlled (Not SH)		-	-	2				?	?	-			-		-	-	-					-	-		-	-		-
	Rural (State Highway)		-	-	-				?	?	-			-		-	-	-					-	-		-	-		-
	Other		-	-	-				?	?	-			-		-	-	-					-	-		-	-		-
5c)	**Others**																												
	Railway Crossing		-	-	-				-	-	-			-		-	-	-					-	-		5	-		-
	Narrow Bridge (one or two way)		-	-	-				-	-	-			-		-	-	-					-	-		-	-		-
6a)	**Routes**																												
	Road side development		Most	-	-				Some	Some	Most			Most		Some	Most	-					-	Most		Some	All		All
	Road cross section		Some	-	All				Some	Some	Most			All		All	Most	-					-	All		Some	All		All
6b)	**Intersections**																												
	Intersection layout		Some	-	All				No	Most	Most			Some		Some	Most	-					-	Some		Some	-		All
	Approach Grades		Some	-	All				No	Most	Most			Some		Most	Some	-					-	No		All	-		All
	Signal phasing and controlled legs		No	-	-				No	Most	All			All		Some	All	-					-	No		All	-		All
	Lane and median widths		Some	-	All				No	Most	Most			All		All	All	-					-	Some		Some	-		All
6c)	**Railway Crossing**																												
	Number of tracks		All	-	-				-	-	All			All		-	Most	-					-	Some		Most	-		-
	Control type		All	-	-				-	-	All			All		-	All	-					-	All		All	-		-
	Safety equipment		All	-	-				-	-	All			All		-	Most	-					-	Most		All	-		-
6d)	**Bridges**																												
	Approach and bridge seal widths		Most	-	-				-	-	All			All		All	Some	-					-	All		Most	-		-
	Length		All	-	-				-	-	All			All		All	Most	-					-	All		Most	-		-
	Abutment and rail protection		No	-	-				-	-	Some			Most		Most	Most	-					-	Some		Some	-		-

29)Buller District Council	30)Transit NZ (Dunedin)	31)Central Otago District Council	32)Montgomery Watson (Dunedin)	33)Western BOP District Council	34)Canterbury Regional Council	35)Waimate District Council	36)Gore District Council	37)Christchurch City Council	38)Montgomery Watson (Christchurch)	39)Transit NZ (Auckland)	40)Transit NZ (Wellington)	41)Rodney District Council	42)Auckland City Council	43)Dunedin City Council	44)Transit NZ (Christchurch)	45)Napier City Council	46)Wellington City Council	47)New Plymouth District Council	48)Rotorua District Council	49)Transit NZ (Wanganui)	50)North Shore City Council	51)Palmerston North City Council	52)Porirua City Council	53)Traffic Design Group	54)Waitakere City Council	55)Marlborough District Council	56)Transit NZ (Napier)	57)Hurunui District Council	58)McKenzie District Council	TOTALS
0	1		1		1	1		1	1	1	1	1		1	1	1	1		1	1		1	1	1	1		1	0	0	32
-	1		1		-	-		-	-	1	1	-		-	-	1	-		-	-		-	-	-	-		1	-	-	7
-	1		1		-	-		1	1	1	1	1		1	-	1	1		1	-		1	-	-	1		1	-	-	20
-	-		-		-	-		1	-	-	1	1		1	-	1	1		-	-		1	-	-	1		-	-	-	13
-	-		-		-	-		1	-	-	-	1		-	-	1	-		-	-		1	-	-	-		-	-	-	8
-	1		1		-	-		-	1	1	1	1		-	-	1	-		1	1		-	-	-	-		1	-	-	15
-	-		-		-	1		-	1	-	-	1		-	-	-	-		-	-		-	-	-	-		-	-	-	9
-	-		-		1	-		-	-	-	-	-		-	-	-	-		-	-		-	-	-	-		-	-	-	3
-	1		-		-	-		-	-	1	-	-		-	-	1	-		-	-		-	-	-	-		-	-	-	3
-	1		1		-	-		1	1	1	1	1		1	1	1	-		1	1		1	-	1	1		1	-	-	22
-	1		1		-	-		1	1	1	1	1		1	1	1	-		1	-		1	1	1	1		1	-	-	25
-	1		1		-	-		1	1	-	1	-		1	-	1	-		-	1		1	-	1	1		1	-	-	18
-	1		1		-	-		1	1	-	1	-		1	-	-	-		-	1		-	-	-	-		1	-	-	12
-	-		-		-	-		-	1	-	1	-		1	-	-	-		-	1		-	-	-	-		-	-	-	6
-	1		-		-	-		-	-	-	-	-		-	-	-	-		-	-		-	-	-	-		-	-	-	1
-	-		-		-	-		-	-	-	-	-		-	-	-	-		-	-		-	-	-	-		-	-	-	2
0	1		1		0	1		1	0	1	1	1		1	1	1	1		1	1		1	1	1	1		1	1	1	34
0	1		1		0	1		1	1	1	1	1		1	1	1	1		1	0		1	1	1	1		1	0	0	27
-	1		1		-	1		1	1	1	1	1		1	1	1	1		1	1		1	1	1	1		1	1	1	34
-	-		?		-	-		?	-	25	-	-		2	-	2	-		?	2		-	-	-	-		-	-	-	31
-	?		?		-	-		?	-	-	-	22		31	-	5	11		?	-		39	-	-	10		4	-	-	365
-	-		?		-	-		?	-	-	-	51		24	-	5	20		?	-		50	-	-	10		-	-	-	547
-	-		?		-	-		?	-	-	-	73		10	-	40	-		?	-		81	-	-	20		-	-	2	717
-	?		?		-	-		?	-	70	75	96		-	-	40	-		?	100		25	-	-	-		23	-	-	824
-	-		?		-	100		?	-	-	-	255		-	-	70	-		?	-		10	-	-	10		-	100	66	1407
-	-		?		-	-		?	-	-	-	-		10	-	1	-		?	-		-	-	-	-		-	-	-	11
-	-		-		-	-		?	3	30	-	-		-	-	1	-		-	-		-	-	-	-		-	-	-	34
-	?		-		-	-		?	1	-	5	-		21	1	4	10		1	-		25	-	6	8		-	-	-	142
-	-		-		-	-		?	1	-	5	-		1	1	4	-		2	-		18	2	2	7		-	-	-	49
-	?		10		-	-		?	-	-	5	-		20	-	2	3		-	-		23	-	13	3		-	-	-	116
-	-		-		-	-		?	-	-	-	-		-	-	1	2		-	-		65	-	28	-		-	-	-	98
-	-		-		-	-		?	-	-	-	-		-	-	-	-		-	-		15	-	-	-		-	-	-	17
-	-		-		-	-		?	-	-	-	-		-	-	-	-		-	-		15	-	-	-		-	-	-	17
-	?		3		-	-		?	1	-	5	-		-	-	-	-		-	-		40	-	-	-		-	-	-	49
-			10		-	-		?	-	-	-	-		-	-	-	-		-	-		-	-	-	-		-	-	-	10
-	-		-		-	-		?	-	-	-	-		-	-	-	-		-	-		10	-	-	1		-	-	-	16
-	?		-		-	-		?	-	-	-	-		-	-	-	-		-	-		14	-	-	-		-	-	-	14
-	Most		-		-	Some		Most	-	No	All	Some		Most	-	Some	Some		All	All		Most	No	-	Most		Some	No	All	
-	Most		-		-	Some		Most	-	No	Most	Most		Most	-	Some	Some		Most	All		Most	No	-	Most		All	Most	All	
-	Some		Most		-	-		Most	All	-	Most	No		Some	Most	Some	Most		All	-		Some	All	?	Most		Some	No	-	
-	-		Most		-	-		All	All	-	Most	No		Some	All	Some	Some		All	-		All	Most	?	Most		Most	No	-	
-	Some		Most		-	-		All	All	-	Most	No		All	Some	Most	Most		All	-		Most	-	?	Most		Some	No	-	
-	Some		Most		-	-		Most	Some	-	Most	No		Most	Most	Some	Some		All	-		Most	Most	?	Most		All	No	-	
-	-		-		-	-		All	-	-	Most	No		Most	-	Most	-		All	All		Some	No	-	Most		Some	No	-	
-	-		-		-	-		All	-	-	Most	No		Most	-	Most	-		All	All		Most	No	-	Most		Some	No	-	
-	-		-		-	-		Most	-	-	-	No		Most	-	Some	-		Most	All		-	No	-	Most		All	No	-	
-	Most		-		-	-		Most	-	-	Most	No		Some	-	Some	-		Most	All		Some	No	-	-		All	No	-	
-	Most		-		-	-		All	-	-	All	No		Some	-	Most	-		Most	All		Most	No	-	-		All	No	-	
-	Most		-		-	-		Most	-	-	Some	No		No	-	Some	-		Most	-		Some	No	-	-		Some	No	-	

Table A6.1 Signalised Cross-road Accident Prediction Equations

Accident Type	LTSA Codes	Equation (accidents per approach)
Crossing (No Turns)	HA	$A = b_0 * q_2^{b1} * q_{11}^{b2}$
Right Turn Against	LB	$A = b_0 * q_2^{b1} * q_7^{b2}$
Rear-end	FA to FE	$A = b_0 * Q_e^{b1}$
Loss-of –control	C & D	$A = b_0 * Q_e^{b1}$
Others		$A = b_0 * Q_e^{b1}$

Figure A6.2 - Conflicting and Approach Flow Types (Cross-roads)

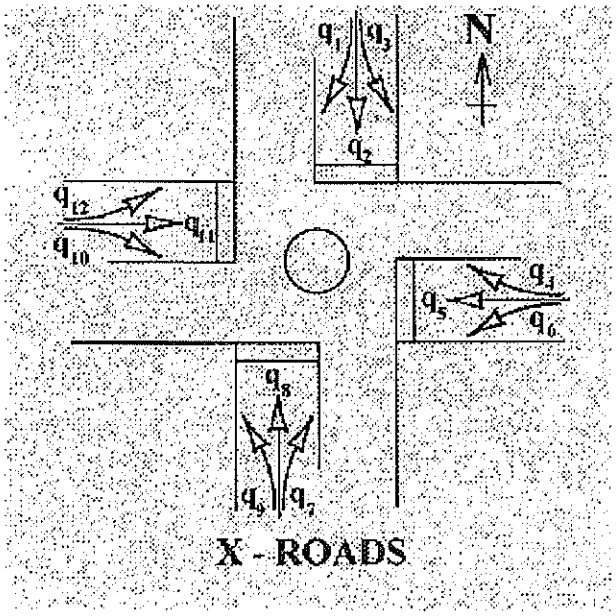


Table A6.2 Signalised Cross-roads – Prediction Model Parameters

Accident Type	LTSA Codes	b_0	b_1	b_2
Crossing (No	HA	$2.00E^{-4}$	0.34	0.37
Right Turn Against	LB	$9.70E^{-5}$	0.49	0.41
Rear-end	FA to FD	$1.70E^{-6}$	1.07	-
Loss-of –control	C & D	$3.12E^{-6}$	0.94	-
Others		$1.22E^{-3}$	0.46	-

A6.4.4.2 Roundabouts

The accident rates at 4-arm roundabouts are predicted by accident type and approach using the equations in Table A6.3, and the parameters in Table A6.4. The circulating flow (Q_c) is the traffic that the entering flow (Q_e) at each roundabout approach must give-way to.

Table A6.3 Roundabout Accident Prediction Equations

Accident Type	LTSA Codes	Equation (accidents per approach)
Entering vs Circulating	HA, LB, JA,	$A = b_0 * Q_e^{b1} * Q_c^{b2}$
Rear-end	FA to FD	$A = b_0 * Q_e^{b1}$
Loss-of –control	C & D	$A = b_0 * Q_e^{b1}$
Others		$A = b_0 * Q_e^{b1}$

Table A6.4 Roundabouts – Prediction Model Parameters

Accident Type	LTSA Codes	b_0	b_1	b_2
Entering vs Circulating	HA, LB, JA, MB, KA & KB	$8.92E^{-5}$	0.42	0.45
Rear-end	FA to FD	$5.76E^{-7}$	1.19	-
Loss-of -control	C & D	$3.02E^{-4}$	0.55	-
Others		$2.28E^{-3}$	0.26	-

A6.4.4.3 Priority Cross-roads (Give-way and Stop Control)

The accident rates at priority cross-roads are predicted by accident type and approach using the equations in Table A6.5 and the parameters in Table A6.6. At priority cross-roads the straight through flows are differentiated into those with priority (q_p) and those which have to give-way ($q_{g/w}$), or stop. For the crossing (no turns) accidents both the q_2 and q_{11} flows are used as predictors, but their order in the equation depends on their priority.

Table A6.5 Priority Cross-road Accident Prediction Equations

Accident Type	LTSA Codes	Equation (accidents per approach)
Crossing (No Turns)	HA	$A = b_0 * q_{p/w}^{b1} * q_p^{b2}$
Right Turn Against	LB	$A = b_0 * q_2^{b1} * q_7^{b2}$
Crossing (Vehicle Turning)	JA	$A = b_0 * q_2^{b1} * q_4^{b2}$
Loss-of –control	C & D	$A = b_0 * Q_e^{b1}$
Others		$A = b_0 * Q_e^{b1}$

Table A6.6 Priority Cross-roads – Prediction Model Parameters

Accident Type	LTSA Codes	b_0	b_1	b_2
Crossing (No Turns)	HA	$3.90E^{-4}$	0.38	0.37
Right Turn Against	LB	$7.50E^{-4}$	0.05	0.53
Crossing (Vehicle Turning)	JA	$1.08E^{-7}$	1.13	0.44
Loss-of –control	C & D	$1.04E^{-3}$	0.30	-
Others		$3.74E^{-4}$	0.57	-

A6.4.4.4 Signalised T-junctions

The accident rates at signalised T-junctions are predicted by accident type and approach using the equations in Table A6.7 and the parameters in Table A6.8. Figure A6.3 illustrates the different conflicting and approach flows at T-junctions.

Table A6.7 Signalised T-junction Accident Prediction Equations

Accident Type	LTSA Codes	Equation (accidents per approach)
Right Turn Against	LB	$A = b_0 * q_5^{b1} * q_3^{b2}$
Rear-end	FA to FD	$A = b_0 * Q_e^{b1}$
Crossing (Vehicles Turning)	JA	$A = b_0 * q_5^{b1} * q_1^{b2}$
Loss-of –control	C & D	$A = b_0 * Q_e^{b1}$
Others		$A = b_0 * Q_e^{b1}$

Figure A6.3 - Conflicting and Approach Flow Types (T-junction)

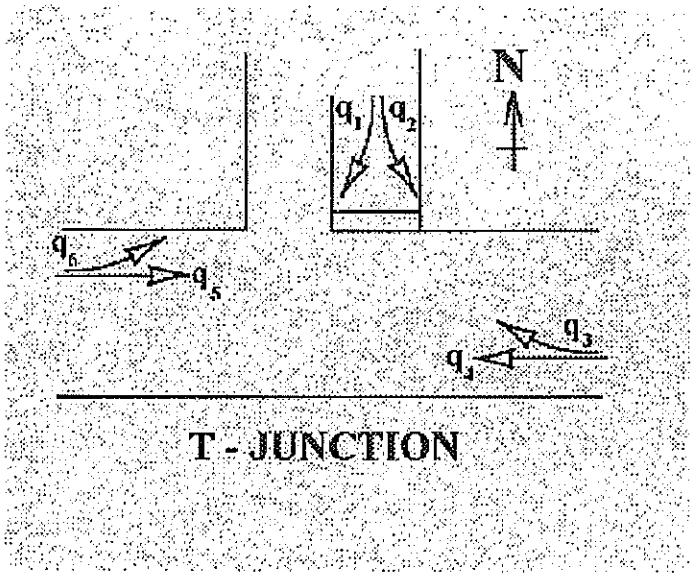


Table A6.8 Signalised T-junctions – Prediction Model Parameters

Accident Type	LTSA Codes	b_0	b_1	b_2
Right Turn Against	LB	0.117	-0.43	0.60
Rear-end	FA to FD	$7.66E^{-8}$	1.45	-
Crossing (Vehicles Turning)	JA	$3.22E^{-2}$	-0.34	0.51
Loss-of –control	C & D	$1.87E^{-3}$	0.17	-
Others		$1.69E^{-2}$	0.15	-

A6.4.4.5 Priority T-junctions

The accident rates at priority T-junctions are predicted by accident type and approach using the equations in Table A6.9 and the parameters in Table A6.10.

Table A6.9 Priority T-junction Accident Prediction Equations

Accident Type	LTSA Codes	Equation (accidents per approach)
Right Turn Against	LB	$A = b_0 * q_5^{b1} * q_3^{b2}$
Rear-end	FA to FD	$A = b_0 * Q_e^{b1}$
Crossing (Vehicles Turning)	JA	$A = b_0 * q_5^{b1} * q_1^{b2}$
Loss-of –control	C & D	$A = b_0 * Q_e^{b1}$
Others		$A = b_0 * Q_e^{b1}$

Table A6.10 - Priority T-junctions Prediction Model Parameters

Accident Type	LTSA Codes	b_0	b_1	b_2
Right Turn Against	LB	$6.66E^{-7}$	0.48	0.42
Rear-end	FA to FD	$2.90E^{-7}$	1.18	-
Crossing (Vehicles Turning)	JA	$7.20E^{-6}$	0.93	0.22
Loss-of–control	C & D	$1.64E^{-3}$	0.30	-
Others		$4.98E^{-4}$	0.51	-

A6.4.4.6 Uncontrolled T-junctions

The accident rates at uncontrolled T-junctions (that is T-junctions that have no give-way, stop or signalised controls) are predicted by accident type and approach using the equations in Table A6.11 and the parameters in Table A6.12. At uncontrolled T-junctions some accident types depend on the volume of three turning movements. This is thought to be due to the confusion caused by the left-turn give-way rule.

Table A6.11 Uncontrolled T-junction Accident Prediction Equations

Accident Type	LTSA Codes	Equation (accidents per approach)
Right Turn Against	LB	$A = b_0 * q_5^{b1} * q_3^{b2}$
Rear-end	FA to FD	$A = b_0 * Q_e^{b1}$
Crossing (Vehicles Turning)	JA	$A = b_0 * q_5^{b1} * q_1^{b2}$
Loss-of–control	C & D	$A = b_0 * Q_e^{b1}$
Others		$A = b_0 * Q_e^{b1}$

Table A6.12 - Uncontrolled T-junctions Prediction Model Parameters

Accident Type	LTSA Codes	b_0	b_1	b_2
Right Turn Against	LB	$2.98E^{-4}$	0.31	0.42
Rear-end	FA to FD	$1.74E^{-8}$	1.50	-
Crossing (Vehicles Turning)	JA	$7.24E^{-5}$	0.22	0.81
Loss-of–control	C & D	$5.02E^{-4}$	0.31	-
Others		$1.25E^{-3}$	0.41	-

A6.4.4.7 Product-of-Link-Flow Models

The models in this section predict the accident rate at an intersection from the link (two-way) flows on each of the intersecting roads. These models should only be used

when turning movement counts are not available, or can not be predicted using transport models.

These models should not be used when the volume of traffic on opposite arms of an intersection differs by more than 25% of the higher flow. If the majority of traffic on a link turns left or right at a cross-roads intersection, so that the opposing arm has a lot less traffic, then this type of model is inappropriate. Where volumes on both approaches of a link are available then the two approach flows should be summed to calculate the link volume.

The total reported accident rate for each intersection types is determined using the equation:

$$A_T = b_0 * Q_{minor}^{\ b1} * Q_{major}^{\ b2}$$

where Q_{minor} is the lowest of the two-way link volumes for cross-roads, and the stem flow for T-junctions.

Table A6.13 Product-of-Link-Flow Models

Intersection Type	b_0	b_1	b_2
Signalised Cross-roads	$4.08E^{-3}$	0.14	0.45
4 –arm Roundabout	$3.62E^{-4}$	0.48	0.37
Priority Cross-roads	$1.42E^{-3}$	0.51	0.21
Signalised T-junctions	0.156	0.13	0.04
Priority T-junction	$7.40E^{-5}$	0.19	0.75
Uncontrolled T-junction	$2.88E^{-3}$	0.19	0.36

A6.4.5 Urban Mid-block Sections, 50 km/h and 70 km/h Speed Limit Areas

For urban arterial, collector and local mid-block accidents, average injury accident rates can be associated with speed limit and, roadside development. The accident types predicted for urban mid-blocks sections, and the model types, are given in Table A6.14. The flow variable used in all models is the two-way traffic volume per day (Q_T).

Table A6.14 Urban Mid-block Accident Prediction Equations

Accident Type	LTSA Codes	Equation (accidents per approach)
Rear-end (both straight)	FA to FF	$A = b_0 * Q_T^{\ b1}$
Rear-end (one turning right)	GC to GE	$A = b_0 * Q_T^{\ b1}$
Loss-of–control	C & D	$A = b_0 * Q_T^{\ b1}$
Manoeuvring & Hit Object	M & E	$A = b_0 * Q_T^{\ b1}$
Other		$A = b_0 * Q_T^{\ b1}$

The accident prediction parameters for the major accident types are given for arterials, collectors and local streets in Tables A6.15 to A6.17.

Table A6.15 – Urban Arterials, 50 and 60 kph Areas

	Commercial/Industrial		Residential	
Accident Type	b_0	b_1	b_0	b_1
Rear-end (both straight)	$1.39E^{-7}$	1.59	$1.21E^{-7}$	1.59
Rear-end (one turning right)	$6.42E^{-4}$	0.64	$4.50E^{-4}$	0.64
Loss-of–control	$8.14E^{-5}$	0.90	$1.18E^{-4}$	0.90
Manoeuvring & Hit Object	$5.84E^{-3}$	0.45	$4.26E^{-3}$	0.45
Other	$2.14E^{-6}$	1.34	$1.74E^{-6}$	1.34

Table A6.16 – Urban Collectors, 50 kph Areas (no median)

	Commercial/Industrial		Residential	
Accident Type	b_0	b_1	b_0	b_1
Rear-end (both straight)	$8.64E^{-9}$	1.96	$4.84E^{-9}$	1.96
Rear-end (one turning	$3.56E^{-4}$	0.70	$2.72E^{-4}$	0.70
Loss-of–control	$1.79E^{-2}$	0.25	$2.90E^{-2}$	0.25
Manoeuvring & Hit Object	$1.24E^{-4}$	0.98	$7.64E^{-5}$	0.98
Other	$6.20E^{-5}$	0.93	$8.92E^{-5}$	0.93

Table A6.17 – Urban Local Streets, 50 kph Areas (no median)

Accident Type	b_0	b_1
Rear-end (both straight)	$2.92E^{-5}$	1.13
Rear-end (one turning right)	$2.50E^{-4}$	0.90
Loss-of–control	$3.98E^{-3}$	0.61
Manoeuvring & Hit Object	$5.46E^{-5}$	1.12
Other	$1.71E^{-5}$	1.33

A6.4.6 Rural Intersections, 80 km/h and 100 km/h Speed Limit Areas

The typical reported injury accident rates (per year) for rural intersections are calculated by using the urban intersection prediction equations with 'rural' parameters

(Table A6.18). Where turning movement counts are available, the accident rate should be predicted by accident type and approach. The total accident rate can then be predicted by summing the predictions by accident type and approach.

Where only approach flows are available, then the total accident rate can be predicted using the parameters at the bottom of each table, and the 'product-of-link-flow' equation (Section 6.4.4.7).

Table A6.18 – Rural Priority & Uncontrolled T-junctions Prediction Model Parameters

Accident Type	LTSA Codes	b_0	b_1	b_2
Right Turn Against	LB	$2.42E^{-8}$	0.54	1.63
Crossing (Vehicles Turning)	JA	$3.96E^{-5}$	0.34	0.93
Turning verses same direction	G	$2.26E^{-4}$	0.58	-
Others		$1.25E^{-3}$	0.34	-
Total (Product-of-link-flows)	All	$2.46E^{-4}$	0.53	0.42

A6.4.7 Rural Mid-block Sections, 80 km/h and 100 km/h Speed Limit Areas

For rural highways (both Transit NZ and district) and local streets (all other 80 and 100 kph streets), the average injury accident rates can be associated with the terrain type (flat, rolling and mountainous). The accident types predicted for rural mid-blocks sections, and the model types, are given in Table A6.19. The flow variable used in all models is the two-way traffic volume per day (Q_T). In the head-on model it is assumed that the traffic vehicle split by direction is 50:50 over 24-hours.

Table A6.19 Urban Mid-block Accident Prediction Equations

Accident Type	LTSA Codes	Equation (accidents per approach)
Head-on	B	$A = b_0*((Q_T/2)^2)^{b1}$
Overtaking	A	$A = b_0*((Q_T/2)^2)^{b1}$
Rear-end (both straight)	FA to FF	$A = b_0*Q_T^{\ b1}$
Rear-end (one turning right)	GC to GE	$A = b_0*Q_T^{\ b1}$
Loss-of-control	C & D	$A = b_0*Q_T^{\ b1}$
Manoeuvring & Hit Object	M & E	$A = b_0*Q_T^{\ b1}$
Other		$A = b_0*Q_T^{\ b1}$

The accident prediction parameters for the major accident types are given for rural highways, and motorways/expressways in Tables A6.20 and A6.21.

Table A6.20 Rural Highway Accident Prediction Equations

	Level		Rolling	
Accident Type	b_0	b_1	b_0	b_1
Head-on	$1.91E^{-4}$	0.33	$2.48E^{-4}$	0.33
Overtaking	$1.03E^{-6}$	0.65	$3.16E^{-7}$	0.65
Rear-end (both straight)	$2.18E^{-8}$	1.72	$1.98E^{-9}$	1.72
Rear-end (one turning right)	$8.50E^{-5}$	0.78	$1.25E^{-5}$	0.78
Loss-of–control	$5.66E^{-3}$	0.48	$3.64E^{-3}$	0.48
Manoeuvring & Hit Object	$8.00E^{-4}$	0.52	$2.14E^{-4}$	0.52
Other	$3.36E^{-5}$	0.84	$2.44E^{-6}$	0.84

Table A6.21 –Motorways & Expressways

Accident Type	b_0	b_1
Rear-end	$1.18E^{-9}$	1.88
Loss-of–control	$5.62E^{-6}$	1.11
Overtaking	$2.50E^{-6}$	1.10
Other	$5.30E^{-3}$	0.41

A6.4.8 Curves in 100 km/h Speed Limit Areas

No changes required

Appendix D
Crash Codings

	TYPE	A	B	C	D	E	F	G	O
A	OVERTAKING AND LANE CHANGE	PULLING OUT OR CHANGING LANE TO RIGHT	HEAD ON	CUTTING IN OR CHANGING LANE TO LEFT	LOST CONTROL (OVERTAKING VEHICLE)	SIDE ROAD	LOST CONTROL (OVERTAKEN VEHICLE)	WEAVING IN HEAVY TRAFFIC	OTHER
B	HEAD ON	ON STRAIGHT	CUTTING CORNER	SWINGING WIDE	BOTH OR UNKNOWN	LOST CONTROL ON STRAIGHT	LOST CONTROL ON CURVE		OTHER
C	LOST CONTROL OR OFF ROAD (STRAIGHT ROADS)	OUT OF CONTROL ON ROADWAY	OFF ROADWAY TO LEFT	OFF ROADWAY TO RIGHT					OTHER
D	CORNERING	LOST CONTROL TURNING RIGHT	LOST CONTROL TURNING LEFT	MISSED INTERSECTION OR END OF ROAD					OTHER
E	COLLISION WITH OBSTRUCTION	PARKED VEHICLE	CRASH OR BROKEN DOWN	NON VEHICULAR OBSTRUCTIONS (INCLUDING ANIMALS)	WORKMANS VEHICLE	OPENING DOOR			OTHER
F	REAR END	SLOW VEHICLE	CROSS TRAFFIC	PEDESTRIAN	QUEUE	SIGNALS	OTHER		OTHER
G	TURNING VERSUS SAME DIRECTION	REAR OF LEFT TURNING VEHICLE	LEFT TURN SIDE SIDE SWIPE	STOPPED OR TURNING FROM LEFT SIDE	NEAR CENTRE LINE	OVERTAKING VEHICLE	TWO TURNING		OTHER
H	CROSSING (NO TURNS)	RIGHT ANGLE (70° TO 110°)							OTHER
J	CROSSING (VEHICLE TURNING)	RIGHT TURN RIGHT SIDE	OBSOLETE	TWO TURNING					OTHER
K	MERGING	LEFT TURN IN	RIGHT TURN IN	TWO TURNING					OTHER
L	RIGHT TURN AGAINST	STOPPED WAITING TO TURN	MAKING TURN						OTHER
M	MANOEUVRING	PARKING OR LEAVING	"U" TURN	"U" TURN	DRIVEWAY MANOEUVRE	PARKING OPPOSITE	ANGLE PARKING	REVERSING ALONG ROAD	OTHER
N	PEDESTRIANS CROSSING ROAD	LEFT SIDE	RIGHT SIDE	LEFT TURN LEFT SIDE	RIGHT TURN RIGHT SIDE	LEFT TURN RIGHT SIDE	RIGHT TURN LEFT SIDE	MANOEUVRING VEHICLE	OTHER
P	PEDESTRIANS OTHER	WALKING WITH TRAFFIC	WALKING FACING TRAFFIC	WALKING ON FOOTPATH	CHILD PLAYING (TRICYCLE)	ATTENDING TO VEHICLE	ENTERING OR LEAVING VEHICLE		OTHER
Q	MISCELLANEOUS	FELL WHILE BOARDING OR ALIGHTING	FELL FROM MOVING VEHICLE	TRAIN	PARKED VEHICLE RAN AWAY	EQUESTRIAN	FELL INSIDE VEHICLE	TRAILER OR LOAD	OTHER

Appendix E
Development of Confidence Interval Equations

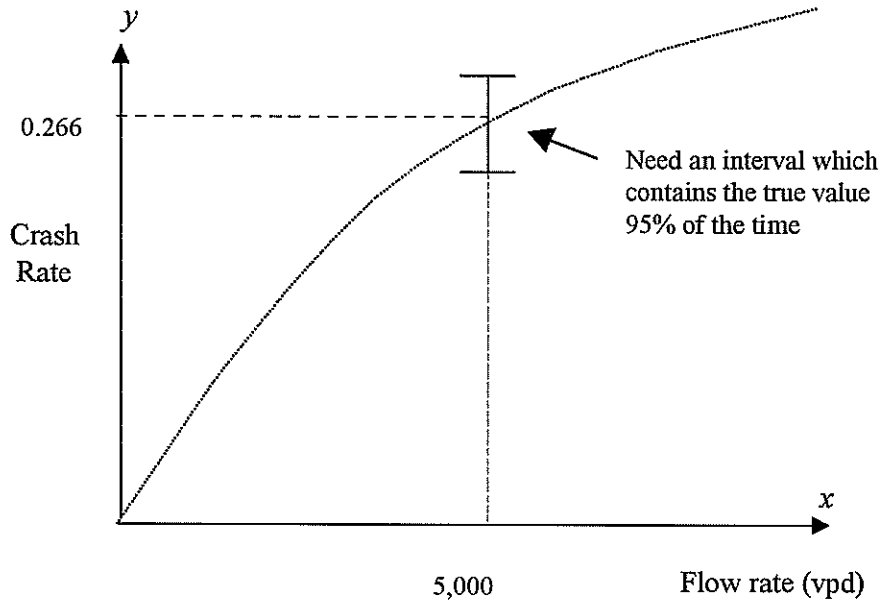
AIM: To develop an expression for the 95% confidence interval for the business hours mean accident rate, for a given vehicle flow rate, using Generalised Linear Models (GLMs).

EXAMPLE: The accompanying graphic shows the pattern of type one (rear end) injury accidents against vehicle flow rate for business hours (Hauer et al. 1989). The fitted curve is:

$$y = 1.2311 \times 10^{-5}\, x^{1.17176} (= b_o x^{b1})$$

where y is the average number of accidents from 1989 to 1991 and x is the vehicle flow rate.

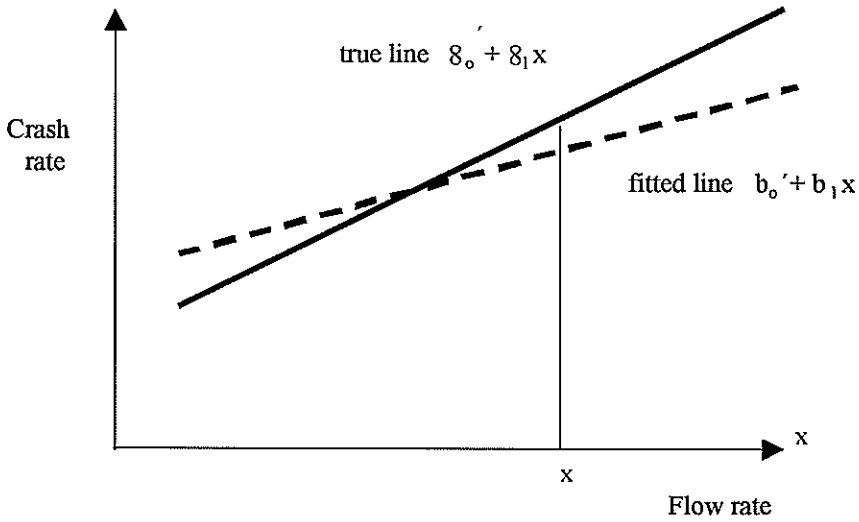
Assume that x = 5000 and y = 0.266. The aim is then to find a 95% confidence interval for the true mean value, $\exists_o x^{\exists 1}$, which this figure estimates.



SOLUTION USING SIMPLE LINEAR REGRESSION:

There are four steps:

1. The business hours mean at x, $\exists_o' + \exists_1 x$, need to be found ($\exists_o'$ is used here to be consistent with later notation).



2. We estimate $\exists_o' + \exists_1 x$ using $b_o + b_1 x$, which is found by using least squares (equivalent to maximum likelihood).

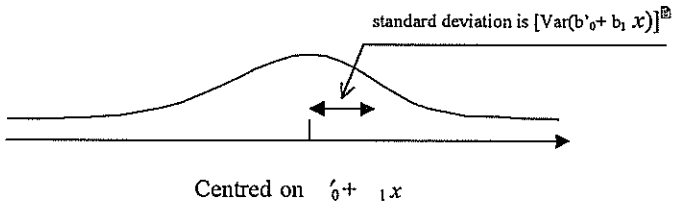
3. $$\begin{bmatrix} b_o' \\ b_1 \end{bmatrix} \sim \underline{N} \left[\begin{bmatrix} \beta_o' \\ \beta_1 \end{bmatrix}, \begin{bmatrix} \text{Var } b_o' & \text{Cov } b_o' b_1 \\ \text{Cov } b_o' b_1 & \text{Var } b_1 \end{bmatrix} \right]$$

A multivariate (in fact bivariance) normal distribution, with the covariance matrix known. The important conclusions from this are:

- $b_0' + b_1 x$ has a normal distribution
- $E[b_0' + b_1 x] = \exists_o' + \exists_1 x$, or $b_0' + b_1 x$ is an unbiased estimator of $\exists_0' + \exists_1 x$
- $\text{Var}(b_0' + b_1 x) = \text{Var } b_0' + 2x \text{ Cov } b_0' b_1 + x^2 \text{ Var } b_1$

So that the distribution of $b_0' + b_1 x$ is normal, as shown:

Distribution of $b_0' + b_1 x$



4. Thus:

$$P\left[\beta'_0 + \beta_1 x - 1.96\sqrt{Var(b'_0 + b_1 x)} \le b'_0 + b_1 x \le \beta'_0 + 1.96\sqrt{Var(b'_0 + b_1 x)}\right] = 0.95$$

from which, in the usual way, we have:

$$(b'_0 + b_1 x) \pm 1.96\sqrt{\mathrm{Var}(b'_0 + b_1 x)}$$

as a 95% confidence interval for $\beta_0' + \beta_1 x$ (for N, the simple size, being sufficiently large).

5. It turns out that:

$$\mathrm{Var}(b'_0 + b_1 x) \approx S^2 \left[\frac{1}{N} + \frac{(x - \bar{x})}{\sum_{i=1}^{N} (x_i - \bar{x})^2} \right]$$

and the construction of the confidence interval can be completed.

The important thing here is the sequence of:

parameter →	estimator →	distribution of estimator →	confidence interval
$\beta_0' + \beta_1 x$	$b_0' + b_1 x$	(normal)	for the estimator

We now follow the same pattern to find a confidence interval for the true mean in the Hauer GLMs. For the single flow model.

$$y = \beta_0 x^{\beta_1}$$

as used in the pattern one example. That the distribution for a given flow rate is negative binomial will be accommodated enroute!

CONFIDENCE INTERVALS FOR TRUE MEANS IN GLM CONTEXT

1. A confidence interval for $y = \beta_0 x^{\beta_1}$ is needed.

2. We use the method of scoring to find maximum likelihood estimates of β_0 and β_1, called b_0 and b_1 (this using the MINITAB macros in Turner (1995)). So $b_0 x^{b_1}$ is our estimator of $\beta_0 x^{\beta_1}$.

3. The next step is the important part. When we maximise the likelihood in our macros, we deal with the model in the form:

$$\begin{aligned} \log y &= \log b_0 + b_1 \log x \\ &= b'_0 + b_1 \log x \end{aligned} \quad \text{(hence the earlier notation)}$$

and estimate b'_0 and b_1. Standard maximum likelihood theory (see for example Dobson 1990), states:

$$\begin{bmatrix} b'_0 \\ b_1 \end{bmatrix} \overset{\text{asympt.}}{\sim} \underline{N}\left(\begin{bmatrix} \beta'_0 \\ \beta_1 \end{bmatrix}, I^{-1}\right)$$

or the covariance matrix is the inverse of the Fisher information matrix:

$$I = \begin{bmatrix} \dfrac{\partial^2 L}{\partial \beta^1{}_0{}^2} & \dfrac{\partial^2 L}{\partial \beta'_0 \, \partial \beta_1} \\ \dfrac{\partial^2 L}{\partial \beta'_0 \, \partial \beta_1} & \dfrac{\partial^2 L}{\partial \beta_1{}^2} \end{bmatrix}$$

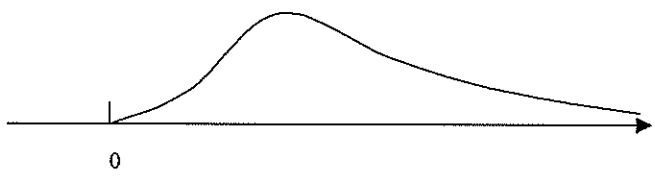
where L is the log - likelihood.

In summary:

(i) Linear combinations of normal random variables are normal, so since b'_0 and b_1 are normal,

$$\log y = b'_0 + b_1 \log x$$

is normal, or y itself (our interest being in y) is "log-normal", i.e. its logarithm has a normal distribution. Log-normal distributions look like:



which is good, since we'd expect this sort of distribution for $b_o x^{b1}$ (go back to the pattern one accidents for $x = 5{,}000$).

(ii) We need to find the covariance matrix I^{-1}. Fortunately this is approximated by $(X'WX)^{-1}$ (see p.65 of Turner (1995)), which is already computed in the MINITAB macros used in this research, (for example as m9 in TRIAL.MTB).

Returning to the main discussion:

4. A 95% confidence interval for $\beta_0' + \beta_1 \log x$ is thus:

$$(b'_0 + b_1 \log x) \pm 1.96\sqrt{\mathrm{Var}(\beta b_0 + \beta_1 \log x)}, \textit{ where}$$

$$\mathrm{Var}(b'_0 + b_1 \log x) = \mathrm{Var}\, b'_0 + 2 \log x \, Cov\, b'_0\, b_1 + (\log x)^2 \,\mathrm{Var}\, b_1$$

$$= (X'WX)^{-1}_{11} + 2\log x (X'WX)_{12} + (\log x)^2 (X'WX)^{-1}_{22}$$

where

$$(X'WX)^{-1} = \begin{bmatrix} (X'WX)^{-1}_{11} & (X'WX)^{-1}_{12} \\ (X'WX)^{-1}_{21} & (X'WX)^{-1}_{22} \end{bmatrix}$$

A caution - we have made two assumptions:

(i) $\begin{bmatrix} b'_0 \\ b_1 \end{bmatrix}$ is only *asymptotically* bivariate normal.

(ii) $(X'WX)^{-1} = I^{-1}$

So a larger sample size will help in each case.

5. Finally, we need a 95% confidence interval for y, not log y, so if we exponentiate the confidence bounds found so far, we will be finished, i.e. a 95% confidence interval for $\exists_0 x \exists^1$ is:

$$\exp\{b'_0 + b_1 \log x \pm 1.96\sqrt{\mathrm{Var}(b'_0 + b_1 x)}\}$$

COMPUTATIONAL SUMMARY (for the $\exists_0 x^{\exists 1}$ model)

1. Find b'_0 (= log b_0), b_1 and $(X'WX)^{-1}$, from the macro.

2. Then an approximate (improving with sample size) 95% confidence interval for $\exists_0 x^{\exists 1}$ is such that:

Lower Limit =

$$\exp\{b'_0 + b_1 x - 1.96\sqrt{(X'WX)^{-1}_{11} + 2\log x (X'\ WX)^{-1}_{12} + (\log x)^2 (X'WX)^{-1}_{22}}\}$$

Upper Limit =

$$\exp\{b'_0 + b_1 x + 1.96\sqrt{(X'WX)^{-1}_{11} + 2\log x (X'WX)^{-1}_{12} + (\log x)^2 (X'WX)^{-1}_{22}}\}$$

A WORKED EXAMPLE

For the pattern one, business hours example described at the start, TRIAL.MTB yields:

$$b'_0 = -11.035, b_0 = 1.2311 \times 10^{-5}, b_1 = 1.17176 \text{ and}$$

$$(X'WX)^{-1} = \begin{bmatrix} 3.54747 & -0.42210 \\ -0.42210 & 0.05047 \end{bmatrix}$$

For $x = 5{,}000$, $b_0 x^{b1} = 1.2311 \times 10^{-5} \times 5000^{1.17176} \approx 0.266$, while

$$\exp\{1.96\sqrt{8.54747 + 2(\log 5000) \times -0.42210 + (\log 5000)^2 \times 0.05047}\}$$

= 1.3052. Then

$$\exp\{-1.96\sqrt{8.54747 + 2(\log 5000) \times -0.42210 + (\log 5000)^2 \times 0.05047}\}\} = \frac{1}{1.3052} = 0.7662$$

So the 95% confidence interval is:

$$[0.7662 \times 0.266\ ,\ 1.3052 \times 0.266]$$

or $[0.204\ ,\ 0.347]$ injury accidents in the 1989 to 1991 period.

FINAL COMMENTS

(i) It was anticipated that the confidence interval width would increase as flow rates increase. This automatically will happen thanks to the "quadratic in log$x$" form of $\mathrm{Var}(b'_0 + b_1 x)$.

(ii) The validity of the interval improves as sample size increases.

(iii) The analysis for two-flow models will follow similarly, e.g. for $b_0\, x_1^{b1} x_2^{b2}$ use:

$$\mathrm{Var}\,(b'_0 + b_1 \log x_1 + b_2 \log x_2) = \mathrm{Var}\, b'_0 + (\log x_1)^2\, \mathrm{Var}\, b_1 + (\log x_2)^2\, \mathrm{Var}\, b_2$$
$$+ 2\,[(\log x_1)\,\mathrm{Cov}\, b'_0\, b_1 + (\log x_2)\,\mathrm{Cov}\, b'_0\, b_2 + (\log x_1)(\log x_2)\,\mathrm{Cov}\, b_1 b_2]$$

and pick these variances and covariances off the $(X'WX)^{-1}$ matrix.