Next generation of rural roads crash prediction models – pilot study

March 2011

S Turner and R Singh (Beca Infrastructure Ltd, Christchurch)
F Tate (MWH New Zealand Ltd, Wellington)

NZ Transport Agency research report 437

ISBN 978-0-478- 37170-3 (print)
ISBN 978-0-478- 37169-7 (electronic)
ISSN 1173-3756 (print)
ISSN 1173-3764 (electronic)

NZ Transport Agency
Private Bag 6995, Wellington 6141, New Zealand
Telephone 64 4 894 5400; facsimile 64 4 894 6100
research@nzta.govt.nz
www.nzta.govt.nz

Turner, S, F Tate and R Singh (2009) Next generation of rural roads crash prediction models – pilot study. *NZ Transport Agency research report 437*. 86pp.

Keywords: crash modelling, crash prediction, point hazards, roadside environment, skid resistance, rural two-lane roads

An important note for the reader

The NZ Transport Agency is a Crown entity established under the Land Transport Management Act 2003. The objective of the Agency is to undertake its functions in a way that contributes to an affordable, integrated, safe, responsive and sustainable land transport system. Each year, the NZ Transport Agency funds innovative and relevant research that contributes to this objective.

The views expressed in research reports are the outcomes of the independent research, and should not be regarded as being the opinion or responsibility of the NZ Transport Agency. The material contained in the reports should not be construed in any way as policy adopted by the NZ Transport Agency or indeed any agency of the NZ Government. The reports may, however, be used by NZ Government agencies as a reference in the development of policy.

While research reports are believed to be correct at the time of their preparation, the NZ Transport Agency and agents involved in their preparation and publication, do not accept any liability for use of the research. People using the research, whether directly or indirectly, should apply and rely on their own skill and judgement. They should not rely on the contents of the research reports in isolation from other sources of advice and information. If necessary, they should seek appropriate legal or other expert advice.

Acknowledgements

The research team would like to acknowledge Alistair Smith (the former project researcher) for his work managing the field data collection surveys that were conducted as part of this research, and for his assistance with initial crash model development.

We would also like to thank Professor Graham Wood from the Department of Statistics at Macquarie University for his valuable input and assistance with the development of a sample-selection procedure for this phase of the research.

Their help and advice has been much appreciated.

Abbreviations

AADT	Annual average daily traffic
AIC	Akaike Information Criterion
ARRB	Australian Road Research Board
BIC	Bayesian Information Criterion
CAS	Crash analysis system
RAMM	Road Assessment Maintenance Management Database
SCRIM	Sideways-force Coefficient Routine Investigation Machine
SH	State highway
TLA	Territorial local authority

Contents

Executive summary

The majority of fatal and serious crashes in New Zealand occur on rural two-lane roads. New Zealand has an extensive rural road network, most of which has relatively low volumes. Data on historic crash patterns is not always sufficient to enable a suitable diagnosis of the safety deficiencies of various sections of this rural road network. It can be difficult to readily identify safety issues on low-volume roads and shorter sections of highway, where the scarcity of crashes may mask the considerable potential for safety improvements.

Crash prediction models are useful tools for evaluating the crash risk of a section of road, and also for evaluating the benefits of any changes to a road section. The benefits of using these models are recognised internationally and many countries have developed comprehensive crash prediction models of road stereotype models for these purposes.

The overall purpose of this research was to quantify the impact of all key road features on the safety of two-lane rural roads, and understand/quantify any interaction between these variables. The research described in this report covers the second phase of a three-phase approach, the three phases being:

1. a **scoping study**, aimed at identifying the salient variable and data collection and initial sampling requirements (this stage was completed in 2006, with a 'Scoping report' produced in July 2006
2. a **pilot study** to test alternative collection methods and to refine the sample size and budget requirements (the focus of this report)
3. the **main study**, which will build the necessary models.

The research documented in this report, which was carried out over the period 2007–2010, covers stage 2 of the project. The key objective was to assess different methods of collecting data and to build preliminary crash prediction models that would identify the key variables required for the final models that will be developed during the 'main study' stage.

Part of this 'pilot' stage involved data collection and integrated modelling of data on a large number of variable types. These variables were broadly classified into categories such as traffic volume, road geometry, lane width, shoulder environment, roadside hazards, road pavement condition and accesses. Data was collected on state highways, for which road alignment information was available in an electronic format. Two hundred sections, each of them 400m in length, were randomly sampled from all two-lane sections of state highway in the Waikato region (Transit Region 3). Twenty-nine of these were surveyed in December 2007, and the rest in February 2008.

Further checks conducted on the data collected for these 200 sections resulted in the exclusion of certain sections because of inaccuracies in the raw roadside, seal-width and crash data, and/or insufficient skid-resistance data. This resulted in a final set of 148 sections that was accepted for input into the models.

Crash data for the surveyed sections, all rural state highways in the Waikato region, and all rural state highways in New Zealand, was extracted from the national Crash Analysis System (CAS) database for the years 1996–2007. The proportion of fatal-injury crashes in the sample set was slightly higher, at 6.6%, than that for Waikato state highways (4.3%) and for all New Zealand state highways (3.8%). The proportions of 'bend – lost control'/'head-on', and 'rear-end'/'obstruction' type crashes were found to be marginally higher and lower, respectively, than those on all Waikato state highways. The proportion of crashes caused by speeding drivers was also lower in the sample set than the average for Waikato and the whole of New

Zealand. The other crash types and factors showed proportions similar to those observed for all state highways in Waikato and New Zealand.

Analysis of the distributions of certain key variables in the sample set showed that a large proportion (87%) of the sites had an AADT of less than 5000 vehicles per day, while the range of values of AADT in the sample set was 295–11,500 vehicles/day. In terms of the shoulder environment, the majority of sites (99%) had an unsealed shoulder width of less than 1m, and 72% of sites did not have any unsealed shoulder. All sites were observed to have a recoverable slope width of less than 4m. The traversable slope width varied across the sites and was found to be between 0 and 9m.

While some point hazards such as wooden poles, culverts, concrete poles and trees were present in significant numbers in the sections included in the sample set, others such as light columns, heavy street poles and sign supports were not commonly present on the sections surveyed. Overall, 38% of the sites had no point hazards, while another 38% had between 0 and 1 point hazard within an average 100m section.

Accesses were not found to be an important issue for a large number of state highway sections. No accesses were present on 44% of the sections in the sample set. The low frequency of accesses suggests that for the purpose of the main study, it may be viable to use a figure of 0.5 accesses per 100m for all sections. Data collection on access densities would need to be done only for areas that are known to have a higher density of accessways, and ideally also for remote areas that have a very low density of accesses, such as the state highway network in the West Coast region.

Analysis was also conducted to assess the variation in the length of straight and curved elements within 20 of the selected 400m sections. It was found that more than half of the sections analysed consisted of elements that had a length of more than 200m. Given the apparent variation in geometric element length within individual sections, it is suggested that during the main study, the state highway network should be split into element lengths, rather than into homogeneous sections, for data collection and analysis purposes.

The predictor variables included in the modelling were checked for correlation to identify pairs of variables that may have a cross-relationship with each other and therefore should not be included together in the models. The following pairs of variables were observed to be significantly correlated:

- **Combined point hazards and traversable slope width/distance to non-traversable slope:** This supports the suitability of using a combined roadside environment rating, such as the KiwiRAP roadside hazard rating, as opposed to looking at the impact of each element of the roadside environment separately.
- **Traffic volume (AADT) and seal width:** Both of these variables are recognised to be important from a crash prediction perspective. During the main study, a suitable approach to overcome this correlation may be to define fixed AADT bands and build individual crash models for each of these bands.
- **Accesses and number of culverts – roadside:** This suggests that accesses on rural roads are often accompanied by drainage culverts. It thus seems reasonable to discard ‘culvert – roadside’ as one of the predictor variables.
- **Combined point hazards and combined accesses were both correlated to the individual categories of point hazards and accesses respectively:** This supports the use of the ‘combined’ variables to describe point hazards and accesses.

- **Recoverable slope width and traversable slope width:** In the main study, a suitable measure to overcome this correlation may involve modifying the definition of traversable slope to refer only to the slope width starting from the edge of the recoverable slope, instead of from the edge of the seal.

The modelling methodology adopted in this study was in accordance with the approach made in many previous studies, and involved fitting generalised regression curves to independent and dependent data. Traffic volume (AADT) was included as a default variable in all the crash models. Starting with the single-variable volume-only model, additional variables were added to assess the best-performing sets of variables. From the initial list of 28 variables for which data (both manual and electronic) was collected, a most-representative and best-performing set of eight was selected to be incorporated into the multi-variable models alongside traffic volume, based on the results of the two-variable models. The variables that were found to perform the best are listed in the table below.

Table Variables in multi-variable models

Type	Variable
Traffic volume	Traffic volume (V)
Shoulder environment	Unsealed shoulder (U)
	Seal width (S)
Point hazards	Combined point hazards (H)
Accesses	Combined accesses (Ca)
Roadside – other	Distance to non-traversable slope/perpendicular deep drain (N)
Road geometry	Absolute curvature (C)
	Absolute gradient (G)
	Skid resistance (Sr)
	% reduction in curve speed (Vc) [1]

1 The results of the two-variable models indicated that horizontal geometry was not a significant variable. However, horizontal geometry is widely considered to have a significant effect on crash rate. A horizontal alignment variable (percentage reduction in curve-negotiation speed of the section compared with the preceding 500m section) was thus later introduced into the best-performing four-, five- and six-variable models.

The overall preferred model was found to involve volume, distance to non-traversable hazard, absolute gradient, SCRIM coefficient, and percentage reduction in curve speed. The preferred model had the following form:

$$A = 2.2E^{-04}\ V^{0.719} \times N^{0.078} \times G^{-0.26} \times Sr^{-2.569} \times Vc^{0.219}$$

where:

A is the predicted number of crashes in five years for a 100m section of rural road

V is the two-way AADT for the road section

N is the distance in metres to the non-traversable hazard (eg row of trees or deep ditch), multiplied by 1000.

G is the absolute gradient

Sr is the average value of SCRIM for the road section

Vc is the percentage reduction in the curve-negotiation speed of the section as compared with the preceding 500m section.

The model results were evidence of the important influence that the roadside environment and road geometry have on predicted crash rate. However, the exponents (powers of model variables) with which most secondary variables (except SCRIM coefficient) contributed to the model were small, indicating a subtle influence only.

No doubt many of the results were sensitive to the small size of the data set. As the pilot aspect of the wider study, this research has identified the key variables and data collection issues that will be critical to progressing the research project in stage 3 (main study). Because of the small sample set, the models developed as part of this stage of the research **should not** be used for crash prediction.

This study has built upon the main recommendations of the 'Scoping report' (from stage 1 of this research) and includes the key predictor variables and sampling methodology identified therein to build initial crash models. The key recommendations for the 'main study phase' of this research, based on the research outcomes of this pilot study, are as follows:

- It may be beneficial to undertake data collection and analysis for homogeneous elements of varying lengths on the state highway network, instead of on fixed section lengths. However, because of inaccuracies in crash location data, both homogeneous road lengths and road elements should be considered during the main study.
- The low density of accesses on a large proportion of sections in the sample set suggests that it may be viable to use a generic figure of 0.5 accesses per 100m for a majority of road sections. Data collection may only be required for state highways in areas that are known to have a high or very low density of accessways.
- There is ample support for the use of the KiwiRAP roadside hazard rating to estimate the quality of the roadside environment. The use of this rating will eliminate correlations between predictor variables as mentioned earlier, while at the same time providing a reasonably detailed and large sample set of data for building the models.
- It may be necessary to build separate crash models for individual AADT bands, to eliminate the correlation between AADT and seal width. Such a model form will also help to better estimate the safety benefits of seal widening.

Abstract

The majority of fatal and serious crashes in New Zealand occur on rural two-lane roads. Data on historic crash patterns is not always sufficient to enable a suitable diagnosis of the safety deficiencies of various sections of this rural road network. It also cannot readily identify safety issues on low-volume roads and shorter sections of highway, where the relative scarcity of crashes may mask the considerable potential for safety improvements.

This pilot study covers the second stage of a three-stage research project that aims to quantify the impact of all key road features on the safety of two-lane rural roads. This stage of the study involved the collection of road alignment, roadside environment, traffic flow, and crash data for 200 sections of rural road, each one 400m long, throughout the Waikato region of New Zealand. The data was used to develop preliminary crash prediction models for two-lane rural roads, using generalised linear regression model techniques developed by Beca.

The data collection exercise covered a total of 28 predictor variables used for developing the preliminary model. The preferred model showed that the crash rate was most influenced by five predictor variables – namely, traffic volume, absolute gradient, distance to non-traversable hazards, skid resistance (SCRIM), and percentage reduction in the curve-negotiation speed of the section as compared with the preceding 500m section.

1 Introduction

1.1 Background

The majority of fatal and serious crashes in New Zealand occur on rural two-lane roads. New Zealand has an extensive rural road network, most of which has relatively low volumes. Data on historic crash patterns is not always sufficient to enable a suitable diagnosis of the safety deficiencies of various sections of this rural road network. It can be difficult to readily identify safety issues on low-volume roads and shorter sections of highway, where the scarcity of crashes may mask the considerable safety improvement potential.

Crash prediction models are useful tools for evaluating the crash risk of a section of road, and also for evaluating the benefits of any changes to a road section. The benefits of using these models are recognised internationally and many countries have developed comprehensive crash prediction models of road stereotype models for these purposes.

Crash prediction models have been increasingly used to identify and evaluate road safety issues since the mid-1990s, when initial forms of crash prediction models were first developed for the majority of intersection and link types in urban and rural areas.

A number of recent studies have developed more detailed crash prediction models, typically for one-off evaluations of specific features or policies. These include:

- McLarin et al (1993), who investigated the impact of terrain characteristics (eg flat, rolling and mountainous) on crash rates
- Chadfield (1993), who looked at the impacts of lane and shoulder width and shoulder slope on crash rates
- Jackett (1992), who looked at crash rates and curve radii
- Koorey and Tate (1997), who looked at how alignment consistency and speed limit impacted on crash rate and severity
- Turner (2000), who investigated the relationship between traffic volume and crash rates
- Cenek et al (2004), who looked at the impact of road surface friction on crash rates
- Turner et al (2004), who looked at the crash rate implications of roadside hazards
- further research by Cenek and Davies (2008), which looked at the safety benefits of engineering measures such as high-friction surfacing and realignment.

While each of the above studies have, in general, answered the questions at hand, the individual models only contain a small subset (typically three to five) of the important predictor variables, and some are based only on traffic volume. As a result, the models cannot be readily or reliably used to evaluate various roading improvement options (eg seal widening compared with hazard removal or realignment).

1.2 Purpose of this research

The overall purpose of this research was to quantify the impact of all key road features on the safety of two-lane rural roads. The research described in this report covers the second phase of a three-phase approach, the three phases being:

1. a **scoping study**, aimed at identifying the salient variable and data collection and initial sampling requirements (this stage was completed in 2006, with a 'Scoping report' produced in July 2006)
2. a **pilot study** to test alternative collection methods and to refine the sample size and budget requirements (the focus of this report)
3. the **main study**, which will build the necessary models.

The objective of the wider study, which is being undertaken by Beca Infrastructure Ltd for the NZ Transport Agency, is to develop a series of crash prediction models (similar to those developed internationally) that can be used to identify and evaluate engineering-related safety issues on rural roads, and will be suitable for a wide number of central government and road controlling authority applications.

1.3 Pilot study objectives

The research documented in this report covers stage 2 of the wider project, with a focus on the coordinated collection and integrated modelling of data on a large number of variable types – traffic volume, road geometry, lane width, shoulder environment, roadside hazards, road pavement condition and accesses – all split into several individually quantifiable variables that can be incorporated separately or in combinations into generalised linear models for crash prediction rates.

The phase 2 research objectives were:

- to collect data for each of the key road features (specified in the 2006 'Scoping report') for 200x200m rural road sections (new data being collected manually through field surveys)
- to collect data for a sample of the rural road sections using video data-capturing techniques, and to determine whether this method is accurate enough to replace the manual field-survey collection approach
- to develop preliminary crash prediction models for rural roads, for the main crash types
- to estimate the sample size that would be required in the 'main study' to produce national and regional rural crash models.

Although the initial objective of this study was to collect data for 200 sections of road, each one 200m in length, the specification for section length was later revised and increased to 400m (see the discussion in section 2).

1.4 Structure of this report

This report consists of the following sections:

- Section 2 presents details of the site-selection procedure and identifies data collection requirements.

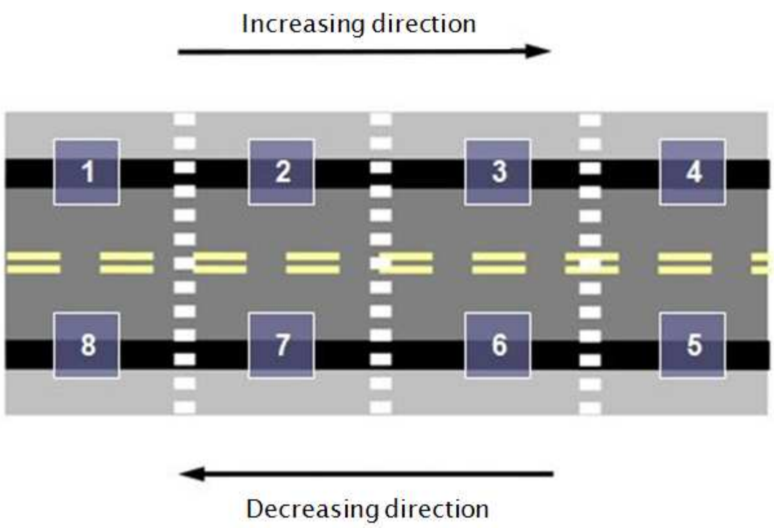
- Section 3 examines the distributions of various variables for which data was collected.
- Section 4 provides the results of a comparative analysis of crashes occurring on the selected sections, as compared with crashes on all state highways in the Waikato region and in the whole of New Zealand.
- Section 5 provides details on the development of predictor variables, and identifies correlated variables.
- Section 6 presents the outcomes of the collection of state highway video data.
- Section 7 details the modelling methodology that was adopted for this study.
- Section 8 provides the results from the modelling process and identifies the preferred preliminary model.
- Section 9 summarises the tasks undertaken during the pilot study and lists the key conclusions
- Section 10 presents recommendations for the main study.

2 Site selection

Data was collected on state highways, for which road alignment information was available in an electronic format. Two hundred sections, each 400m long, were randomly sampled from all two-lane sections of state highway in the Waikato region (Transit Region 3). Twenty-nine of these were surveyed in December 2007, and the rest in February 2008. The procedure that was used to select road sections is described in section 2.1 below.

The 400m road sections were segmented into eight subsections of 100m each, with four subsections on each side of the road, as depicted in the figure below. Data related to the road geometry and roadside environment was collected for each of the eight 100m subsections. This data for individual directional 100m subsections within each 400m section was later combined, and the final data for the development of the model was based on directional 400m sections.

Figure 2.1 Road section showing division into eight subsections



The original plan was to collect data from 200 sections, each of them 200m in length, which would later be added to 1000 road sections of the same length that would be collected during the main study (the third phase of this project). However, the length of these sections was increased to 400m to enable the research team to test whether it would be best to use homogeneous lengths (100m long) or element lengths (curves and straights of variable lengths) for the analysis and modelling to be undertaken during the main study. Section 2.2 reports the outcomes of an analysis to assess the variability in element lengths within the 400m sections.

2.1 Sampling procedure

The process for randomly sampling a 400m section from the entire list of state highway sections for which data was available consisted of the following steps:

1 Assign a probability of selection for each road section on the list that was proportional to its length – thus a longer section would have a higher probability of being selected.

2 Select a road section according to that distribution (of length L, say) and then select a position ‘x’ along that road uniformly on the interval from zero to (L-1).

3 Sample sections from position ‘x’ on the road for the next 400m.

4 Repeat steps 2 and 3, as above, until 200 lengths, each of them 400m in length, had been sampled, avoiding overlaps as necessary.

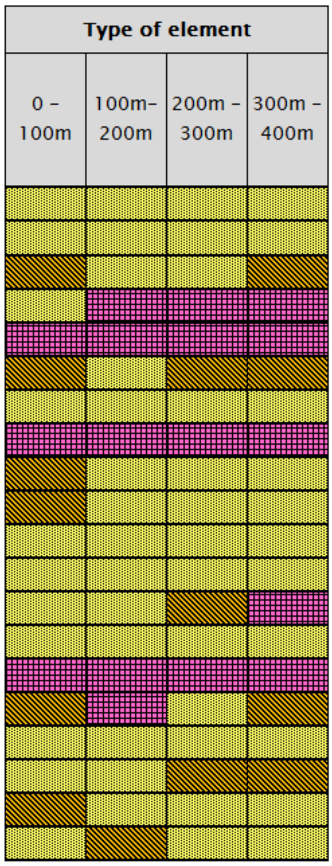
2.2 Variations in geometric element length

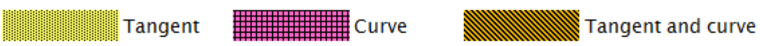
An analysis was undertaken to assess the variation in element length (curves/tangents) in the sections included in the sample set. Twenty sections were randomly selected from the list of 400m sections in the sample set. Geometry and alignment data, by element type (curves/tangents), for these sections was extracted from the RAMM database. Figure 2.2 shows the results of the analysis.

The lengths of individual curve and tangent elements within each of the selected sections were calculated from the RAMM data. These are shown in the table on the left in figure 2.2, in a sequential manner according to the occurrence of an element in the ‘increasing’ direction along the highway. The selected sections were also divided into 100m subsections and the type of road geometry element in each subsection was noted. These are shown in the grid in the right half of figure 2.2.

Figure 2.2 Sections according to element length

Section	Element length along increasing direction (m)				
	Geometry element in direction of travel →				
	Curve 1	Tangent 1	Curve 2	Tangent 2	Curve 3
1		400			
2		400			
3	43	350	7		
4		100	300		
5	400				
6	79	200	60	50	11
7		400			
8	400				
9		54	40	306	
10		20	80	300	
11		400			
12		400			
13		217	183		
14		400			
15	135		150		115
16		74	120	170	36
17		400			
18		235	90	75	
19		41	60	299	
20		132	268		





It was observed that seven of the 20 sections analysed were located on mostly straight sections of state highway, which was shown by the dominance of tangent elements for the said sections. Two of the sections were found to lie entirely along curved portions of the highway. The remaining sections were found to have a mix of tangent and curved elements, often in the form of a tangent segment transforming into a curve (and vice versa), or a tangent element located between two curved elements.

The above analysis indicates that more than half of the sections analysed consisted of straight or curved elements that had a length of more than 200m. Thus, increasing the section length from 200m to 400m was shown to be suitable in selecting sections with more 'complete' elements. However, there was also a case for using element lengths rather than 100m sections, given the presence of multiple elements and parts of different element types within individual 100m sections. This can also be seen in figure 2.2, where nine out of the 20 400m sections analysed had one or more individual 100m lengths that consisted of a combination of straight and curved elements.

2.3 Roadside hazard variables

Data on the following categories of roadside hazards was collected for each of the 100m sample sections:

- shoulder/roadside environment
- point hazards
- access densities.

Additional details on the variables included in each of the above categories are provided in sections 2.3.1–2.3.3 below.

Photographs were also taken for each of the 400m sample sections, showing the general roadside environment as well as the road alignment and cross section.

Figure 2.3 Example roadside environment

2.3.1 Shoulder/roadside environment

2.3.1.1 Seal width

Although seal width is known to affect the crash rate, the relationship between these two factors is complex and unclear. There is some evidence to suggest that a narrower seal width decreases the separation between opposing streams of vehicles, which may have a negative impact on the number of head-on crashes. A wider seal width can influence a driver's choice of speed, but may also help a driver who has lost control of their vehicle to regain control before leaving the roadway and hitting a hazard in the roadside environment.

Data was collected on the following seal width variables:

- seal width (m)
- unsealed shoulder width (m)

Figures 2.4 and 2.5 below depict a typical rural road cross section indicating the seal width, sealed shoulder and unsealed shoulder width.

Figure 2.4 Rural road cross section

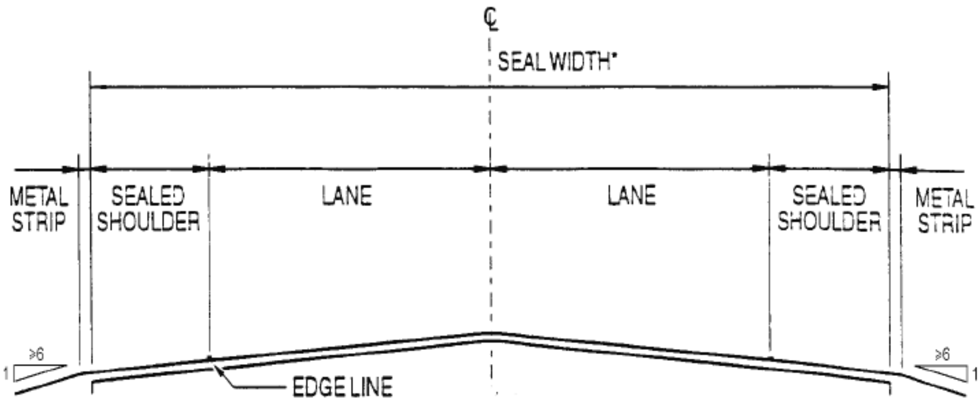


Figure 2.5 Seal width and unsealed shoulder width

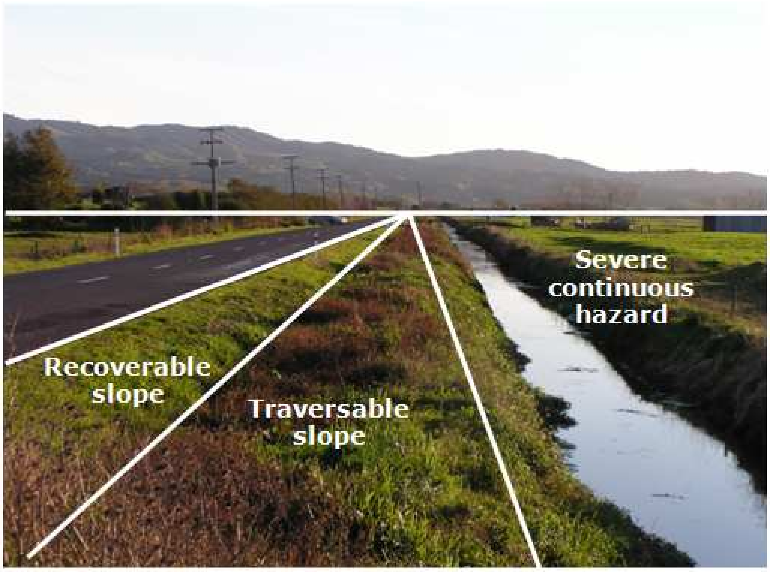
2.3.1.2 Slope width

The shoulder environment, and in particular the slope width, is important as it affects the ability of a driver who has lost control of the vehicle to regain some control and bring the vehicle to a safe stop, or re-enter the traffic lane.

Data on the following categories of slope width was collected for this research:

- recoverable slope width (m):
  - measured from the edge of seal where the slope is flatter than, or equal to, about 1:6 and clear of any continuous severe hazards – this gradient allows a motorist exiting the roadway to regain control of the vehicle and steer back to the roadway
- traversable slope width (m):
  - measured from the edge of seal where the slope is flatter than, or equal to, about 1:3 and clear of any continuous severe hazards – this gradient allows a motorist to slow or stop safely, but most likely not steer back to the roadway
  - often includes an area of recoverable slope.

Figure 2.6 Example of roadside shoulder environment showing recoverable slope and traversable slope



The width of the recoverable slope and the traversable slope can vary depending on the presence of continuous hazards in the roadside environment. Table 2.1 lists various kinds of continuous hazards and illustrates examples of each.

Table 2.1 Types of continuous hazards

Ditch		
Fence		
Continuous trees		

Cut-face/retaining bank		

2.3.2 Point hazards

Turner et al (2004) found a relationship between the density and type of roadside point hazards with the number and severity of loss-of-control crashes. When a vehicle leaves the roadway and moves into the shoulder environment, it may collide with a roadside hazard if one is present. If the hazard is close to the roadway, the likelihood of a crash increases and the severity of the crash will vary depending on the type of hazard.

For this study, data on a selection of roadside point hazards was collected. When assessing the shoulder environment, the traversable slope measurement represents the distance to severe continuous hazards – therefore only point hazards within this distance were collected (out to a maximum of 9m). A list of the point hazards collected in the field survey is provided below, along with the unit of measurement[1]:

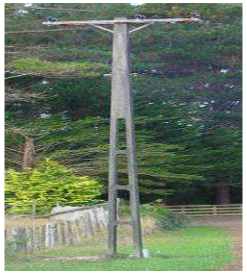
- wooden pole >200mm (no.)
- light column >300mm (no.)
- concrete pole – usually 'I' section (no.)
- heavy street pole >300mm without slip base (no.)
- sign support >120mm without slip base (no.)
- tree – trunk >100mm diameter (no.)
- culverts – roadside (no.)
- culverts – road with non-traversable headwall (no.)
- high-impact roadside furniture (eg transformers) (no.)
- distance to non-traversable slope/perpendicular deep drain (m)
- end concrete barrier/bridge parallel to road (m)
- concrete fence/barrier perpendicular to road (m).

1 This list, and the severity level of each hazard, was developed for this study by Turner and Tate (2009).

Figure 2.7 Example of roadside shoulder environment with point hazards

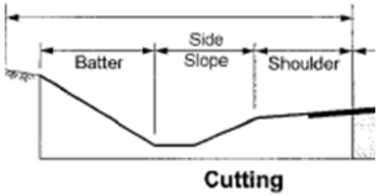
Table 2.2 illustrates examples of varying scales of severity for each of the point hazard categories listed above. The column corresponding to a ‘high’ severity level illustrates examples of the criteria used to identify point hazards during the data collection exercise.

Table 2.2 Severity levels of point hazards

Hazard	Severity level		
	Low	**Medium**	**High**
Wooden pole	 Post <100mm diameter	 Thin pole 100-200mm diameter	 Pole >200mm diameter
Concrete pole	 Vierendeel pole	 Concrete lighting columns	 Concrete service pole (I section)

<table>
<tr><th rowspan="2">Hazard</th><th colspan="3">Severity level</th></tr>
<tr><th>Low</th><th>Medium</th><th>High</th></tr>
<tr><td>Other poles</td><td colspan="2">
Slip base or frangible</td><td>
Steel, hollow, ground planted</td></tr>
<tr><td>Sign support</td><td>
<=100mm wood, <=60mm steel, all slip bases</td><td>
100-150mm wood, 60-120mm steel/aluminium (no slip base)</td><td>
>150mm wood, >120mm steel/aluminium (no slip base)</td></tr>
</table>

Hazard	Severity level		
	Low	**Medium**	**High**
Tree	 Small tree 50-100mm	 Medium tree 100-300mm	 Large tree >300mm
Culvert	 Culvert over shallow ditch		 Culvert over deep ditch or with headwall

Hazard	Severity level		
	Low	**Medium**	**High**
High impact roadside furniture	 Lightweight structures eg isolated rural letterboxes, hoardings	  Medium-weight structures eg telephone boxes, solid letterbox, rural bus shelter, cattle box	  High impact structures eg buildings, bus shelters, transformers
Non-traversable slope/ perpendicular deep drain	 Near-traversable (1:4) slopes	 Recovery of control not possible (0.5–1m deep)	 Deep and dangerous – >1m deep, >2m wide

Hazard	Severity level		
	Low	Medium	High
End concrete barrier/bridge parallel to road	  BCT or NCHRP type	  Partial – fishtail or 'Texas Twist'	  High-impact type
Concrete fence/barrier perpendicular to road			 

2.3.3 Driveway accesses

The number of driveway accesses on a rural road influences the number of intersection/driveway crashes, rear-end collisions, and crashes that involve collision with an object (eg culvert headwall). For this research, data was collected on the number and type of road accesses. For commercial accesses, an assessment of usage was undertaken using a three-point scale – low, medium, and high.

The data involved the following parameters:

- type of access:
  - farm/residential (no.)
  - commercial (no.)
- usage (for commercial accesses only) – L/M/H (no.).

Table 2.3 below illustrates examples of various vehicle accesses.

Table 2.3 Examples of vehicle access points

2.4 Traffic volumes

Traffic-volume data for the surveyed sections was extracted from the national State Highway Traffic Volumes (2003–2007) database. Where data on the precise section was not available, estimates based on volume levels at nearby sites were used.

2.5 Road alignment and cross section

Road alignment and cross-section data was initially obtained from the State Highway RAMM database. This data was subsequently assessed for accuracy by the project team. The following variables were used:

- road crossfall (percentage)
- road curvature (1000/R, where R is the radius of curvature)
- road gradient (percentage).

Sections that had been selected to be in the data set used for the development of crash models were also analysed for horizontal consistency. Previous research, undertaken by Turner and Tate (2009), found that the horizontal consistency of a section of state highway can be estimated by comparing the speed environment of the section, V_{500} (85th percentile speed averaged over a distance of 500m approaching the curve) with the 85th percentile curve negotiation speed V_c averaged over the entire length of the study section. V_c and V_{500} of the sections are calculated as per the equations below:

$$V_{500} = 2.1019(AS_{500})^{0.8432} \quad \text{(Equation 2.1)}$$

$$V_c = -24.967 + 0.397V_{500} + 0.741e^{(4.7142 - 26.736/R)} \quad \text{(Equation 2.2)}$$

where AS_{500} is the mean advisory speed over a distance of 500m approaching the study section.

The speed environment V_{500} and 85th percentile curve negotiation speed V_c were calculated for each of the sections included in the sample set.

2.6 Surface characteristics

Skid resistance, which describes the surface condition of the section of road, is an important factor affecting a driver's safety level. Various parameters measuring the condition of the road surfaces were analysed to determine which ones best described the level of skid resistance. These included the 'SCRIM coefficient', 'macrotexture' and 'microtexture':

- Macrotexture and microtexture are widely recognised as important parameters influencing the skid resistance of a section of road. Microtexture refers to irregularities in the surfaces of the stones in the pavement, while macrotexture relates to larger irregularities in the road surface, and takes into account the effects of stone particle size, spacing and arrangement.
- The Sideways Force Coefficient Investigation Machine (SCRIM) has been the basis for measurement of skid resistance for most of the recent work being done in the UK and Australia to assess the quality of road surfaces. The SCRIM coefficient of a road surface is measured by the 'SCRIM+' vehicle, which consists of a truck fitted with either one or two extra wheels along the left and right wheelpaths and aligned at an angle of 20 degrees to the direction of travel. Water is sprayed in front of the test wheel(s) in a controlled manner. This arrangement allows measurement of the sideways force generated by the wheel as it travels in a particular direction. The skid resistance offered by the pavement is computed by taking the ratio of the force created at right angles to the plane of the wheel to the vertical load on the wheel. Data on the sideways forces generated is stored continuously on an onboard computer.

> The SCRIM coefficient was selected as an appropriate measure of the skid resistance because of its wide acceptability and close relationship with surface macrotexture, and the fact that it can reasonably estimate the effects of both macrotexture as well as microtexture. The SCRIM coefficient data for the roadway is a measure of the road surface friction, which influences the ability of drivers to avoid a conflict situation, negotiate curves in the road and regain control of their vehicle during a crash or a near miss.

SCRIM data for each wheelpath (left and right) for the selected sections, adjusted for annual factors, was extracted from the RAMM database for the period 2003–2007. The average value of the SCRIM coefficient over both wheelpaths and over the five-year period was used as an input in the model development.

Other surface characteristics that can be used to understand the role that surfacing has on crashes include roughness, rutting and flushing. Although these are important factors, particularly on certain sections of the network such as low-radius curves, they are not expected to have as important an effect on crash rates as skid resistance. These factors are more likely to be an issue on local roads, and as such need to be considered in crash models looking at those categories of roads.

Research undertaken previously by Cairney et al (2008) utilised data from major road links in the Australian state of Victoria to assess the relative importance of various measures of road skid resistance. The study found no significant correlation between macrotexture, roughness and rutting. However, a good fit was found to exist between plots of crashes and both macrotexture and roughness. This suggests that both macrotexture and roughness are factors that should be considered in crash models.

3 Data collection

The first part of the data collection process involved the identification and selection of suitable sites on the rural road network. Once suitable sites were identified, data on the roadside environment variables described in section 2 was collected via a field survey using survey sheets.

Data collection sheets and a Work Instruction document were prepared for the surveys, which were carried out during December 2007 and February 2008 by a road inventory team based in Tauranga. The full 'Work Instruction for Surveys' document is included in the appendix.

The raw data from the survey sheets was then entered into a spreadsheet for analysis. MATLAB (numerical computation software) was also utilised for data visualisation and further analysis.

A final set of 148 sections, each of them 400m in length, was accepted for input into the models (see section 5.4). The sections were separated according to direction, which resulted in a total of 296 directional 400m sections in the sample set.

3.1 Distribution of variables

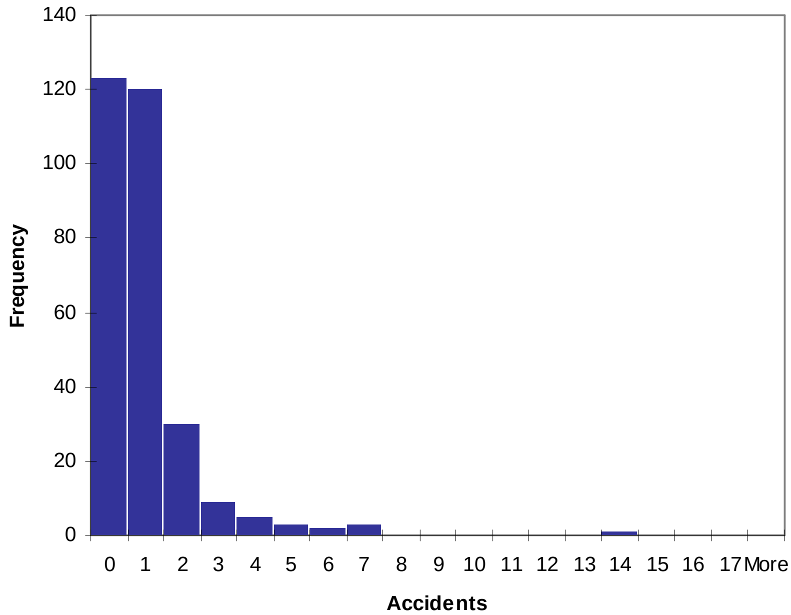
The distributions of key variables within the sample set consisting of 296 directional 400m sections are shown in figures 3.1 to 3.11.

3.1.1 Crashes

Figure 3.1 depicts the distribution of crashes at the sites included in the sample set. Crash data for the surveyed sections was extracted from the national Crash Analysis System (CAS) database for the years 1996–2007, inclusive.

Most of the selected sites had had either zero or one crash during the analysis period. Only 18% of the sites had had more than one crash.

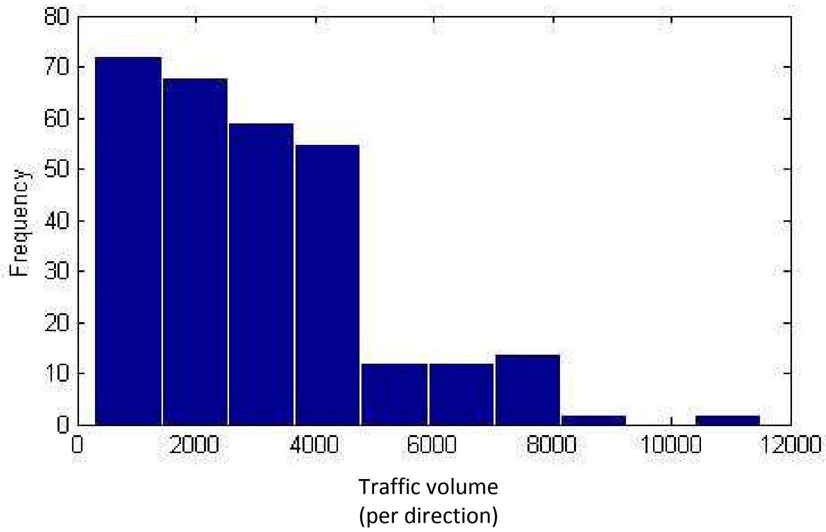
Figure 3.1 Crashes (number per directional 400m section)



3.1.2 Traffic volume

Figure 3.2 illustrates the distribution of average AADT for the selected sites. Eighty-seven percent of the sites had an AADT of less than 5000 vehicles per day.

Figure 3.2 Traffic volume (vph) per directional 400m section

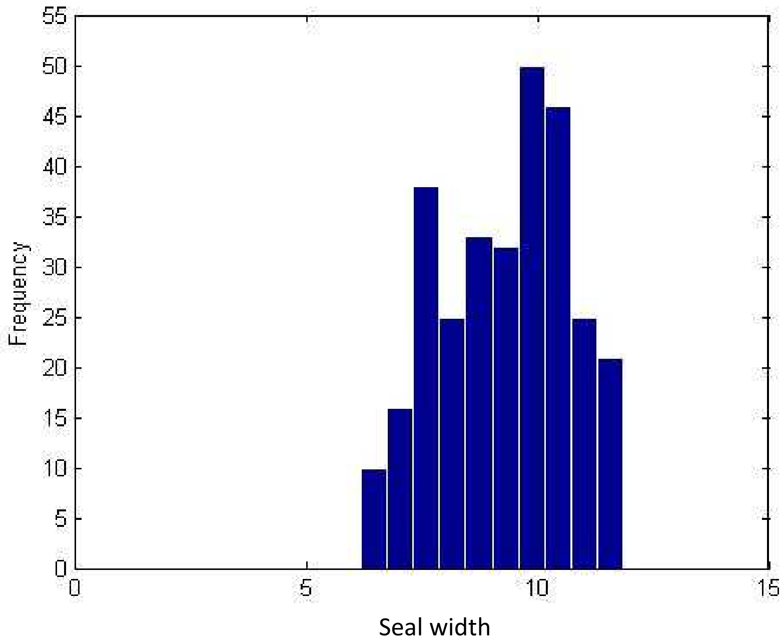


3.1.3 Seal width

Figure 3.3 depicts the distribution of seal width for the selected sites. Sites with seal widths greater than 12m were excluded from the sample set, as an analysis of RAMM geometry data for the state highway network showed that sections with a seal width greater than 12m were mostly either four-lane sections, sections with passing lanes, or two-lane sections located in urban areas – ie not suitable for this study.

Forty-eight percent of the sites in the sample set for this study had a seal width between 9m and 11m.

Figure 3.3 Seal width (m)



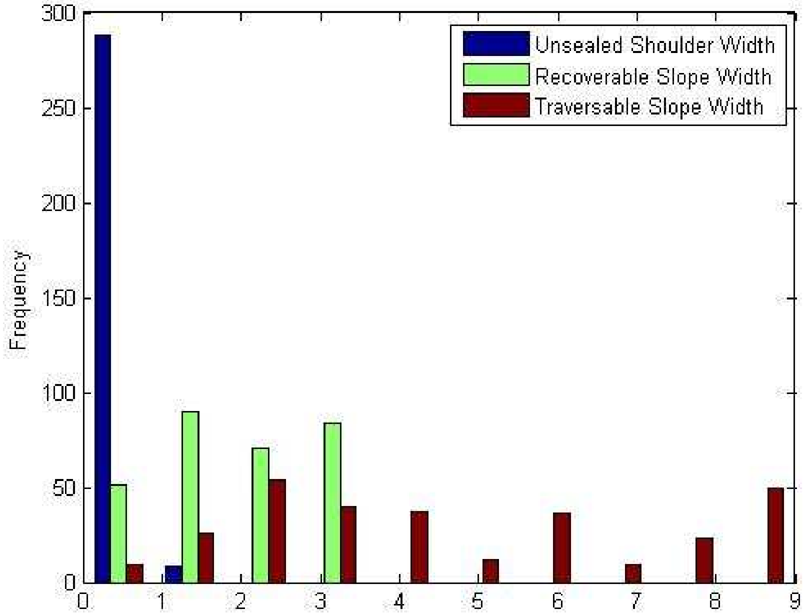
3.1.4 Shoulder environment

Figure 3.4 shows the distributions of the following shoulder environment variables:

- unsealed shoulder width
- recoverable slope width
- traversable slope width.

The majority of sites (99%) had an unsealed shoulder width of less than 1m, and 72% of sites did not have any unsealed shoulder. All sites were observed to have a recoverable slope width of less than 4m. The traversable slope width varied across the sites and was found to be between 0m and 9m.

Figure 3.4 Shoulder environment (m) per directional 400m section



3.1.5 Point hazards

Figure 3.5 illustrates the distribution of the combined sum of all point hazards for the selected sites. It was observed that 38% of the sites had no point hazards, while another 38% had between 0 and 1 point hazard within an average 100m length.

Figure 3.5 Number of point hazards per 100m (average value for each directional 400m section)

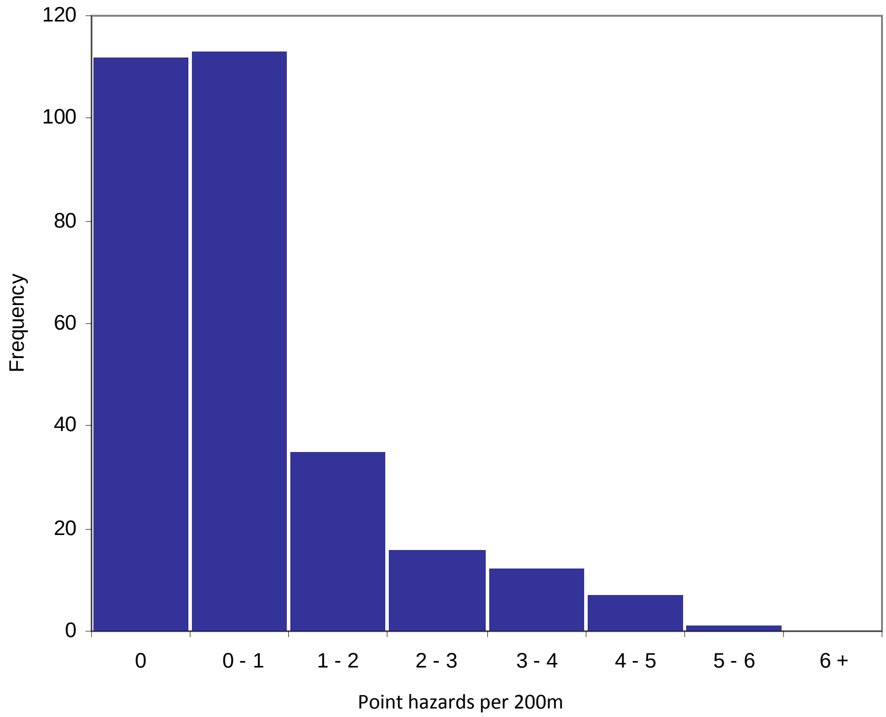


Table 3.1 shows the frequency of occurrence of individual types of point hazards per 100m for each of the 400m sections surveyed. While some point hazards such as wooden poles, culverts, concrete poles and trees were present in significant numbers, others such as light columns, heavy street poles and sign supports were not often present on the sections surveyed.

Table 3.1 Point hazards by type

Frequency	Type of point hazard									
	Wooden pole	Light column	Concrete pole	Heavy street pole	Sign support	Tree	Culvert – roadside	Culvert – roadside with non-traversable headwall	High-impact roadside furniture	Combined point hazards
0–1	75	1	40	0	3	30	71	89	4	**113**
1–2	16	0	18	0	0	11	2	2	0	**35**
2–3	4	0	0	0	0	0	0	0	0	**16**
3–4	2	0	0	0	0	1	0	0	0	**12**
4–5	0	0	0	0	0	1	0	0	0	**7**
5–6	0	0	0	0	0	0	0	0	0	**1**
6+	0	0	0	0	0	0	0	0	0	**0**

3.1.6 Accesses

Figure 3.6 depicts the distribution of the total number of accesses (the sum of farm/residential and commercial accesses) for each of the 400m sections included in the sample set. No accesses were present in 44% of the sections in the sample set, and the data for the remaining sections indicated a fairly low proportion of accesses.

Figure 3.6 Combined accesses per directional 400m section

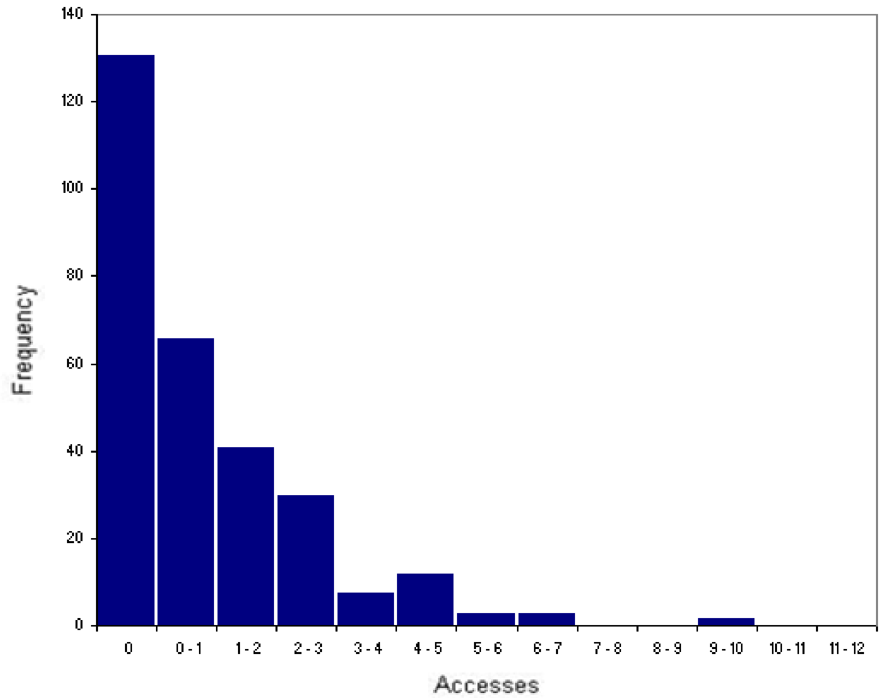


Table 3.2 below shows the distribution of the number of accesses by type. The data was sorted according to the two types of accesses for which data collection was undertaken; ie farm/residential and commercial accesses.

Table 3.2 Number of accesses per 400m by type

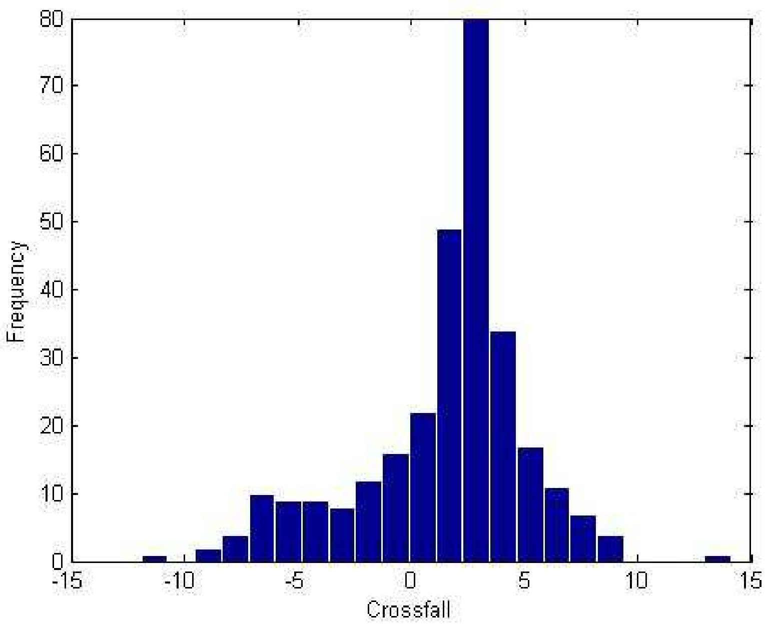
Number of accesses per 400m	Frequency	
	Farm/residential accesses	Commercial accesses
0–1	67	9
1–2	40	3
2–3	27	0
3–4	7	0
4–5	13	0
5–6	2	0
6–7	3	0
7–8	0	0
8–9	0	0
9–10	2	0

3.1.7 Road geometry

Figures 3.7, 3.8 and 3.9 depict the range of values of crossfall, curvature and gradient, respectively, of the sites included in the sample section.

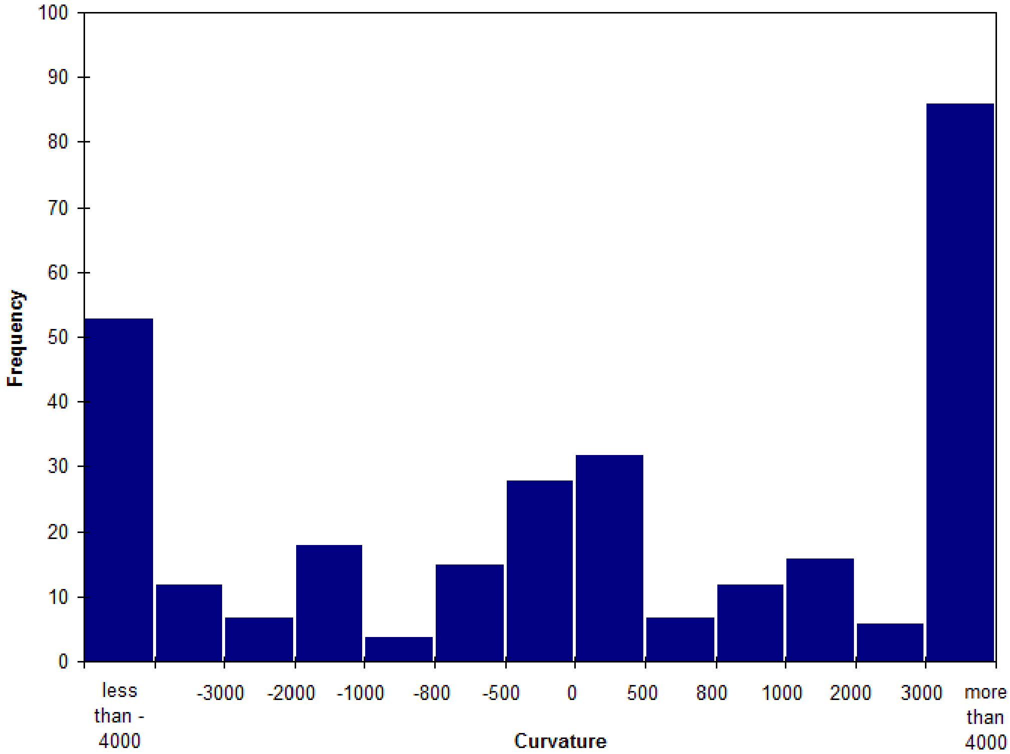
Among the sites included in the sample set, 76% had a positive value of crossfall, while the rest had a negative value of crossfall, as shown in figure 3.7.

Figure 3.7 Road crossfall per directional 400m section



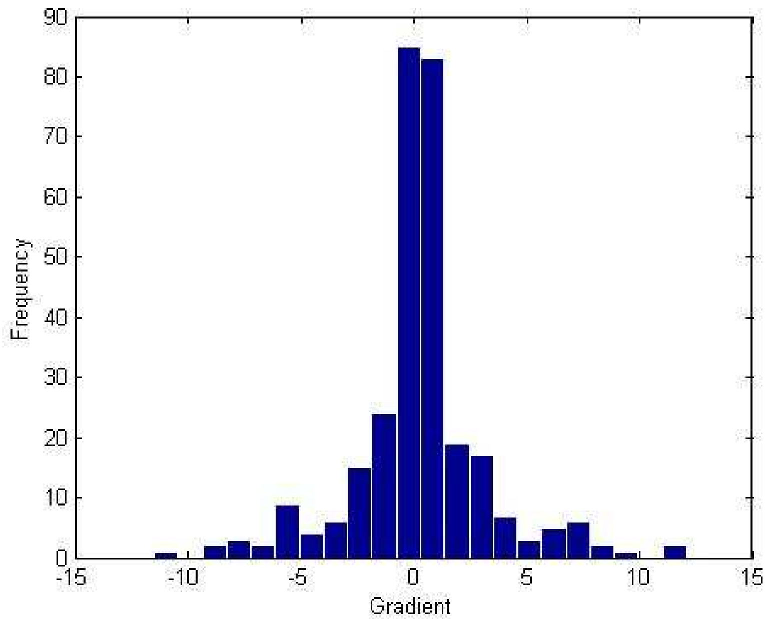
Twenty-nine percent of the sections had an average value of radius of curvature that was less than 800m, indicating the presence of curves, while the rest of the sections were relatively straight.

Figure 3.8 Road curvature (metres) per directional 400m section



The value of gradient for 59% of sections was between -1 and 1, indicating that these sections were quite flat. The rest of the sections had higher values of gradient, as shown in the figure below.

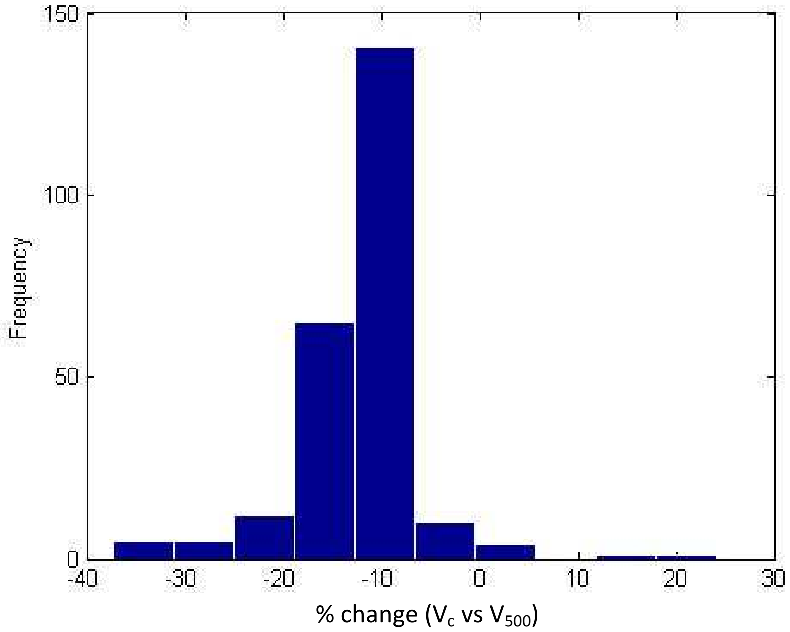
Figure 3.9 Road gradient per directional 400m section



3.1.8 Horizontal consistency

The values of V_{500} (85th percentile speed averaged over a distance of 500m approaching the curve) and V_c (85th percentile curve negotiation speed) were estimated by using the road geometry data. Variations in the values of curve negotiation speed and 85th percentile speed for the sections were then used to estimate their horizontal consistency, as described in section 2.4. Figure 3.10 depicts the percentage change in V_c with respect to V_{500}.

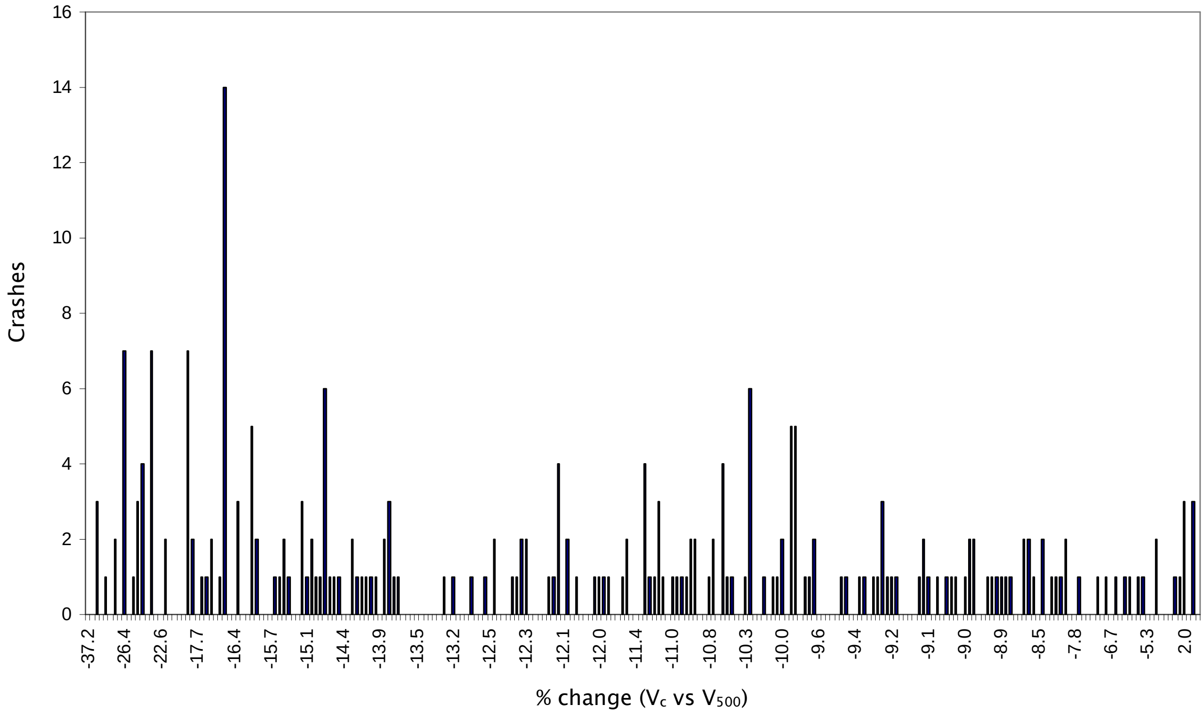
Figure 3.10 Variability in curve speed as compared with the speed on the approaching 500m section



From the predicted values of curve negotiation speeds, it can be seen that drivers at a majority of these sections would need to reduce their speeds by less than 20%, which indicated that the selected sections had a good degree of horizontal consistency. However, a small percentage of sections were found to require a larger percentage change in the mean speed, which suggested that these sections were subject to more pronounced variations in horizontal geometry elements, such as curve radius.

Figure 3.11 below compares the variability in curve negotiation speed with the number of crashes for each of the selected sections. The figure indicates that sections that were associated with a higher variation in speed appeared to have a higher number of crashes. This was consistent with the results found by Koorey and Tate (1997).

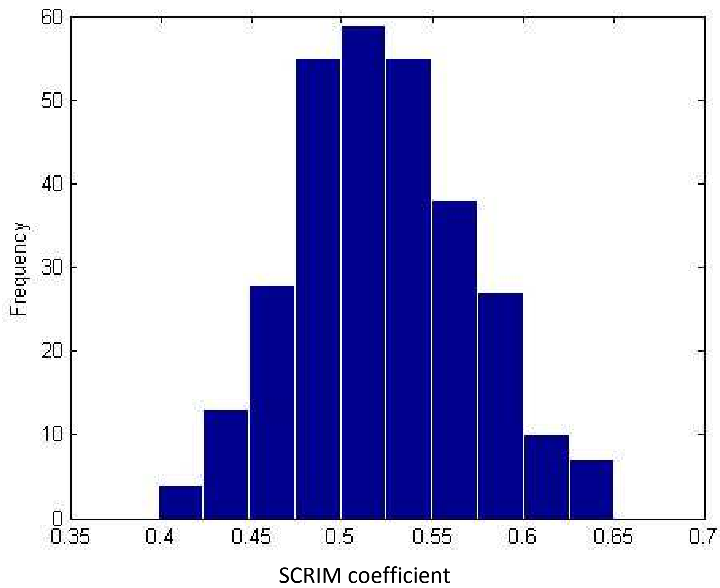
Figure 3.11 Variability in curve negotiation speed vs number of crashes



3.1.9 SCRIM coefficient

Figure 3.12 depicts the distribution of SCRIM values for sections included in the sample set. The set of SCRIM values varied between 0.4 and 0.65. This indicated that the sample set had a good range of sections with low skid resistance (SCRIM coefficient of 0.5 or lower) and high skid resistance (SCRIM coefficient of 0.6 and above).

Figure 3.12 Distribution of SCRIM coefficient per directional 400m section



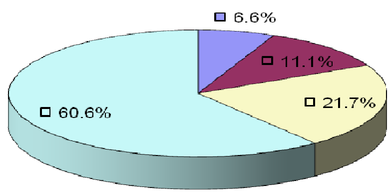
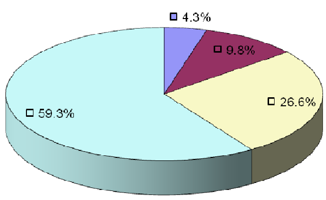
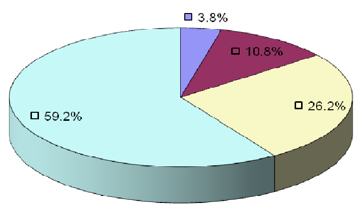
4 Crash analysis

Crash data for the surveyed sections was extracted from the national CAS database for the years 1996–2007, inclusive. Data on the direction of travel of the main vehicle involved in the crash (vehicle 1) was also extracted from CAS, and this was cross-referenced with the corresponding direction of the highway ('increasing'/'decreasing'), which resulted in separation of crashes according to direction of travel along the highway. Crashes within each of the 400m sample sections were identified. Crash locations from CAS were also cross-referenced against the crashes in Koorey (2009), which corrected a number of the crash locations produced in CAS.

Crash data for all rural state highways in the Waikato region, and all rural state highways in New Zealand, was also extracted from the CAS database. This was used to compare the crashes occurring on the selected sections with the crash patterns on state highways in the Waikato region and in the whole of New Zealand. Results of this analysis are given in tables 4.1 and 4.2 and figure 4.1.

Table 4.1 compares the proportion of crashes, according to severity in the sample set, against all Waikato and New Zealand state highways. The proportion of fatal-injury crashes in the sample set was slightly higher, at 6.6%, than for Waikato state highways (4.3%) and for all New Zealand state highways (3.8%). The proportion of serious and minor crashes was comparable, while the number of non-injury crashes in the sample set was similar to those occurring in the Waikato region, which in turn was lower than the national average.

Table 4.1 Comparison of crashes, by severity

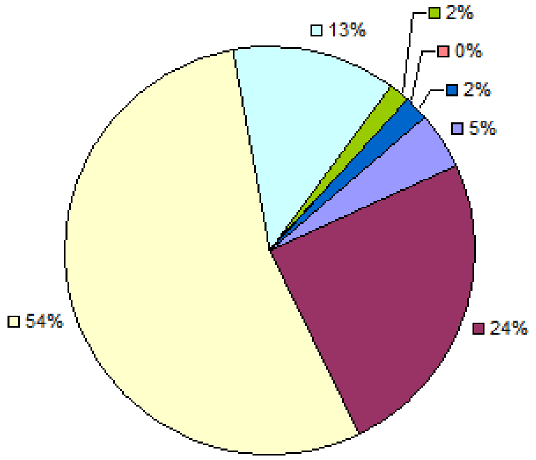
Sample set	
	
All Waikato state highways	**All New Zealand state highways**
	

Key:
Fatal Serious Minor Non-inj

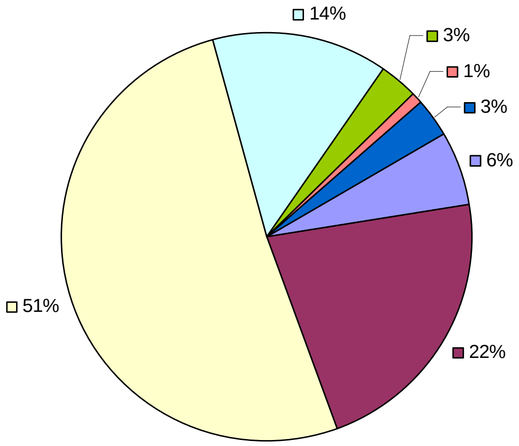
Table 4.2 depicts proportions of crashes by type for the sample set, Waikato State Highways and New Zealand State Highways. While the proportions of most of the crash types in the sample set were comparable to those in the Waikato region as a whole, the proportions of 'bend – lost control'/ head-on' and 'rear-end'/'obstruction' type crashes were marginally higher and lower, respectively, than those on all Waikato state highways.

Table 4.2 Comparison of crashes, by type

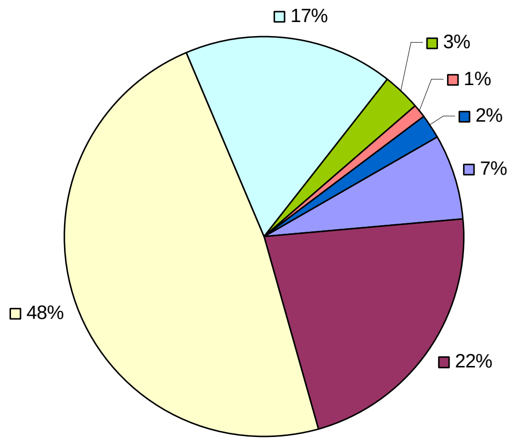
Sample set



All Waikato state highways



All New Zealand state highways

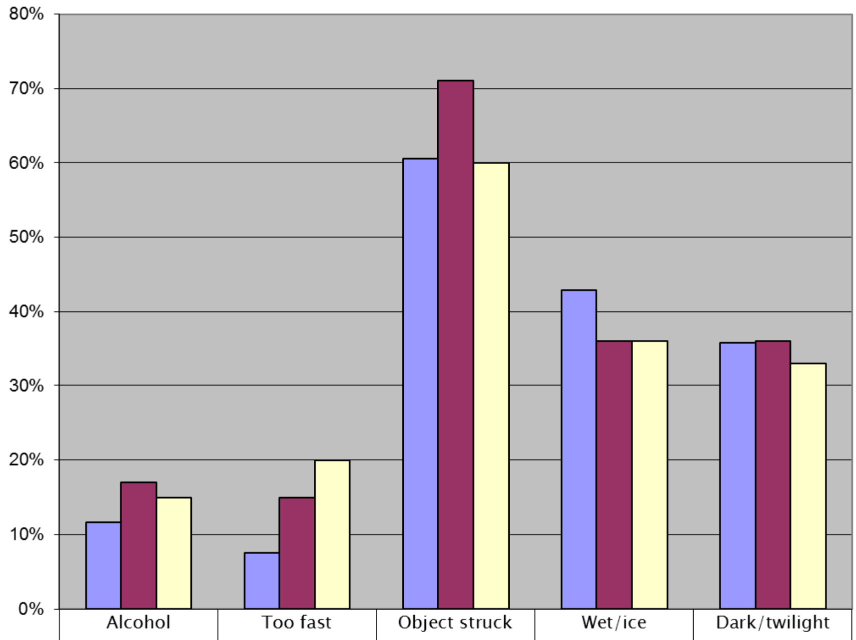


Key:

- Overtaking Crashes
- Straight Road Lost Control/Head On
- Bend - Lost Control/Head On
- Rear End/Obstruction
- Crossing/Turning
- Pedestrian Crashes
- Miscellaneous Crashes

Figure 4.1 illustrates differences in key crash-causing factors for the three data sets. The proportion of crashes caused by speeding drivers was lower in the sample set than the average for Waikato and the whole of New Zealand. The other factors showed proportions similar to those observed for all state highways in Waikato and New Zealand.

Figure 4.1 Crash causal factors



	Alcohol	Too fast	Object struck	Wet/ice	Dark/twilight
Sample set	12%	8%	61%	43%	36%
All Waikato state highways	17%	15%	71%	36%	36%
All New Zealand state highways	15%	20%	60%	36%	33%

5 Data preparation

Data on each variable for each 400m sample section was input into a Microsoft Excel spreadsheet. The data was broken down further into 100m segments on both sides of the road, where relevant (ie 4 or 8 values for each variable per section). Table 5.1 below shows the format of the spreadsheet with a select set of variables for a single 400m section. Table 5.2 shows a screenshot of the data collection spreadsheet.

Table 5.1 Format of data entry spreadsheet

Code100[1]	Code400[2]	Incr/ decr	Road-RS	RP100[3]	RP400[3]	AADT 400[4]	Accidents (400m)[4]	Seal width[5]	Unsealed shoulder width[5]	High-impact roadside furniture[5]
I-032-0015/1363	I-032-0015/1363	I	032-0015	1363	1363	1350	1	7.8	0	0
-032-0015/1463	I-032-0015/1363	I	032-0015	1463	1363	1350	1	7.4	0	0
I-032-0015/1563	I-032-0015/1363	I	032-0015	1563	1363	1350	1	7.8	0	0
I-032-0015/1663	I-032-0015/1363	I	032-0015	1663	1363	1350	1	7.8	0	0
D-032-0015/1763	D-032-0015/1763	D	032-0015	1763	1763	1350	0	7.8	0	0
D-032-0015/1663	D-032-0015/1763	D	032-0015	1663	1763	1350	0	7.8	0	0
D-032-0015/1563	D-032-0015/1763	D	032-0015	1563	1763	1350	0	7.4	0	0
D-032-0015/1463	D-032-0015/1763	D	032-0015	1463	1763	1350	0	7.8	0	0

1 Code100: RS/RP location for start of 100m sections within each 400m section

2 Code400: RS/RP location for start of 400m sections

3 RP100, RP400: Route position for start of 100m and 400m sections respectively

4 AADT400, Accidents 400: AADT and number of reported injury crashes for each 400m section

5 Seal width, unsealed shoulder width etc: Average value of variable (eg seal width) for each 100m section

The following sections describe adjustments that were made to individual predictor variables during the data analysis process.

5.1 Point hazards

In addition to the individual hazard types specified below, the total number of hazards in each 100m section were combined to produce a combined point hazards variable. The following variables were included in this combination:

- wooden pole
- light column

- concrete pole
- heavy street pole
- sign support
- tree
- culvert – roadside
- culvert – roadside with non-traversable headwall
- high-impact roadside furniture.

Table 5.2 Format of data entry spreadsheet

Code400	AADT	Accidents	Seal Width	Unsealed Shoulder Width	Recoverable Slope Width	Traversable Slope Width	Wood Pole	Light Column	Concrete Pole	Heavy Street Pole	Signs Supports	Trees	Culverts-roadside	Culverts-road with non-traversable headwall	High Impact Roadside Furniture	Non-traversable slope / Perpendicular deep drain	End Concrete Barrier/ Bridge Parallel to road	Concrete Fence/ Barrier Perpindicular to road	Farm/Residential	Commercial	Crossfall	Curvature	Gradient	Abs Crossfall	Abs Curvature	Abs Gradient	Scaled Crossfall	Scaled Curvature	Scaled Gradient	Average SCRIM
D-002-0037/2533	2438	0	10.4	0	0.5	2	0	0	0	0	0	0	0.25	0.25	0	2	0	0	0.25	0	0.2	3745	1.1	0.2	3745	1.1	0.928	1.004	1.054	0.4727
D-002-0037/4008	4000	2	10.2	0	0	9	0.5	0	0	0	0	0.75	0	0	0	0	0	0	0	0	14.1	121	-3.2	14.1	121	3.2	1.679	0.968	0.760	0.4479
D-002-0048/4277	4000	0	10.45	0.125	0.875	9	0.75	0	1.5	0	0	0	0.25	0.25	0	0	0	0	0.5	0	-1.2	-3745	-0.7	1.2	3745	0.7	0.853	0.930	0.931	0.5720
D-002-0061/5432	3040	1	10.25	0.25	1.75	4	0	0	0	0	0	0	0	0	0	1.75	0	0	0.75	0	-5.7	-714	0.9	5.7	714	0.9	0.609	0.960	1.040	0.5626
D-002-0061/6397	5000	1	10	0	0.25	2.25	0	0	0	0	0	0	0	0	0	2.25	0	0	0.5	0	2.6	-14925	0.5	2.6	14925	0.5	1.058	0.819	1.013	0.5446
D-002-0073/6231	6500	3	9.325	0	2.25	3.75	0	0	0	0	0	0	0	0	0	0.75	0	0	0	0	6.5	461	3.4	6.5	461	3.4	1.269	0.971	1.211	0.5147
D-002-0100/1280	4017	1	11	0	1.75	6.5	0.25	0	0	0	0	2	0	0	0	0	0	0	0.25	0	1.4	3745	0.5	1.4	3745	0.5	0.993	1.004	1.013	0.5647
D-002-0189/13516	2514	1	10.475	0.125	1.75	2	0	0	0	0	0	0	0	0	0	2	0	0	0	0	6.4	909	0.6	6.4	909	0.6	1.263	0.976	1.020	0.5791
D-002-0189/17051	4200	1	10.725	0	1.5	1.5	0	0	0	0	0	0	0	0	0	1.5	0	0	0	0	0	-811	0.1	0	811	0.1	0.917	0.959	0.985	0.5948
D-002-0189/868	2516	0	10.3	0.125	2.5	6	0.75	0	0	0	0	0	0.25	0	0	1.5	0	0	0.75	0	1.1	-30303	1.9	1.1	30303	1.9	0.977	0.667	1.108	0.5592
D-002-0209/13124	4079	0	10.2	1.25	2.5	5.5	0.5	0	0.5	0	0	0	0	0	0	1	0	0	0.5	0	3.8	-5000	-1.3	3.8	5000	1.3	1.123	0.917	0.890	0.5575
D-002-0223/3826	2250	0	9.4	0	0.5	1.25	0	0	0	0	0	0	0	0	0	1.25	0	0	0	0	2.4	1364	1.2	2.4	1364	1.2	1.047	0.980	1.061	0.5986
D-002-0223/8000	1297	1	8.7	1	2.75	9	1.5	0	1.25	0	0	0	0	0.25	0	0	0	0	1.25	0	8.4	101	0.2	8.4	101	0.2	1.371	0.968	0.992	0.5565
D-002-0232/2904	1640	0	8.225	0	1.75	2.25	0	0	1	0	0	0	0	0	0	2.25	0	0	0.25	0	-2.6	-1667	-0.5	2.6	1667	0.5	0.777	0.950	0.945	0.5486
D-002-0232/6202	3100	1	7.525	0	1.25	2	0	0	0	0	0	0	0	0	0	2	0	0	0	0	-6.2	-103	-0.6	6.2	103	0.6	0.582	0.966	0.938	0.5413
D-002-0243/7993	1600	1	9.3	0	1.75	5.5	1	0	1	0	0	0	0	0.25	0	1	0	0	0.5	0	2.9	5000	0.5	2.9	5000	0.5	1.074	1.016	1.013	0.5903
D-002-0258/740	2750	0	10	0.125	1.5	3	0	0	0	0	0	0	0	0	0	3	0	0	0.5	0	-6.8	-508	3.6	6.8	508	3.6	0.550	0.962	1.224	0.5787
D-002-0271/1185	696	1	8.2	0	1.5	1.75	0	0	0	0	0	0	0	0	0	1.75	0	0	0.25	0	2.4	7519	0.4	2.4	7519	0.4	1.047	1.041	1.006	0.5135
D-002-0271/4126	1000	0	9.75	0.25	1.375	2.875	0.25	0	0	0	0	0.25	0	0.25	0	0.625	0	0	0	0	5.5	390	-2.7	5.5	390	2.7	1.215	0.971	0.795	0.6250
D-002-0283/5392	3100	0	8.625	0	1	2	0	0	0	0	0	0	0	0	0	2	0	0	0.5	0	2.7	10000	3.6	2.7	10000	3.6	1.063	1.066	1.224	0.6189

5.2 Accesses

The number of farm and commercial accesses was specified for each 100m section. For farm and residential accesses, the total number was specified, based on the number of letterboxes. Commercial accesses were classified as per the following level of activity:

- low (approximately 1–30 vehicle movements per day)
- moderate (approximately 30–100 vehicle movements per day)
- high (approximately 100+ vehicle movements per day).

These categories were then combined by using a weight of 0.5, 1.0 and 1.5 respectively for each level of activity.

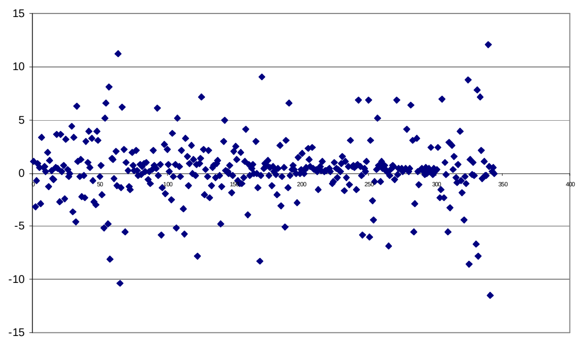
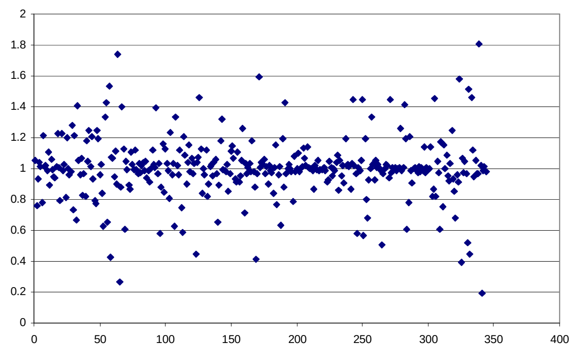
Farm/residential and commercial accesses were also combined to produce a variable denoting the total number of accesses in each 100m section.

5.3 Road geometry

All of the road geometry variables (crossfall, curvature and gradient) can take on a negative value (eg decreasing gradient, and left-hand vs right-hand curves), which is problematic for generalised linear regression modelling. Furthermore, in this study the data was being considered directionally, so taking the absolute values of these variables did not retain the magnitude of variation within each of the variables.

Therefore, scaled variables were also included. The variables were scaled such that their mean was 1 and their standard deviation was 0.2. This is a straightforward linear transformation and therefore does not affect correlation. As an example of the outcome of the scaling process, figure 5.1 below shows scatter plots of the values of 'gradient' for sections included in the sample set, before and after scaling.

Figure 5.1 Actual vs scaled values of gradient

Gradient – Actual values	Gradient – Scaled values
	

The scatter plot on the right in figure 5.1 depicts the scaled gradient values obtained by linear transformation of the original data set. As is obvious from the above diagrams, the degree of variation among the different data points has been preserved, while resulting in a set of values which are all positive and thus can be used as inputs in the linear regression modelling process.

From the raw data on crossfall, curvature and gradient, the following six variables were derived and used in the modelling:

- absolute crossfall
- absolute curvature
- absolute gradient
- scaled crossfall
- scaled curvature
- scaled gradient.

5.4 Accuracy of the data

From the initial list of 200 road sections, 52 sections were rejected for the following reasons:

- Data interrogation and outlier analysis revealed errors in the raw roadside and crash data. These were corrected or omitted, and this resulted in the removal of 28 sections.
- Six sections were omitted because of lack of sufficient SCRIM data.
- Eighteen sections were omitted because they had a seal width of more than 12m.

A final set of 148 sections was accepted for input into the models. Since all data had been separated by direction, this resulted in a total of 296 data points for each model build.

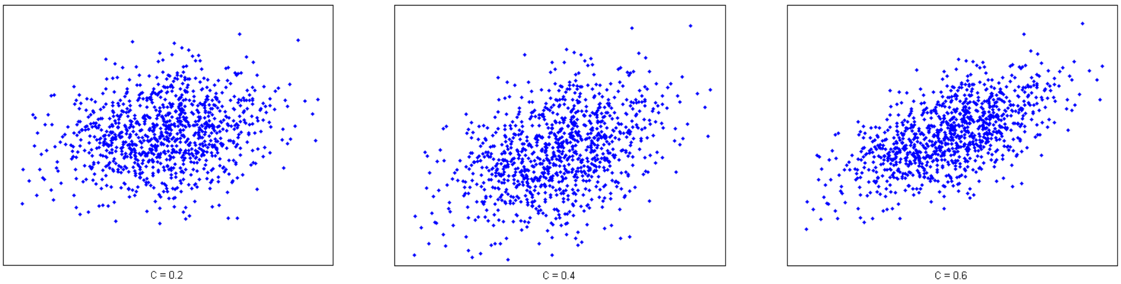
The main study phase of this project will focus on developing improved processes for measuring the accuracy of the data to be used for developing the crash models. Part of this exercise will focus on checking the accuracy of existing KiwiRAP data by comparing the KiwiRAP predicted roadside hazard rating with the roadside hazard data that was collected during this pilot study. Other sources of data that may lead to improvements over the data that is available at present will also be explored during the course of the main study.

5.5 Checking for correlation

The large number of predictor variables in the rural roads crash model adds to the complexity of the modelling task, in particular through the cross correlation of variables. A high degree of correlation between two variables can result in the relationship between some variables and crashes not being reported correctly in the model. When two variables are correlated, the addition of the second variable to the model may show a weak relationship with crashes, whereas the first variable might have a strong relationship. In some instances, the second variable may show a totally different relationship with crashes (eg a negative rather than positive relationship) because of the correlation. Typically, it is best to choose only one of two highly correlated variables for a model. This normally does not have an impact on the predictive power of the model, as both variables tend to explain the same amount of variability in the predictor variable.

As an illustration of the implication of the correlation value, figure 5.2 shows scatter plots of two variables with correlations of 0.2, 0.4 and 0.6 respectively.

Figure 5.2 Examples of correlation



Of the many combinations of variables, the 32 variable pairs listed in table 5.3 produced a correlation with absolute value greater than 0.2.

Table 5.3 Independent variable cross correlation

Serial number	Variable A	Variable B	Correlation
1	Combined accesses	Farm/residential	0.99
2	Wooden pole	Combined point hazards	0.77
3	Combined point hazards	Concrete pole	0.63
4	Combined point hazards	Tree	0.58
5	Traversable slope width	Combined point hazards	0.54
6	Traversable slope width	Wooden pole	0.53
7	AADT	Seal width	0.52
8	Recoverable slope width	Traversable slope width	0.51
9	Traversable slope width	Distance to non-traversable slope/ perpendicular deep drain	-0.49
10	Combined accesses	Culvert – roadside	0.47
11	Farm/residential	Culvert – roadside	0.46
12	Distance to non-traversable slope/ perpendicular deep drain	Combined point hazards	-0.40
13	Combined point hazards	Culvert – road with non-traversable headwall	0.40
14	Wooden pole	Distance to non-traversable slope/ perpendicular deep drain	-0.40
15	Wooden pole	Concrete pole	0.38
16	Combined accesses	Combined point hazards	0.36
17	Farm/residential	Combined point hazards	0.35
18	Traversable slope width	Concrete pole	0.35

19	Combined point hazards	Culvert – roadside	0.31
20	Traversable slope width	Tree	0.29
21	Combined accesses	Concrete pole	0.28
22	Farm/residential	Concrete pole	0.28
23	Wooden pole	Culvert – road with non-traversable headwall	0.26
24	Distance to non-traversable slope/ perpendicular deep drain	Concrete pole	-0.25
25	AADT	SCRIM	-0.25
26	Absolute curvature	Absolute crossfall	-0.23
27	Wooden pole	Combined accesses	0.23
28	Wooden pole	Farm/residential	0.23
29	% reduction in curve speed	Concrete pole	0.22
30	% reduction in curve speed	Combined accesses	0.22
31	% reduction in curve speed	Farm/residential	0.21
32	AADT	Concrete pole	0.20

Of the variables assessed, the only correlations that were of concern were the first 14 in the table, where the correlation was above 0.4 (see the middle plot, figure 5.2, which demonstrates the amount of correlation). Ideally, variables that have a reasonable amount of correlation (above 0.4) should not be used in the same model.

Section 5.5.1 below provides a more detailed discussion of some of the high correlations reported in table 5.3.

5.5.1 Discussion on correlations

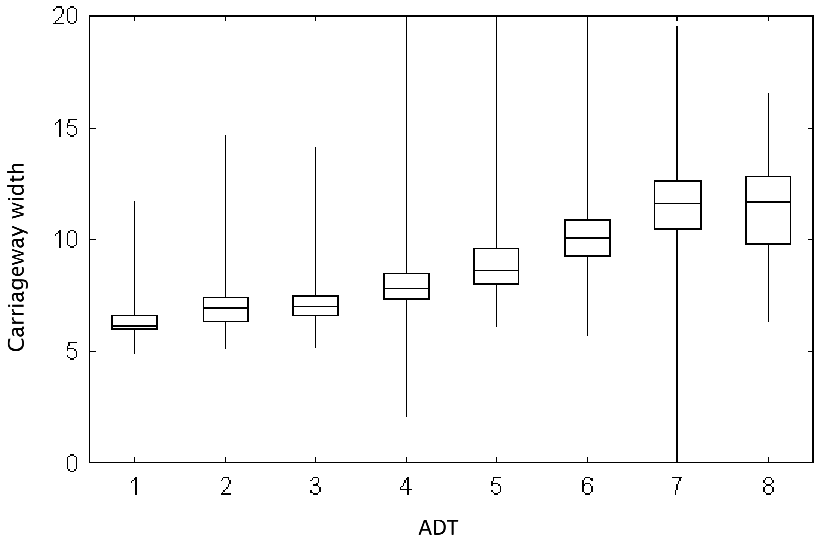
5.5.1.1 Correlations involving traversable slope width

Table 5.3 suggests a high correlation between 'traversable slope width' and both 'combined point hazards' and 'wooden pole' (serial numbers 5 and 6 in table 5.3). These correlations seem intuitive, given the fact that the presence of point hazards in the roadside environment would be likely to result in a reduction in the distance within which a motor vehicle could safely stop or slow down. These high correlations justify the suitability of using an integrated hazard scoring system, such as the KiwiRAP roadside hazard rating, as opposed to looking at each element of the roadside environment and individual types of point hazards separately. This is also reinforced by the high correlation between 'combined point hazards' and 'distance to non-traversable slope' (serial number 12).

5.5.1.2 Correlation between AADT and seal width

Table 5.3 also shows a high correlation coefficient of 0.52 between daily traffic volume (AADT) and the seal width of the sample section. This is not entirely unexpected, since roads that are associated with a high traffic volume are often upgraded and widened. Research conducted previously by Cenek et al (2004) also found a correlation between AADT and carriageway width – they found that higher-volume roads are often characterised by higher carriageway widths, though with a slightly higher degree of variation as well, as shown in figure 5.3.

Figure 5.3 Box-plot of carriageway width versus ADT (Cenek et al 2004)



However, previous research has shown that both AADT and seal width are important predictor variables and both should be included in the list of predictor variables used for crash modelling. During the main study, a suitable approach to overcome the limitation of the high correlation between AADT and seal width may be to define fixed AADT bands and build individual crash models for each of these bands.

5.5.1.3 Correlation between ‘accesses’ and ‘culverts – roadside’

‘Combined accesses’ and ‘farm/residential accesses’ both seem to be highly correlated with the number of ‘culverts – roadside’ (serial numbers 10 and 11 in table 5.3). These correlations suggest that accesses on rural roads are often accompanied by culverts. It thus seems reasonable to discard ‘culvert – roadside’ as one of the predictor variables used for modelling. However, future research could look into the more hazardous culvert headwalls and how such culverts may impact on safety.

5.5.1.4 Correlations involving combined point hazards

Table 5.3 shows a high correlation between ‘combined point hazards’ and variables representing individual point hazards (such as concrete poles, wooden poles, trees and culverts – roadside). These high correlations are intuitive, given that the combined point hazards variable is a sum of the various individual point hazard types. However, these correlations will ultimately not have an effect on the final model, since the combined point hazards variable will be used to estimate the cumulative impact of various point hazard types on the roadside environment.

5.5.1.5 Correlations involving combined accesses

As with ‘combined point hazards’ above, the ‘combined accesses’ variable can be used to estimate the total number of accesses (farm/residential and commercial) within the sections under analysis. This can neutralise the high correlation coefficient between ‘combined accesses’ and ‘farm/residential accesses’ seen in table 5.3 above.

5.5.1.6 Correlations involving shoulder environment variables

Another significant correlation is the one between ‘recoverable slope width’ and ‘traversable slope width’ (serial number 8 in table 5.3). This is again expected, given that the traversable slope often contains a part of the recoverable slope. However, both of these variables are considered to be important predictor variables for use in modelling. To avoid the complexities that may arise from including both of these variables as inputs in the model development process, the definition of traversable slope should be altered to refer only to the slope width starting from the edge of the recoverable slope, instead of from the edge of the seal.

6 Video data capture and analysis

6.1 Video and image data capture and analysis in Australia and New Zealand

Existing use of image-capture technology overseas indicates that such technology is primarily used for asset management of roading infrastructure. In New Zealand it is primarily used as a reference tool for network managers and is not being used to collect roadside hazard information (Turner et al 2005).

ARRB Road Info Ltd and Pavement Management Services Ltd are the two companies that currently provide data collection by image-capture services in New Zealand. ARRB Road Info, which is a joint venture with ARRB Transport Research, uses ARRB software and hardware. Pavement Management Services (PMS) has collected digital image-capture data for a number of road networks, including Manukau City, Palmerston North and all of Transit Region 2 (Auckland). While this data is largely used by New Zealand clients as a reference tool, in Australia the same hardware and software has been used to collect data for asset management inventories, and can be used for roadside hazard studies as well.

Turner et al (2005) also conducted a literature search to identify studies where roadside hazard information had been collected using video data-capture techniques. Only one such study was found. This study, by Prinsloo and Goudanas (2003), sought to review and update the rural roads crash rates for the Roads and Traffic Authority (RTA) of New South Wales. In addition, the project aimed to create a spatial database linking road crashes with other roadway attributes and parameters. A large amount of road attribute data was collected through the interrogation of GipsiCam imagery data. For the viewing and processing of the GipsiCam digital imagery, Prinsloo and Goudanas used an RTA program called 'AssetLoc' (see figure 6.1), which enabled the processing of road inventory data during the accident analysis. AssetLoc uses images taken at 10m intervals. Figure 6.1 shows a screenshot of the AssetLoc software.

Figure 6.1 Screenshot of AssetLoc software

Another related study by Thomsen and O'Brien (2003) collected roadside hazard data as part of asset management. However this study did not investigate the crash risk implications of roadside hazards. Thomsen and O'Brien used a Geographic Information System (GIS) interface and also incorporated Geographic Positioning System (GPS) technology using the DRIVE software (see the screenshot in figure 6.2). Using highway centreline data and aerial photos, the DRIVE analysis software is able to simultaneously display the assessed location on aerial photos and the video capture.

Figure 6.2 DRIVE software (Thomsen and O'Brien 2003)

Thomsen and O'Brien identified the following core benefits of this technology (Turner et al 2005):

- The process could be broken down into stages, such as by item type, locations, data/assessment steps, and then the data could be accumulated progressively as resources allowed.
- The data capture and assessment could be directly overseen and managed, and was repeatable and auditable from the primary data source, giving confidence regarding the quality of the information.
- At a later date, the images could be used to assess aspects that were not foreseen at the time of data capture.

6.2 ARRB data collection and analysis methodology

The ARRB Group were commissioned by the research team to determine whether roadside and access data could be collected using the KiwiRAP video files for the Waikato State Highway network, and if so, how accurate the results of the video analysis method were, compared with manual data collection.

The ARRB Road Info procedure involves the production of a GIS database where images can be accessed at points along routes. This has been used by New Zealand Territorial Local Authorities (TLAs) as a reference tool for customer inquiries, and also for before-and-after studies and safety audits. To date, ARRB Road Info has not used this system for large-scale collection of asset management data that could be used for roadside hazard studies in New Zealand.

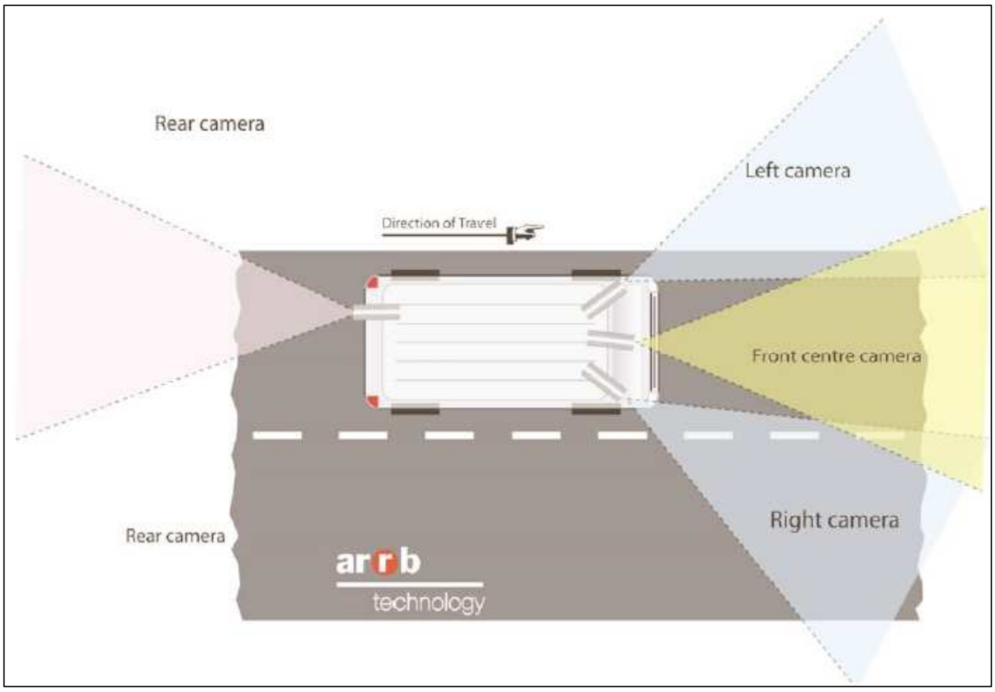
The software used by ARRB makes use of multiple cameras mounted on vehicles that travel the length of the state highway network. These cameras can be set to take digital photos at variable pre-set intervals. Playing back the photos at speed converts the series of still photos (which can be referenced using GPS) into a video.

The cameras mounted on each vehicle require calibration prior to being used for referencing roadside objects. Once calibrated, a grid is overlaid on the images (whether taken from the side or front of the vehicle), and each polygon in this grid approximates a set distance on the ground. However, if the terrain is significantly different to the area where the cameras were calibrated (usually a flat pad), the accuracy of the measurement of offset of roadside objects is reduced. At the time of writing, the images captured by the camera were not of a sufficiently high resolution to enable useful measurement of roadside distances.

6.3 Data analysis outcomes

The ARRB video-processing software was not able to collect the roadside environment, roadside hazard and access data from the videos, mainly because the camera had been focused on the road surface instead of the roadside environment, as shown in figure 6.3 below. This, along with the fact that the videos were not of a high enough resolution, resulted in significant difficulties in calibration during the data analysis process, since measurements such as distance away from the vehicle could not be accurately estimated.

Figure 6.3 Orientation of cameras used for video data collection (www.arrb.com.au)



Hence, although the extraction of the number of point hazards in a given section was possible through the video data analysis process used by ARRB, an accurate estimate of elements such as traversable and non-traversable slope widths and hazard offsets was not possible with the technology and video data that was available at the time.

There were two other options for collecting the level of detail required:

- manual data collection
- utilising the roadside hazard rating score produced by ARRB (as part of the KiwiRAP assessment) from video data, rather than the detailed roadside data.

The KiwiRAP roadside hazard rating index consists of a 4 X 4 matrix-based system of rating roadside hazards within fixed offset distance bands (0–4m, 4–9m, 9–15m, 15m+) and according to the following levels of severity:

- negligible severity – involves the presence of a kerb, wire-rope barrier, or level slope with no hazards
- rigid barriers – eg W-sections and concrete barriers with end treatments
- moderate severity – involves the presence of small trees, shallow drains or frangible poles
- high severity – involves the presence of steep slopes, large trees and unrecoverable slopes steeper than 3:1; also includes divided roads where an errant vehicle could completely cross the median and cause a head-on crash.

This rating score should be considered satisfactory for use as a measure of the quality of the roadside environment during the main study phase of this research, because of the reasonable level of detail that takes into account the quality of the roadside environment and the severity of hazards at varying distances along the roadside. However, the accuracy of the KiwiRAP roadside hazard rating will also be investigated during the main study by comparing the rated score with hazard data that was collected manually during this pilot stage of the research.

Although the KiwiRAP roadside hazard rating appears to be of sufficient quality and accuracy as to enable its use in the main study, it is not expected to surpass the quality of detailed roadside hazard data. However, the major benefit of using the KiwiRAP roadside hazard score will be its ease of use, and the fact that it will allow a much larger sample set to be used for developing the models, owing to the elimination of the requirement for data collection. Use of the KiwiRAP rating will enable easy application of the crash prediction models without the need for investing a significant amount of time and resources into collecting detailed roadside hazard information.

7 Modelling methodology

7.1 Introduction

The aim of crash prediction modelling in this case is to develop relationships between the flow variables (ie mean number of crashes and AADT) and the non-flow predictor variables such as road geometry, roadside hazards, frequency of accesses, and skid resistance of the road surface.

The models are called ‘generalised linear models’ and typically have a negative binomial or Poisson error structure. Generalised linear models were first introduced to modern road crash studies by Maycock and Hall (1984) and extensively developed in Hauer and Hakkert (1989). Turner (2000) further developed and fitted these models for the New Zealand context using crash data and traffic counts for motor vehicle only crashes.

Over recent years, the process has been refined to allow for incorporation of non-flow variables, which allow different functional forms, improved ‘goodness of fit’ statistics, and the selection of ‘preferred’ models. This chapter provides an outline of the modelling processes that we used to develop the model.

7.2 Selecting functional form

The methodology for this research was in accordance with the approaches made in many previous studies that involved fitting generalised regression curves to independent and dependent data. In this study we limited the regression curves to the following form:

$$F = aX_1^{\,b1} X_2^{\,b2} ... X_m^{\,bm} \quad \text{(Equation 7.1)}$$

We considered that the variability around this regression curve was describable with a negative binomial (or, in its special case, Poisson) error structure.

The negative binomial distribution is given by:

$$y(x,r,k) = \frac{\Gamma(x+r)}{\Gamma(x+1)\Gamma(r)} \left(\frac{1}{1+k}\right)^r \left(\frac{k}{1+k}\right)^x \quad \text{(Equation 7.2)}$$

where *r* and *k* are distribution parameters.

Whenever the resultant value for *k* is negative or zero, this is interpreted as meaning that the correct error structure is in fact not negative binomial, but Poisson. In this case, the optimisation algorithm is repeated, with:

$$y(x,\lambda) = \frac{e^{-\lambda}\lambda^x}{\Gamma(x+1)} \text{ and } \lambda = aX_1^{\,b1} X_2^{\,b2} ... \quad \text{(Equation 7.3)}$$

7.3 Fitting crash prediction model parameters

Once the functional form for each variable had been determined, generalised linear models could then be developed using either a negative binomial or Poisson distribution error structure. Software had been developed in Minitab in order to fit such models (ie to estimate the model coefficients) – this could also be done in various other commercial packages, such as GENSTAT, LIMDEP or SAS.

7.4 Adding variables to the models

The log-likelihood, *LL*, is the measure of fit of the model to the observed number of crashes. It is calculated by:

$$LL = \sum sum(\log(y)) \qquad \text{(Equation 7.4)}$$

where *y* is the distribution function, as given in section 7.2. The best-fit function is found by varying these parameters in such a way as to maximise *LL*.

In this research, combinations of predictor variables were modelled to test the goodness of fit of the various possible multi-variable models. While the log-likelihood function was useful for finding the best-fit function and for comparing functions having the same number of variables, comparing functions with different numbers of variables required further analysis. We chose to use the popular Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC). The BIC and AIC are given by:

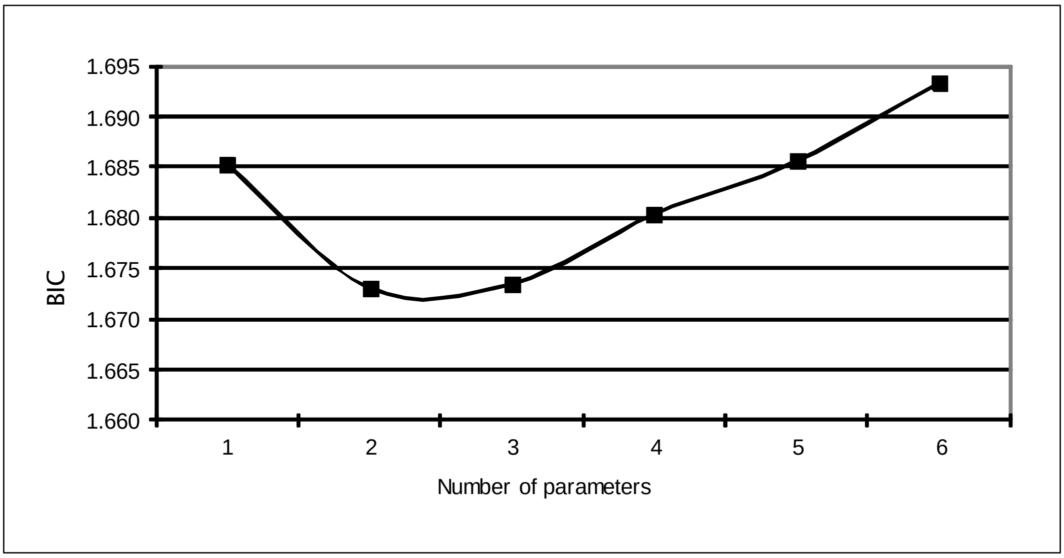
$$BIC = \frac{-2LL + p\ln(n)}{n} \qquad \text{(Equation 7.5)}$$

$$AIC = \frac{-2LL + 2p}{n} \qquad \text{(Equation 7.6)}$$

where *LL* is the log-likelihood, *p* is the number of parameters in the function (the number of exponents plus one for the constant term), and *n* is the cardinality (ie sample size) of the data set from which the function is generated.

Typically, the model with the lowest BIC is the preferred model from a statistical perspective. It indicates that the number of variables explains most of the variability in the data (that can be explained by the full set of variables) and that there is not a lot of merit in adding further variables. Figure 7.1 provides an example of this. As shown in the figure, the BIC values indicated that the parsimonious number of parameters in this case was two, ie addition of further variables would not explain a lot more variability.

Figure 7.1 Graph used to determine the number of parameters yielding the optimal BIC

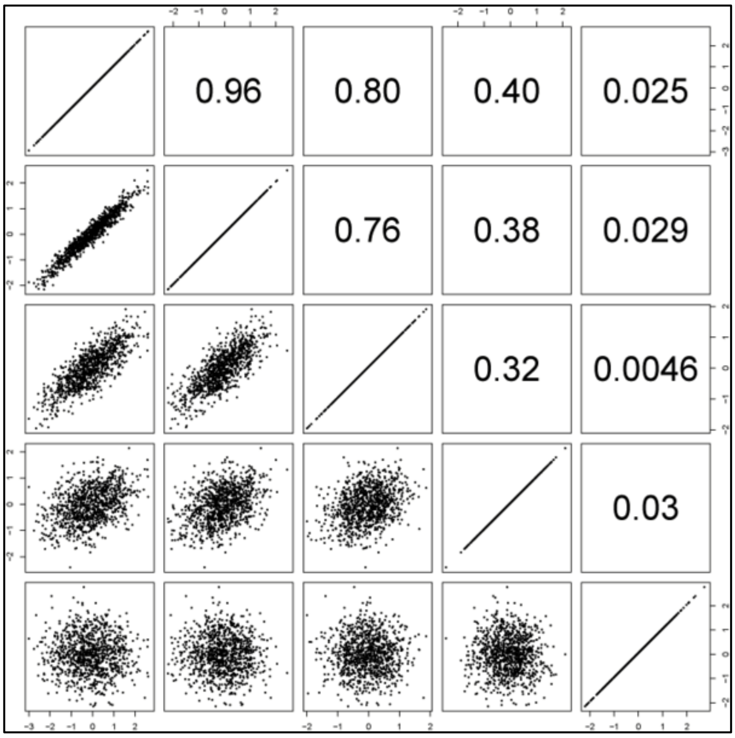


While the BIC (and AIC) are useful measures from a statistical perspective, engineers generally want to consider the effects of variables, even if these variables don't have a major effect on crashes. Hence from an engineer's perspective, the models should include all the key variables that are of interest in design and planning of the road, except in cases where the variables are highly correlated.

Where variables are correlated, adding both variables into the model can lead to a poor description/prediction of each of the variables. A choice needs to be made on whether to exclude one of the two variables, or to make an adjustment on how the variables are added to the model (ie by orthoganalizing the correlated variables).

The correlation between different variables can be determined by examining the correlation matrix, which is a matrix of correlation coefficients between the variables used for modelling. Correlation coefficients indicate the strength and direction of a linear relationship between two random variables, where a value of one indicates a perfect positive correlation between two variables, and a value of zero indicates statistical independence. Figure 7.2 illustrates an example of different values of linear correlation.

Figure 7.2 Examples of linear correlation



8 Modelling results

This section describes the preliminary models that were developed for this pilot study. Given the small sample set, these models **should not** be used for crash prediction – they were developed to give an indication of the more important predictor variables to focus on in the main study in the third phase.

8.1 Volume-only model

Traffic volume, V, defined as AADT, was included as a default variable in all the crash models. This was because a significant body of research over the last 20 years, including work by the authors of this research, has shown that traffic volume is by far the most important variable to include in crash prediction models, be they for rural roads or other road elements.

The first and only one-variable model developed in this study was therefore of the form $A = aV^{b1}$, where V is the traffic volume. The results are given in table 8.1.

Table 8.1 One-variable crash model – volume

Single variable (X_1)	a	b1	k	LL
Volume	0.003744	0.700435	2.5	-381.283

8.2 Two-variable models

This stage of the analysis required searching for the other important variables to include along with traffic volume. To achieve this, each of the other variables was added to a model along with traffic volume to develop a set of two-variable models. The models with the highest (least negative) log-likelihood (LL) indicated that the variable was important.

Secondary variables relating to the road environment were included alongside traffic volume, V. Models developed here were of the form $A = aV^{b1}X_2^{b2}$. Tables 8.2 to 8.6 show the results of the two-variable models according to the general category of the second variable.

Table 8.2 Two-variable crash models – shoulder environment

Second variable (X_2)	a	b1	b2	k	LL
Unsealed shoulder width	0.003782	0.7026	-0.0174	2.5	-381.079
Traversable slope width	0.005647	0.7092	-0.0582	2.5	-381.131
Recoverable slope width	0.003978	0.7010	-0.0092	2.5	-381.259
Seal width	0.007054	0.7097	-0.0773	2.5	-381.274

None of these variables showed a significant increase in the LL. As noted previously, a number of these variables were correlated with other variables. Of particular note was the correlation between seal width and traffic volume.

Table 8.3 Two-variable crash models – point hazards

Second variable (X_2)	a	b1	b2	k	LL
High-impact roadside furniture	0.005519	0.6474	0.3327	3	-377.378
Culvert – road with non-traversable headwall	0.002976	0.7210	0.0609	2.7	-380.331
Wooden pole	0.003843	0.7023	-0.0303	2.5	-380.923
Sign support	0.003897	0.6958	-0.1534	2.5	-381.124
Concrete pole	0.003425	0.7141	-0.0233	2.5	-381.129
Culvert – roadside	0.003781	0.6971	0.0178	2.5	-381.203
Combined point hazards	0.003769	0.7024	-0.0082	2.5	-381.248
Tree	0.003671	0.7037	-0.0096	2.5	-381.261
Light column	*Insufficient convergence*				
Heavy street pole	*Insufficient convergence*				

The high-impact roadside furniture variable had the greatest impact on the LL, despite very few occurrences in the sample set. While 'combined point hazards' did not appear to have a significant impact on crashes, it was considered the best variable from this set, given the limited number of hazards of some types in the sample set.

Table 8.4 Two-variable crash models – accesses

Second variable (X_2)	a	b1	b2	k	LL
Farm/residential	0.004035	0.7111	-0.0813	2.7	-378.688
Combined	0.004225	0.7051	-0.0786	2.7	-378.834
Commercial	0.004147	0.6893	-0.1489	2.6	-380.668

Both the 'farm/residential' and 'combined' variables appeared to reduce the LL and were important variables.

Table 8.5 Two-variable crash models – roadside – other

Second variable (X_2)	a	b1	b2	k	LL
Distance to non-traversable slope/ perpendicular deep drain	0.001953	0.7322	0.0668	2.7	-377.487
End concrete barrier/bridge parallel to road	*Insufficient convergence*				
Concrete fence/barrier perpendicular to road	*Insufficient convergence*				

There were limited observations of these variables in the small sample set. The 'distance to non-traversable slope/perpendicular deep drain' was shown to be important, although this variable was found to be correlated with a number of other roadside environment variables. This again supports the development of an overall roadside classification, as provided in the KiwiRAP scoring.

Table 8.6 Two-variable crash models – road geometry

Second variable (X_2)	a	b1	b2	k	LL
Absolute gradient	0.003134	0.7151	-0.2670	3.1	-371.292
Absolute curvature	0.002542	0.6993	0.0498	2.5	-380.6
Absolute crossfall	0.004277	0.6925	-0.0789	2.5	-380.776
Skid resistance	0.001415	0.6220	-2.3890	2.8	-376.895

Both ‘absolute gradient’ and ‘skid resistance’ were shown to be important variables. ‘Curvature’ was shown to be less important than these other two variables. This may have been a case of out-of-context curves being important, rather than the actual curvature being of less or greater importance (see section 2.2). The issue of alignment consistency will be included as a variable in the main study.

Some variables sets contained a large number of zeros, as shown in figures 3.1–3.10 in section 3. These variables, while potentially generating a plausible best-fit function, depending on the initial parameter estimates, were not numerically stable in their convergence dynamics. They were therefore discarded from further analysis in the pilot study. However, it is important to note that these variables may be considered important in the main study, which will have significantly more data available.

The combining of variables (eg roadside hazards and accesses) did not appear to produce an improved model fit. However, this may have been because of the small size of the data set.

The variables standing out as yielding the significantly highest LL values were ‘high-impact roadside furniture’, ‘distance to non-traversable slope/perpendicular deep drain’, ‘farm/residential accesses’ and ‘absolute gradient’.

It is also worth noting that several variables had a negative exponent, which indicated that they contributed towards reducing the crash rate. This result may have reflected the change in driver behaviour arising from noticing and responding to these features of the road environment. Alternatively, it could have been a reflection of the small size of the data set. Another possibility was that the exponent of the variable in question may have actually reflected the effect of a dependence on a separate variable that may not have been measured. This issue will receive considerable attention in the main study stage of this research, as we attempt to build a more robust rural road model for widespread use in New Zealand.

8.3 Multi-variable models

Table 8.7 lists the exhaustive list of variables that were selected for developing multi-variable models (ie three or more variables), along with their type.

Table 8.7 Variables in multi-variable models

Type	Variable
Traffic volume	Traffic volume (V)
Shoulder environment	Unsealed shoulder (U)
	Seal width (S)
Point hazards	Combined point hazards (H)
Accesses	Combined accesses (Ca)
Roadside – other	Distance to non-traversable slope/perpendicular deep drain (N)
Road geometry	Absolute curvature (C)
	Absolute gradient (G)
	Skid resistance (Sr)
	% reduction in curve speed (Vc)[1]

1 The results of the two-variable models indicated that horizontal geometry was not a significant variable. However, horizontal geometry is widely considered to have a significant effect on crash rate. A horizontal alignment variable (percentage reduction in curve-negotiation speed of the section compared with the preceding 500m section) was thus later introduced into the best-performing four-, five- and six-variable models, as described later in this section.

This list contains variables that yielded a high LL, or were deemed to be important variables from other studies and/or from a road-design perspective. For example, although 'high-impact roadside furniture' was observed to have a higher LL, 'combined point hazards' was instead included in the final list of variables to be used for multi-variable modelling, since this provided a more holistic representation of the hazards actually present in the roadside environment. Similarly, 'combined accesses' was chosen in place of 'farm/residential accesses', which had a marginally higher value of LL.

All 247 combinations of these variables that include 'traffic volume' (V) (excluding horizontal alignment) were tested. Models were of the form $A = aV^{b1} X_2^{\ b2} ... X_m^{\ bm}$.

The preferred models, that is, the ones yielding the highest LL for a given number of variables, are given in table 8.8. The variable coding/naming convention is specified in table 8.7.

Table 8.8 Multi-variable models

Quantity	1-var	2-var	3-var	4-var	5-var	6-var	7-var	8-var	9-var
Multiplier a	0.003744	0.003134	0.001492	0.000565	0.000715	0.000679	0.000669	0.000165	0.000162
V exp	0.700435	0.7151	0.7520	0.6780	0.6789	0.6709	0.6732	0.6512	0.6513
S exp	0	0	0	0	0	0	0	0.1709	0.1657
U exp	0	0	0	0	0	0	-0.0113	-0.0126	-0.0121
H exp	0	0	0	0	0	0.0613	0.0606	0.0602	0.0604
Ca exp	0	0	0	0	-0.0641	-0.0916	-0.0886	-0.0888	-0.0891
N exp	0	0	0.0747	0.0768	0.0720	0.0848	0.0844	0.0841	0.0839
G exp	0	-0.2670	-0.2826	-0.2756	-0.2789	-0.2877	-0.2882	-0.2888	-0.2863
C exp	0	0	0	0	0	0	0	0.0033	0.0082
Sr exp	0	0	0	-2.3159	-2.1879	-2.0727	-2.0920	-2.1161	-2.1238
k	2.5	3.1	3.6	4	4.4	5.2	5.2	5.2	5.2
LL	-381.283	-371.292	-366.276	-361.945	-360.194	-358.543	-358.448	-358.401	-358.381
N	296	296	296	296	296	296	296	296	296
P	2	3	4	5	6	7	8	9	10
BIC	2.615	2.566	2.552	2.542	**2.549**	2.557	2.576	2.595	2.614
AIC	2.590	2.529	2.502	2.479	**2.474**	2.470	2.476	2.482	2.489

The resultant BIC and AIC measures (see table 8.8) indicated that the models with 4, 5, and 6 variables were preferred. Table 8.9 below lists the variables comprising the four-variable, five-variable, and six-variable models.

Table 8.9 Predictor variables in 4-, 5- and 6-variable models (excluding horizontal alignment)

4-variable model	5-variable model	6-variable model
Traffic volume (V)	Traffic volume (V)	Traffic volume (V)
Distance to non-traversable slope/perpendicular deep drain (N)	Distance to non-traversable slope/perpendicular deep drain (N)	Distance to non-traversable slope/perpendicular deep drain (N)
Skid resistance (Sr)	Skid resistance (Sr)	Skid resistance (Sr)
Absolute gradient (G)	Absolute gradient (G)	Absolute gradient (G)
	Combined accesses (Ca)	Combined accesses (Ca)
		Combined point hazards (H)

The four-variable model was found to have the lowest BIC measure. However, this model excluded the important predictor variables of combined point hazards and combined accesses. On the other hand, the six-variable model was found to have the lowest AIC measure, although this model included both the predictor variables of combined point hazards and combined accesses. These two variables were found to have a high correlation coefficient of 0.358 (refer to table 5.2), and it was therefore not deemed suitable to include both variables in the same model**.** None of the above models contained a road geometry/horizontal alignment variable.

Figure 8.1 shows the correlation between combined point hazards (H) and combined accesses (Ca). The correlation between combined point hazards (H) and distance to non-traversable slope/perpendicular deep drain (N) is depicted in figure 8.2.

Figure 8.1 **Correlation between combined point hazards and combined accesses**

Combined accesses

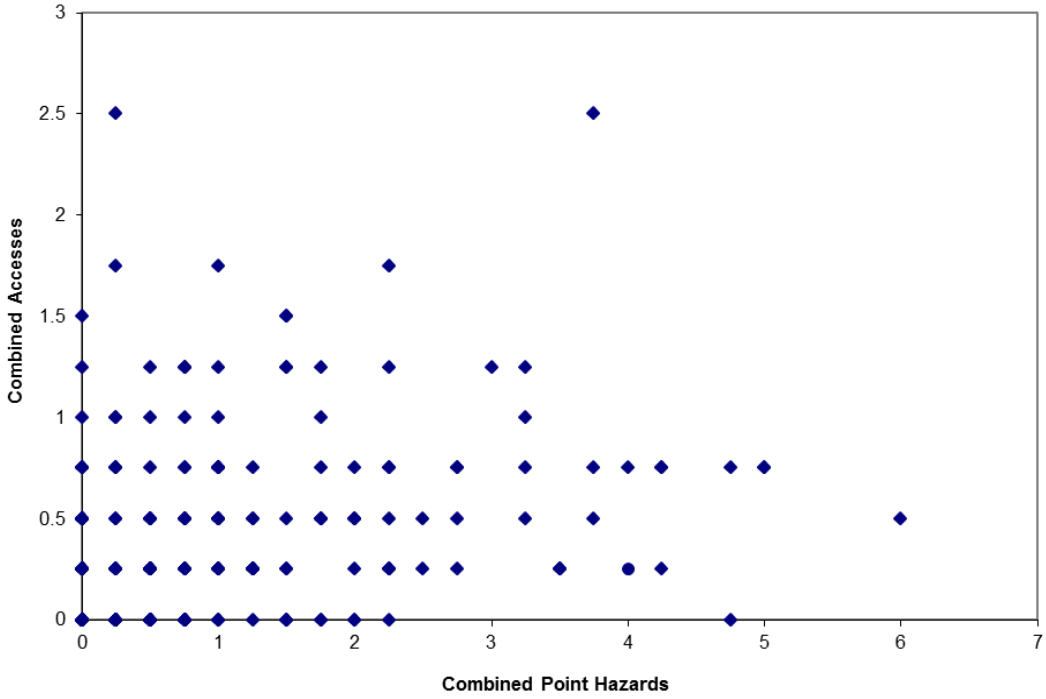
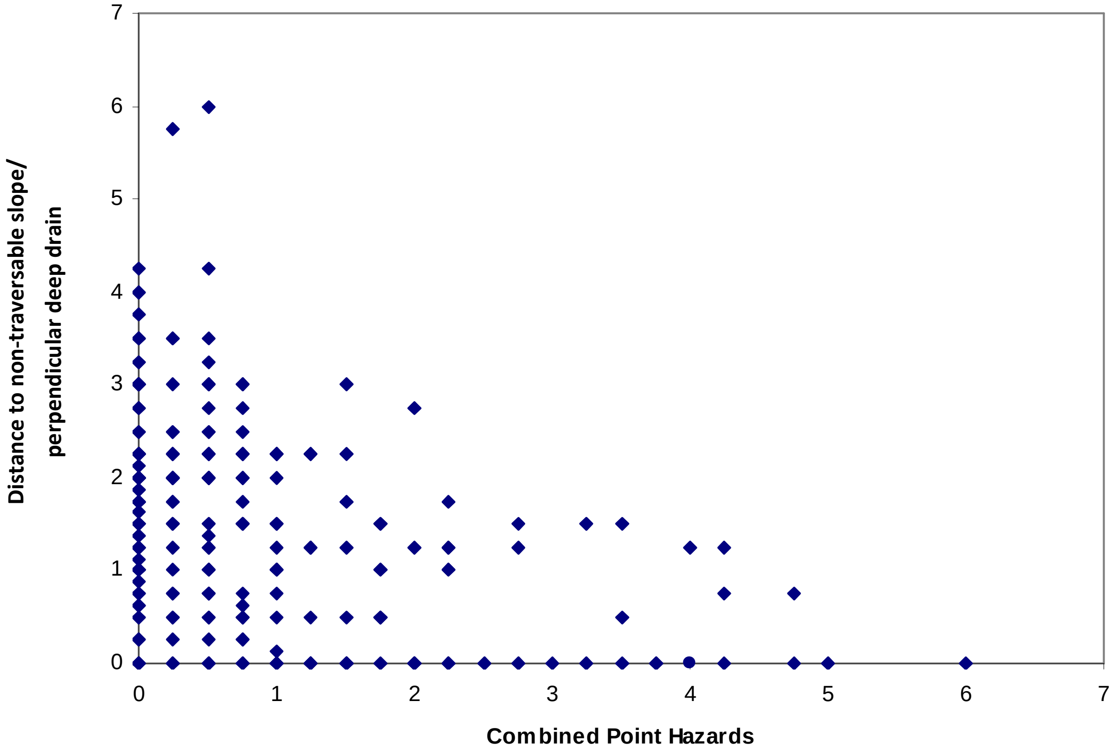


Figure 8.2 **Combined point hazards and distance to non-traversable slope/perpendicular deep drain**



As noted earlier, the best-performing four-, five- and six-variable models shown in table 8.8 do not contain a road geometry variable, which is considered to be important from a crash-risk perspective. The horizontal alignment variable (percentage reduction in curve speed, Vc) was thus added to the above best-performing models, so that horizontal geometry could be given consideration in the final model form. The BIC and AIC figures were re-assessed after the addition of the horizontal alignment variable to determine the preferred final model.

Subsequent to the addition of the horizontal alignment variable, the models that were found to perform the best are shown in table 8.10 below, along with their BIC and AIC values.

Table 8.10 Variables in multi-variable models (after the addition of horizontal alignment variable)

	4-variable model	5-variable model	6-variable model
Variables included	Traffic volume (V)	Traffic volume (V)	Traffic volume (V)
	Distance to non-traversable slope/perpendicular deep drain (N)	Distance to non-traversable slope/perpendicular deep drain (N)	Distance to non-traversable slope/perpendicular deep drain (N)
	Skid resistance (Sr)	Skid resistance (Sr)	Skid resistance (Sr)
	Absolute gradient (G)	Absolute gradient (G)	Absolute gradient (G)
		% reduction in curve speed (Vc)	Combined accesses (Ca)
			% reduction in curve speed (Vc)
BIC	2.54	**2.53**	2.56
AIC	2.48	**2.45**	2.47

8.4 Preferred model

The five-variable model was found to have the lowest BIC and AIC values, and also included a horizontal alignment variable. It was thus considered to be a better model than the four-variable or six-variable models.

The preferred preliminary model has the following form:

$$A = 2.2E^{-04}\ V^{0.719} \times N^{0.078} \times G^{-0.26} \times Sr^{-2.569} \times Vc^{0.219} \qquad \text{(Equation 8.1)}$$

where:

A is the predicted number of crashes in five years for a 100m section of rural road

V is the two-way AADT for the road section

N is the distance in metres to the non-traversable hazard (eg row of trees or deep ditch), multiplied by 1000

G is the absolute gradient

Sr is the average value of SCRIM for the road section

Vc is the percentage reduction in the curve-negotiation speed of the section as compared with the preceding 500m section.

As specified earlier, this model is based on a relatively small sample size and is therefore not suitable as a prediction model for general use. This is also reflected in the values of the exponents of certain variables, such as that of absolute gradient (G), which appear to be counter-intuitive. It does, however, give some indication of the most important variables that affect crash occurrence on rural roads, at least for roads in the Waikato. The larger sample set to be used in the main study will lead to a more accurate representation of the effects of the important variables, and thus to the production of a much more robust model that can be used for general crash prediction by practitioners.

8.5 Sample size for stage 3 main study

At the time of writing, the sample size for the third stage (main study) of this research had not yet been confirmed. While KiwiRAP video data is available for the entire two-lane rural state highway network, data regarding a small number of variables, such as accessway densities, which are not measured by KiwiRAP, may need to be collected manually in the field.

The accuracy of the KiwiRAP roadside hazard rating will also be assessed as part of the main study stage of this research. While efforts will be made to undertake this assessment using data already collected during this pilot stage of the study, plus any data available from past research undertaken by Beca, there may still be a requirement to collect additional data in the field.

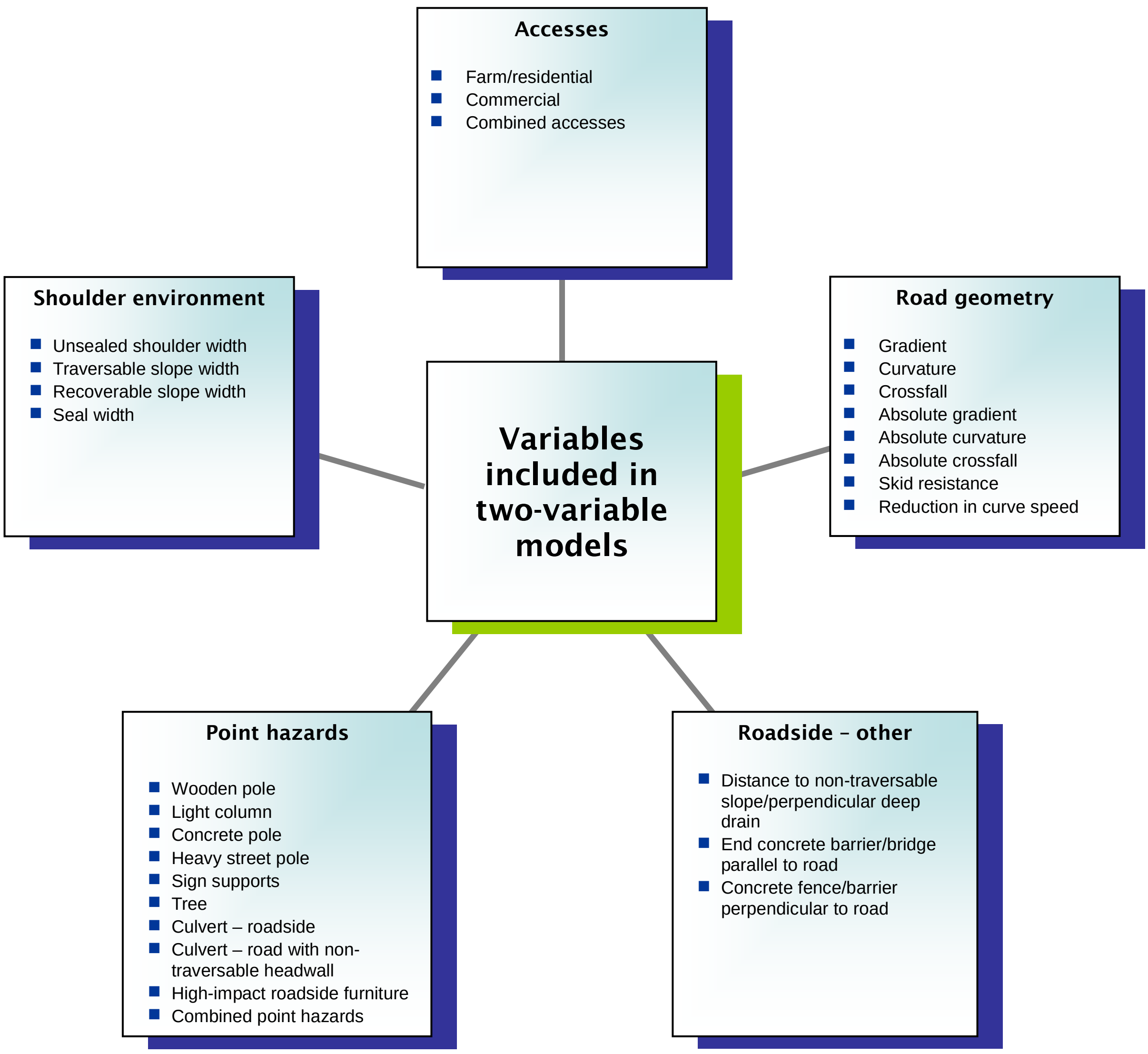
A sample-size assessment will thus need to be undertaken as part of the main study stage of this research project.

9 Summary and conclusions

Data on road alignment, roadside environment and traffic flow has been collected for 200 x 400m rural road sections in the Waikato region of New Zealand. Using this data along with crash data, multi-variable cross-section models can be developed to predict the crash rate. This pilot study has developed preliminary crash prediction models for rural roads using generalised linear regression model techniques developed by Beca. Data from 148 of the 200 sample sections was used for building the models. Full models that can be used for crash prediction will be developed in stage 3 of this research project (main study).

In addition to traffic volume, 28 other predictor variables (divided into five categories) were included in the two-variable models. These are listed in figure 9.1 below.

Figure 9.1 Variables included in two-variable models



From those 28 variables, a most-representative and best-performing set of eight was selected to be incorporated into the multi-variable models alongside traffic volume, based on the results of the two-variable models. The variables that were found to perform the best were:

- unsealed shoulder width
- seal width
- combined point hazards
- combined accesses
- distance to non-traversable slope/perpendicular deep drain
- absolute gradient
- absolute curvature
- SCRIM coefficient.

In addition, a horizontal alignment variable (percentage reduction in curve speed) was added to the best-performing variables in the above list, in order to develop a more robust model.

Out of the nine variables listed above, the overall preferred model was found to involve volume, distance to non-traversable hazard, absolute gradient, SCRIM coefficient, and percentage reduction in curve speed. The preferred model had the following form:

$$A = 2.2E^{-04}\ V^{0.719} \times N^{0.078} \times G^{-0.26} \times Sr^{-2.569} \times Vc^{0.219} \qquad \text{(Equation 8.1)}$$

where:

A is the predicted number of crashes in five years for a 100m section of rural road

V is the two-way AADT for the road section

N is the distance in metres to the non-traversable hazard (eg row of trees or deep ditch), multiplied by 1000

G is the absolute gradient

Sr is the average value of SCRIM for the road section

Vc is the percentage reduction in the curve-negotiation speed of the section as compared with the preceding 500m section.

From the above model form, it is noted that the value of 0.719 on the exponent of traffic volume (V) implies that the relationship between traffic volume (AADT) and number of crashes on a section of rural road shows a significant deviation from linear behaviour; ie the crash rate does not increase at a rate proportional to the rate of increase of traffic volume.

It must be noted that this model is based on a relatively small sample size and is not suitable as a prediction model for general use.

The model results were evidence of the important influence that the roadside environment and road geometry have on predicted crash rate. However, the exponents (ie the numerical powers to which model variables are raised) with which most secondary variables (except SCRIM coefficient) contributed to the model were small, indicating a subtle influence only. The SCRIM coefficient was observed to have an exponent equal to -2.569, which implied that an improvement in the skid resistance of the road surface would be likely to have a significant impact on the crash risk of a given section of rural road.

No doubt many of the results were sensitive to the small size of the data set. As the pilot aspect of the wider study, this research has identified the key variables and data collection issues that will be critical to progressing the research project in stage 3 (main study).

10 Recommendations for the main study

The sub-sections below outline key recommendations for data collection, data analysis and crash model development during the main study, to build on the results of this pilot study.

10.1 Use of 'element' lengths rather than 'section' lengths

Analysis conducted to assess the lengths of individual elements (straights or curves) within 400m sections included in the sample set for this study revealed that 45% of the sections analysed were located on mostly straight or mostly curved sections of highway. Of the remaining 55%, more than half consisted of elements that had a length of more than 200m. The remaining sections consisted of straight or curved elements that were less than 200m in length. A combination of straight/curved elements was also observed in a number of individual 100m lengths.

This apparent variation in the lengths of individual elements suggests that for the purpose of data collection and analysis during the main study, it may be beneficial to split the state highway network into elements of varying lengths, instead of using homogeneous sections of fixed lengths as has been done in this pilot stage of the study. Research undertaken by Cenek et al (2004) led to the development of a process for splitting the state highway network into element lengths, and this process should be used for identifying sections according to element type during this project's main study stage.

However, inaccuracies in the recording of crash data may mean that there may be an error of more than 100m in the crash locations. It is therefore proposed that both homogeneous lengths and road elements be looked at during the main study.

10.2 Estimation of accessway densities

The study team found it beneficial to use the combined accesses variable to estimate the total number of accesses (farm/residential and commercial) on a section of rural road. In addition, the frequency of accesses on the sections included in the sample set for this study was observed to be quite low, with no accesses present on 44% of sections.

This suggests that for the purpose of the main study, it may be viable to use a generic figure of 0.5 accesses per 100m for the majority of sections, instead of undertaking field data collection to measure the density of accesses on the entire state highway network. This is considered to be a reasonable estimate in the absence of existing data on access densities. However, field data collection may still need to be undertaken to estimate access densities on state highways in areas that are known to have a higher density of accessways, and ideally also for remote areas that have a very low density of accesses, such as the West Coast region.

10.3 Estimation of roadside environment: use of KiwiRAP roadside hazard rating

A high degree of correlation was observed between several variables describing the roadside and shoulder environments. Of particular interest were the correlations between 'traversable slope width' and 'combined point hazards' (0.54), and 'traversable slope width' and 'wooden poles' (0.53). Significant correlations were also observed between 'combined point hazards' and variables representing individual point hazards, such as concrete poles, wooden poles, culverts and trees. Additionally, while some point hazards such as wooden poles, culverts, concrete poles and trees were present in significant numbers in the sections included in the sample set, others such as light columns, heavy street poles and sign supports were not commonly found on the sections surveyed. Overall, 38% of the sites had no point hazards, while another 38% had between 0 and 1 point hazard within an average 100m section.

The high correlations mentioned above confirm the suitability of using an integrated hazard-scoring system, such as the KiwiRAP roadside hazard rating, as opposed to looking at each element of the roadside environment and individual types of point hazards separately. The roadside hazard rating system employed by KiwiRAP provides a reasonably detailed rating that takes into account the quality of the roadside environment and the severity of hazards at varying distances along the roadside. It is intended that the KiwiRAP system will be used for the large sample set available in the main study phase of this research. Any concerns with the accuracy of the KiwiRAP roadside hazard rating can be investigated during the main study by comparing the raw KiwiRAP rated score with a similar rating using the hazard data that was collected manually during this pilot stage of the research, as well as any data available from past research undertaken by Beca.

10.4 Separate models for AADT bands

Traffic volume (AADT) and seal width were found to have a high correlation coefficient of 0.52. Previous research has shown that both AADT and seal width are important predictor variables and it is therefore important to include both of them in the crash modelling. One approach to overcome the limitation of the high correlation between AADT and seal width is to define fixed AADT bands and build individual crash models for each of these bands. Such a model form would neutralise the effects of the correlation between seal width and AADT, and would also better reflect the safety impact of changes in seal width, such as when seeking to identify the benefits of seal widening.

10.5 Sample size for the main study

At the time of writing, a sample size for the main study stage of this research had not yet been confirmed – the sample size assessment will be an integral part of that stage.

KiwiRAP video data is available for the entire two-lane rural state highway network in New Zealand. However, data on additional variables that are important for this research (eg density of accessways), which are not measured by KiwiRAP, may still need to be collected manually.

11 References

ARRB data collection services. Accessed 15 December 2009. *www.arrb.com.au/Equipment-services/Data-collection-services.as*px

Cairney, P and P Bennett (2008) Relationship between road surface characteristics and crashes on Victorian rural roads. *23rd ARRB Conference*, Adelaide, Australia.

Cenek, PD and RB Davies (2008) Crash risk relationships for improved safety management of roads. *Transfund NZ research report PR3-0709*.

Cenek, PD, RB Davies and RJ Henderson (2004) The effect of skid resistance and texture on crash risk. In *Proceedings Sustainable Transport Conference*, Wellington, New Zealand.

Chadfield, E (1993) *Review of cross-section guidelines for two-lane rural roads*. Wellington: Transit NZ. 32pp.

Hauer, E and S Hakkert (1989) The extent and some implications of incomplete crash reporting. *Transportation Research Record 1185*: 1–10.

Jackett, M (1992) *Accident prediction: rural curves.* Wellington: Land Transport NZ. 3pp.

Koorey, GF (2009) Incorporating safety into rural highway design. PhD thesis, Department of Civil and Natural Resources Engineering, University of Canterbury, New Zealand.

Koorey, G F and FN Tate (1997) Review of accident analysis procedures for project evaluation manual. *Transfund research report 85*. 56pp.

Maycock G and RD Hall (1984) Crashes at four-arm roundabouts. *TRRL laboratory report LR1120*.

McLarin et al (1993) *Impact of terrain carriageway characteristics*. New Zealand: Transit NZ.

Transit New Zealand (2008) *State highway traffic data booklet (2003–2007).* New Zealand: Transit NZ. 47pp.

Prinsloo, B and C Goudanas (2003) Development of a crash prediction model for rural roads in NSW. *21st ARRB Transport Research Conference*, Queensland, Australia.

Thomsen and O'Brien (2003) The role of digital imaging in road asset management: a Blue Mountain City Council experience. In *Proceedings of the 21st ARRB and 11th REAAA Conference*, Queensland, Australia.

Turner, S (2000) Accident prediction models. *Transfund research report 192*. 78pp.

Turner, S and F Tate (2009) Relationship between road geometry, observed accident speed and rural accidents. *NZ Transport Agency research report 371.* 72pp.

Turner, S, A Dixon and G Wood (2004) Assessing the crash risk implications of roadside hazards. *Traffic Management Workshop: Technical Conference*, Wellington, New Zealand. 15pp.

Turner, S, F Tate and A Roozenberg (2005) Roadside hazard management – stage 1. Land Transport NZ research report. 39pp.

Turner, SA, AP Roozenburg, F Tate, and GR Wood (2006) Rural accident prediction model – a new generation – stage 1 (scoping) report. Unpublished Land Transport NZ research report. Wellington: Land Transport NZ.

Appendix: Survey work instruction

A1 Scope of work

This work involves the field survey for the research study 'Rural crash prediction model' for Land Transport New Zealand.

A2 General

The survey involves two surveyors walking and/or driving along selected segments of rural road and recording the number of accessways, the shoulder environment, and any roadside hazards – each side of the road should be observed in a separate drive/walk.

- Driving – one person drives the survey vehicle while the other observes and records the data required.
- Walking – one person observes and records the required data while the other keeps lookout for traffic.

All data collected should be clear and legible. The following RAMM database information will be provided on pre-printed data collection forms, and should be transcribed onto blank sheets if more are needed:

- road number
- road name
- distance reference to side roads, bridges or culverts
- section start displacement
- section end displacement
- authority.

A3 Equipment

The following equipment will be provided and is to be used while undertaking these surveys:

- roof-mounted revolving amber light
- rear- or roof-mounted 'Road Inspection' sign
- terra-trip odometer
- calibration record sheets
- clipboard and data entry sheets

- two glow jackets
- measurement wheel
- ‘dazzle’ spray paint
- writing equipment.

A4 Safety

Temporary traffic management is required for all road-inspection activities on the state highway. An approved traffic management plan must be carried when undertaking these surveys.

The general principle is that the driver of the survey vehicle must not expect traffic to slow down for the survey vehicle. The survey vehicle should travel at the speed of the rest of the traffic, in most instances. When the vehicle slows down it should travel on the shoulder of the road where practical. Where the shoulder is narrow, the vehicle should travel as far as practicable from the ‘live’ traffic lane.

A roof-mounted revolving amber beacon should be fitted on the road side of the survey vehicle, and should operate throughout the survey. A TW27 ‘Road Inspection’ sign should be mounted on either the tow bar or the roof. Where the road visibility is less than 300m, the vehicle should be parked clear of the roadway and the data collection completed on foot, using a lookout person.

As the seal width will be measured and the surveyors are required to walk into the ‘live’ traffic lane, one of the two personnel must be a Site Traffic Management Supervisor (STMS).

Additionally, company personnel must also comply with all their company’s health and safety procedures.

A5 Data collection procedure

1. Locate the start of the section to be surveyed by measuring the distance from the landmark specified, using the trip meter. It is important that the measurement is performed in the right direction along the road (either ‘increasing’ or ‘decreasing’). Mark this position by spraying the dazzle paint on the berm. Then mark the middle of the section (200m from the start) and the end of it, so that the start, middle and end points can be located on subsequent runs without the further use of the trip meter.
2. Assess the shoulder environment for each of the four 100m segments **on each side** of the road, either by driving the length of the section, or on foot with a lookout person. If you need to perform a U-turn, choose a location that has clear sight distance. It may be practical to undertake the assessment of roadside hazards simultaneously – otherwise, do this after the assessment of the shoulder environment.
3. Assess the accesses by a drive-over if possible – as before, assess four segments on either side of the road, every 100m.
4. Use the measuring wheel to measure the typical seal width at a location in each 100m section. This measurement should be done where there is clear sight distance in both directions, with one person acting as a lookout. If the STMS deems that the sight distance or traffic volumes prohibit

this measurement, it can be obtained in the Beca office from aerial photos. In this situation, the reason behind the decision must be specified.

A6 Information to be collected

A6.1 Shoulder environment

When examining the shoulder environment, the area up to 9m from the edge of seal should be considered. For each 400m section, a total of eight shoulder environments should be assessed – ie 4 x 100m segments on each side of the road.

In these 100m segments, select the environment that is applicable to the majority of the segment (greater than 50%) and record the 'typical' width.

For each of the four segments, record the following width estimates:

- **Unsealed shoulder width:** ie the width of the unsealed shoulder from the edge of seal, estimated as 0, 0.5, 1, 2, 3, or greater than 3m
- **Recoverable slope width:** ie the width from the edge of seal (including unsealed shoulder) where the slope is flatter than, or equal to, about 1:6 and clear of any continuous severe hazards (eg drains or rows of trees), and would allow vehicles to recover and then re-enter the roadway. If a continuous severe hazard is present on a recoverable slope, then the recoverable width is the width between the edge of seal and the hazard (for definitions of 'severe hazards' see the next section). Estimate the recoverable width as 0, 0.5, 1, 2, 3, or greater than 3m.
- **Traversable slope width:** ie the width from the edge of the seal (including unsealed shoulder and recoverable slope) where the slope is flatter than, or equal to, about 1:3, and clear of any continuous severe hazards (eg drains or rows of trees). If a continuous severe hazard is present on a traversable slope, then the traversable width is the width between the edge of the seal and the hazard (for definitions of 'severe hazards' see the next section). Estimate the traversable width to the nearest metre up to 9m from the edge of seal.

It may be difficult at times to distinguish between recoverable slopes and unsealed shoulders, and recoverable slopes and traversable slopes. Unclear cases should be noted as such. Figures A6.1, A6.2 and A6.3 provide examples of slopes.

Figure A6.1 Shoulder environment with unsealed shoulder and traversable slope > 9 m

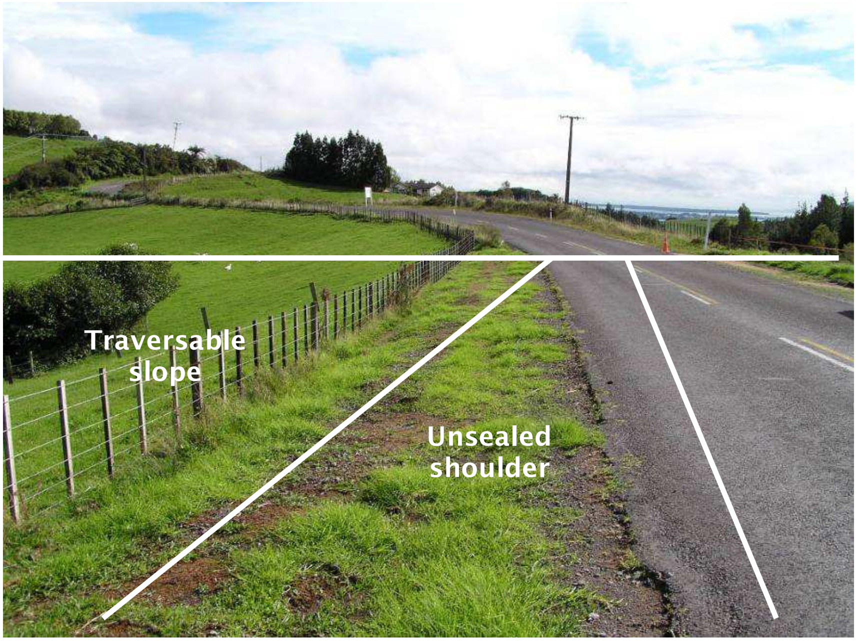
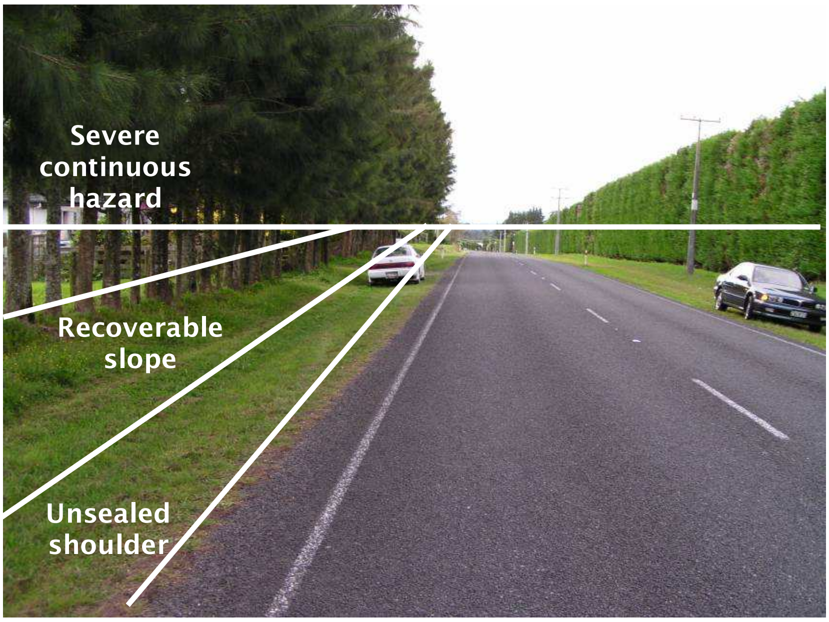


Figure A6.2 Shoulder environment without unsealed shoulder traversable slope limited by severe continuous hazard

Figure A6.3 Shoulder environment with unsealed shoulder and recoverable and traversable slopes limited by severe continuous hazard



A6.2 Roadside hazards

For this study, a limited hazard inventory is required, recording only a selection of roadside hazard data.

As the traversable slope measured when assessing the shoulder environment represents the distance to severe continuous hazards, only point hazards within this distance should be measured (out to a maximum of 9m) – ie if the traversable width is 4m, then only hazards within 4m of the edge of the seal should be included.

Record only the following severe point hazards:

- wooden utility pole with diameter >200mm (includes most wooden poles)
- steel lighting column – ground-planted with diameter <300mm (includes most hollow-section poles)
- solid concrete service pole – usually ‘I’ section and strongly reinforced
- heavy-weight steel pole without a slip base and with diameter >300mm
- tree with trunk >100mm diameter
- culvert for traversing side drain, or culvert under the roadway, with non-traversable headwall
- sign with supports with diameter >120mm and without a slip base
- non-traversable slope or deep drain perpendicular to the roadway
- concrete fence or barrier perpendicular to the roadway

- end of concrete barrier or bridge parapet parallel to the roadway
- high-impact roadside furniture (eg roadside transformer box).

Count the number of severe point hazards within each 100m segment on either side of the road – ie eight counts of severe point hazards for each section.

Where point hazards are closely grouped (such as rows of trees), define them as a severe continuous hazard and rather than count them, consider them in the shoulder environment classification as the point where the recoverable and/or traversable slope ends (see figures A6.3 and A6.4).

Figure A6.4 Example of severe point hazard data collection

A6.3 Access density

Collect the access density in a drive-by of the 400m road sections (divided into eight segments, ie four segments of equal length on either side of the road).

The following two different classes of access are to be considered in this study:

- **Farm and residential accesses:** Simply count these. Where accesses are shared, count the number of adjoining properties; eg one access shared with two properties should be counted as two accesses, and so on. Counting the mailboxes at an access should be an adequate method.
- **Commercial accesses:** Estimate the activity eg:
  - dairy/cafe or other roadside store
  - dairy farm
  - greenhouse
  - garden centre

  - sawmill
  - winery
  - cattery or kennel
  - quarry
  - stable
  - fruit stand
  - chicken/pig farm
  - school
  - lodge/bed and breakfast
  - gas station
  - motel/hotel.

For each access, class the level of activity as follows:

- low (approximately 1–30 vehicle movements per day)
- moderate (approximately 30–100 vehicle movements per day)
- high (approximately 100+ vehicle movements per day).

For example, a typical residential property in a rural area will generate 6–10 trips per day.

Note also the condition of the access. For example, long grass obstructing an access may indicate infrequent use and therefore it should not be included in the survey.

A6.4 Seal width

Seal width measurements should be taken at four locations along the 400m section, one in each 100m section. The exact locations should be chosen as points with 'typical' seal widths that would be representative of the 100m segment, and should also have safe clear distance to allow the surveyor to enter the 'live' traffic lane. If no safe location exists, this must be noted and the width obtained in the Beca office from Transit Aerial Strip Maps.

Use the measurement wheel to measure the total width from the edge of the seal on one side of the road to the edge of the seal on the opposite side. Also record the sealed shoulder width and lane width on both sides of the road.

While one surveyor is measuring the width, the other must act as a lookout for approaching vehicles.

A6.5 Photographs

Use a camera with a resolution of no less than 3.0 megapixels to take a photo of each 100m segment in each direction – ie a total of eight photographs for each 400m section. Take these photos in the following order (to enable later identification):

- the first 100m section looking in the 'increasing' direction
- the second 100m section looking in the 'increasing' direction
- the third 100m section looking in the 'increasing' direction
- the fourth 100m section looking in the 'increasing' direction
- the fourth 100m section looking in the 'decreasing' direction
- the third 100m section looking in the 'decreasing' direction
- the second 100m section looking in the 'decreasing' direction
- the first 100m section looking in the 'decreasing' direction.