# Safety Evaluations using a Real-Time Crash Potential Model: Sensitivity to Model Calibration

By
Bruce Hellinga, Amir Samimi

Proceedings of the ITE Canadian District Annual Conference
held May 6-10, 2007 in Toronto, Canada.

ABSTRACT:

Traditional safety analysis focuses on determining locations that have experienced higher than expected number of collisions (termed "black spots") and then identifying actions, typically geometric improvements (e.g. centre median barrier, etc.) that can improve safety. More recently, attention has been given to identifying, in real-time, traffic conditions that are associated with high crash potential, and when such conditions are observed, to intervene using various freeway traffic management options (e.g. ramp metering, variable speed limits, etc.). These models are often referred to as "real-time crash prediction models".

Real-time crash prediction models are often structured as general log-linear categorical models which must be calibrated using an extensive database containing roadway geometric characteristics, weather conditions, crash records, and real-time traffic conditions (typically determined from loop detector data). The task of calibrating a crash prediction model requires considerable effort and expense. Furthermore, the calibration process does not necessarily result in the optimal set of model parameters. This raises the question of how important the calibration is to the safety impacts estimated when using the crash potential model.

In this paper, we examine the impact that the process used to calibrate the crash potential model has on estimates of safety impacts of a variable speed limit system. Two calibration methods are compared, namely (1) heuristic ad hoc method, and (2) near-optimal method. Both methods are applied to a crash potential model calibrated using data from an urban freeway in Ontario. The calibrated crash potential models are used to evaluate the safety benefits of a candidate variable speed limit system under three different traffic demand levels (Peak, Near-Peak, and Off-Peak). It is found that safety improvements estimated by the two calibrated crash potential models are within approximately 13% of each other for the Peak and Near-Peak scenarios, but differ by a much larger amount for the Off-Peak scenario. Nevertheless, for each scenario the sign of the safety impact (i.e. increase in safety versus decrease in safety) was the same regardless of the calibration method used.

The results suggest that the safety impacts provided by the crash potential model are robust in that they are relatively insensitive to the optimality of the calibration. This raises the level of confidence that can be placed in the safety impact estimates which is encouraging, particularly when costly investment decisions may be made, in part, on the basis of these model estimates.

# 1. INTRODUCTION

Traditional safety analysis focuses on determining locations that have experienced higher than expected number of collisions (termed "black spots") and then identifying actions, typically geometric improvements (e.g. centre median barrier, etc.) that can improve safety. More recently, attention has been given to identifying, in real-time, traffic conditions that are associated with high crash potential, and when such conditions are observed, to intervene using various freeway traffic management options (e.g. ramp metering, variable speed limits, etc.). These models are often referred to as "real-time crash prediction models".

Real-time crash prediction models can be used in combination with micro-level traffic simulation models to estimate the safety impacts of proposed traffic control strategies or geometric changes. However, the real-time crash prediction models, which are often structured as general log-linear categorical models, must be calibrated using an extensive database containing roadway geometric characteristics, weather conditions, crash records, and real-time traffic conditions (typically determined from loop detector data). The task of calibrating a crash prediction model requires considerable effort and expense. Furthermore, different approaches to model calibration may result in different values for model parameters. The determination of model adequacy is made on the basis of goodness of fit statistics; however, several models all with different parameter values may pass these statistical tests. The important practical question is what impact do these different parameter values have on the safety evaluations that are conducted using the crash potential models? Are the differences small? If so, then the method used to calibrate the crash potential model is not as important. However, if the differences are large, then much more care must be applied in selecting and employing a calibration method, as the final safety evaluation results will be highly dependent on the calibration process.

Furthermore, it may be hypothesised that if the safety evaluation is relatively insensitive to the calibration process, then there is greater confidence that a crash potential model calibrated using data from one freeway section may be valid for application to other similar freeway sections.

In this paper we explore the sensitivity of safety impacts to the safety model calibration. Specifically, we use two methods to calibrate a log-linear crash potential model to data from a section of an urban freeway near Toronto, Ontario. These two crash potential models are then used in combination with a simulation model to evaluate the safety impacts of a candidate variable speed limit system. Differences in the safety impact estimates are examined to provide insights into the sensitivity of the safety impacts to the crash potential model calibration methods.

# 2. CRASH POTENTIAL MODEL

## A. Crash Potential Model Framework

The crash potential model employed in this study, originally developed by Lee et al. (2003 & 2004), calculates crash frequency as a function of traffic conditions, external control factors and exposure. A log-linear model was chosen for the development of the crash potential model since it allows the investigation of the relationships between the input variables for different levels of crash frequency. Also, it can incorporate a value of exposure that is associated with each traffic condition and external control factors. The log-linear function developed by Lee et al. to calculate crash potential is shown in Equation 1. The inputs to the log-linear model are described in the following sections.

$$\text{Crash Potential} = \frac{F}{EXP^{\beta}} = \exp\left(\theta + \lambda_{CVS(i)} + \lambda_{Q(j)} + \lambda_{COVV(k)} + \lambda_{R(l)} + \lambda_{P(m)}\right) \tag{1}$$

Where,

| | |
|---|---|
| $F$ | : expected number of crashes; |
| $EXP$ | : exposure (veh-km); |
| $\beta$ | : parameter for exposure; |
| $\theta$ | : constant; |
| $\lambda_{CVS(i)}$ | : effect of the crash precursor variable *CVS* having *i* levels; |
| $\lambda_{Q(j)}$ | : effect of the crash precursor variable *Q* having *j* levels; |
| $\lambda_{COVV(k)}$ | : effect of the crash precursor variable *COVV* having *k* levels; |
| $\lambda_{R(l)}$ | : effect of road geometry (control factor) having *l* levels; |
| $\lambda_{P(m)}$ | : effect of time of day (control factor) having *m* levels. |

## B. Crash Precursors

The measures of traffic conditions, termed crash precursors, represent the traffic flow conditions prior to a crash occurrence. More turbulent levels of crash precursors correspond to a higher likelihood of an impending crash situation. Lee et al. (2002, 2003) explored a number of traffic flow characteristics believed to be related to crash occurrence, including:

- temporal variation of speed at a fixed location;
- longitudinal variation in speed along road sections;
- variation in speed across lanes;
- lane changing behaviour; and,
- traffic density.

Three of these characteristics were found to be the most statistically powerful in predicting crash occurrence: 1) temporal variation of speed at a fixed location; 2) longitudinal variation in speed; and 3) lane changing behaviour.

First, the temporal variation of speed at a fixed location represents the stability of speeds between vehicles in a traffic stream. Low variation of speed is an indication of smooth traffic flow in which vehicles are traveling at nearly constant speeds. An increase in this speed variation indicates more variability in the speed choice among drivers. This in turn requires drivers to adjust speeds more frequently, leading to the deterioration of flow stability and a higher risk of driver error and an impending crash situation. The temporal variation of speed is measured by the coefficient of variation of speed (precursor CVS), calculated at the nearest detector station upstream of a crash location. CVS is a measure of dispersion which normalizes the standard deviation and is expressed by Lee et al. (2002) through Equation 2.

$$CVS = \frac{1}{n}\sum_{i=1}^{n}\frac{(\sigma_s)_i}{\bar{s}_i} = \frac{1}{n}\sum_{i=1}^{n}\left( \sqrt{\frac{\frac{\Delta t}{t_p}\sum_{t=t^*-\Delta t}^{t^*}(s_i(t))^2 - \left(\sum_{t=t^*-\Delta t}^{t^*} s_i(t)\right)^2}{\frac{\Delta t}{t_p}\left(\frac{\Delta t}{t_p}-1\right)}} \Bigg/ \frac{\sum_{t=t^*-\Delta t}^{t^*} s_i(t)}{\left(\frac{\Delta t}{t_p}\right)} \right) \tag{2}$$

Where,

$t^*$ : time of crash occurrence;
$\Delta t$ : observation time slice duration (seconds);
$(\sigma_s)_i$ : standard deviation of speed on lane *i* computed over period $\Delta t$;
$\bar{s}_i$ : average speed on lane *i* computed over period $\Delta t$ (km/hour);
$t_p$ : time interval of observation of speed (seconds);
$s_i(t)$ : speed on lane *i* at time *t* upstream of a location (km/hour);
$n$ : total number of lanes.

Second, the spatial (longitudinal) variation of speed along road sections measures the difference in averages travel speeds between two consecutive loop detector stations. A small spatial variation indicates constant speeds and a traffic state of little acceleration. However, a large spatial variation in speed indicates traffic will experience an abrupt change in travel speed, requiring either sudden acceleration or deceleration. A state of sudden deceleration is most likely to cause crashes and often occurs as a traffic queue is formed during recurrent or non-recurrent congestion. Spatial variation of speed, represented by precursor *Q* is expressed through Equation 3 (Allaby, 2006).

$$Q = \bar{s}_1 - \bar{s}_2 = \frac{t_p}{\Delta t}\sum_{t=t^*-\Delta t}^{t^*}\left(\frac{\sum_{i=1}^{n_1} s_{1i}(t)v_{1i}(t)}{\sum_{i=1}^{n_1} v_{1i}(t)}\right) - \frac{t_p}{\Delta t}\sum_{t=t^*-\Delta t}^{t^*}\left(\frac{\sum_{i=1}^{n_2} s_{2i}(t)v_{2i}(t)}{\sum_{i=1}^{n_1} v_{2i}(t)}\right) \tag{3}$$

Where,

$Q$ : average speed difference between upstream and downstream locations (km/hour);
$\bar{s}_1, \bar{s}_2$ : average speeds computed over period of $\Delta t$ upstream and downstream of a location, respectively (km/hour);
$s_{1i}(t)$ : speed on lane *i* at time *t* upstream of a location (km/hour);
$s_{2i}(t)$ : speed on lane *i* at time *t* downstream of a location (km/hour);
$v_{1i}(t)$ : volume on lane *i* at time *t* upstream of a location (km/hour);
$v_{2i}(t)$ : volume on lane i at time t downstream of a location (km/hour);
$n_1, n_2$ : numbers of lanes upstream and downstream of a location, respectively.

The third traffic characteristic, lane changing behaviour, is estimated by the average covariance of volume difference between upstream and downstream locations across adjacent lanes (precursor COVV). The covariance of volume captures the correlation of traffic volume changes between two lanes (i.e. traffic moving from lane 1 to lane 2 creates a volume reduction on lane 1 and subsequently a volume increase on lane 2). Thus, COVV is a surrogate measure of lane changing activity. High levels of COVV indicate frequent lane changing and thus more turbulence within the traffic stream that increases the likelihood of a crash occurrence. Lee (2004) expresses COVV through Equation 4.

$$COVV = \frac{1}{n-1}\sum_{i=1}^{n-1}\left|\mathrm{cov}(\mathrm{V}_i, \mathrm{V}_{i+1})\right| = \frac{1}{n-1}\sum_{i=1}^{n-1}\sum_{t=t^*-\Delta t}^{t^*}\left[\left(\Delta v_i(t) - \Delta\bar{v}_i\right)\left(\Delta v_{i+1}(t) - \Delta\bar{v}_{i+1}\right)\right] \quad (4)$$

Where,

$\Delta v_i(t)$ : volume difference between upstream and downstream locations on lane $i$ at time $t$;

$\Delta \bar{v}_i$ : average volume difference on lane $i$ over period $\Delta t$;

$\mathbf{V}_i$ : vector of $\Delta v_i(t)$ over period $\Delta t$ (i.e. $\mathbf{V}_i = \{\Delta v_i(t_0), \Delta v_i(t_0+1), \ldots, \Delta v_i(t_0+\Delta t\,)\}$).

* Lane $i$ and lane $i$+1 are adjacent lanes.

The three crash precursors, CVS, Q, and COVV can be calculated on the basis of traffic flow measures such as volume, speed, and occupancy, which are easily extracted from dual loop detector data. The period over which the precursors are calculated is called an observation time slice, Δt. The selection of Δt involves identifying the time over which the impact of crash precursors on crash occurrence is maximized. For example, is the impact of CVS most significant when it is computed over periods of 2-minutes, 5-minutes, or 10-minutes prior to a crash? To establish the best observation time slice for each precursor, Lee et al. applied a method to determine the maximum difference in precursor values between crash and non-crash cases. For a number of time intervals, they measured the difference in the frequency distribution of the precursor values between crash and non-crash data. They found that the frequency differences in CVS, Q, and COVV, were maximized at Δt = 8 minutes, 2 minutes, and 2 minutes, respectively. These observation time slice durations were carried forward for use in this study.

## C. External Control Factors

External control factors include road geometry, time of day and environmental conditions. These factors alone can affect driver behaviour, so it was necessary to include them in the crash potential model to identify the isolated effects of the crash precursors. Road geometry refers to the lane configuration of a freeway segment with regard to entrances and exits, rather than horizontal or vertical alignment. Lee et al. (2002, 2003, and 2004) found that freeway segments with merging or diverging traffic contribute more to crash potential than straight freeway segments with no changes in lane configuration. Time of day refers to peak and off-peak periods. Typically, traffic volumes and congestion are higher during peak periods and drivers, particularly commuters, may react more aggressively to maintain their schedules. These factors are likely to increase the likelihood of a crash occurrence during the peak periods. Lastly, environmental conditions include such factors as local weather, road surface quality, and lighting. Due to the limited amount of available environmental data, the effect of these factors can be difficult to capture.

## D. Exposure

Exposure forms a relationship between the frequency of traffic and environmental events and the associated crash frequency. For example, consider two traffic scenarios. The first scenario arises 20% of the time and experiences 20% of all crashes. The second scenario also experiences 20% of all crashes, but arises only 5% of the time. Although the crash frequencies are identical, the crash rate (crash frequency divided by level of exposure) is clearly higher for the second scenario.

Exposure is expressed in the crash model as the number of vehicle-kilometres exposed to each combination of traffic characteristics and external control factors. In other words, the average vehicle-kilometres present over a road section are multiplied by the probability of a certain time of day (peak or off-peak) at a certain type of road geometry (merge/diverge or straight) under a certain range of crash precursor values. The expression for exposure is shown in Equation 5.

$$\mathrm{EXP} = p(CVS) \cdot p(Q) \cdot p(COVV) \cdot p(P) \cdot p(G) \cdot V \cdot L \cdot T \qquad (5)$$

Where,

| | | |
|---|---|---|
| $EXP$ | : | exposure (veh-km); |
| $p(CVS)$, $p(Q)$, $p(COVV)$ | : | probabilities that precursors (i.e. $CVS$, $Q$ and $COVV$) calculated from normal daily traffic will fall within the range of values; |
| $p(P)$ | : | proportion of normal daily traffic volume associated with the peak and off-peak periods; |
| $p(G)$ | : | proportion of road section associated with merging/ diverging and straight geometry; |
| $V$ | : | average annual daily traffic of the road section (veh/day); |
| $L$ | : | length of section (km); |
| $T$ | : | Number of days over which calibration data obtained. |

Since the crash potential model is log-linear, crash precursors will be categorized. The probabilities of the precursors in the expression for exposure will depend on the selection of categories. A category spanning a larger range of crash precursor values will result in a higher probability for exposure than a category spanning a small range of values.

# 3. SAFETY MODEL CALIBRATION

## A. Introduction

Calibration of the crash potential model consists of the following 6 steps:

1. Selection of Model Calibration Site

2. Development of Crash Database
3. Identification of External Control Factors
4. Calculation of Crash Precursors
5. Categorization (Selection of Boundary Values)
6. Log-Linear Analysis

In this paper, we focus on the impact of Step 5 – Categorization on the model fit and on safety estimates obtained from the model. In particular, we compare two categorization methods, namely an ad hoc method, and a near-optimal method.

The steps in the model calibration are described in more detail in the following sections.

## B. Field Data

A segment of the Queen Elizabeth Way (QEW) in Mississauga, Ontario was chosen to calibrate the crash potential model for this study. The QEW is a multilane freeway located in south western Ontario, Canada. The freeway begins near the Canadian/American border at Fort Eerie and, following the coastline of Lake Ontario, passes through several urban centres such as Niagara Falls, St. Catharines, Hamilton, Burlington, Mississauga, and finally into Toronto. The QEW near Toronto services a large volume of commuter traffic in the morning and evening peak periods, resulting in heavy congestion and a high frequency of crashes,

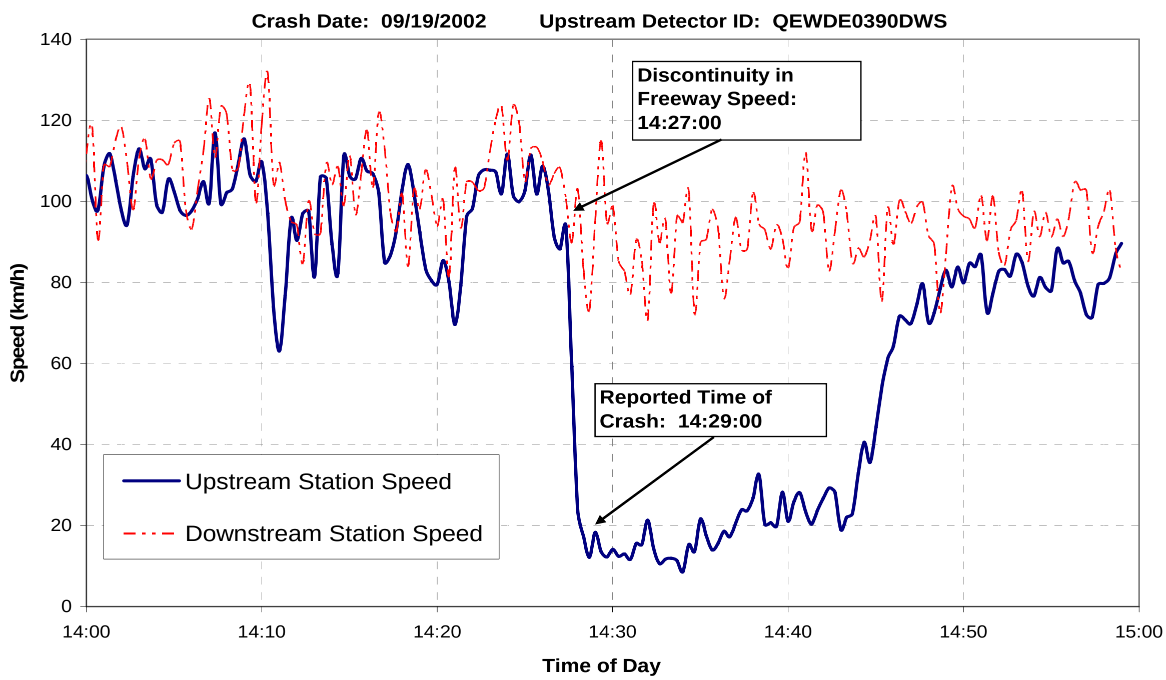
The segment used for calibration was a 13 km section between Royal Windsor Dr. and Highway 427 including both directions of travel. This freeway segment features a posted speed limit of 100 km/hr, has 3 to 4 mainline lanes, and experiences a directional AADT of about 70,000 vehicles. The section is instrumented with dual loop detector stations in each mainline lane spaced at approximately 600m. The study segment contains 26 loop detector stations in each travel direction. Every 20 seconds, speed, volume, and occupancy are recorded for all mainline stations.

Detector data and Freeway Traffic Management System (FTMS) incident logs were obtained from the Ministry of Transportation of Ontario (MTO). Crash records were compiled for the period of January 1998 through February 2003. The FTMS incident logs provided several pieces of information on every incident detected on the highway. Of most importance to this study were:

- Date and time incident was reported;
- Identity of upstream loop detector station; and
- Type of incident (e.g. accidents, breakdowns, etc.)

Using the information provided, the FTMS incident logs were filtered to form a crash database with records appropriate for the crash model calibration. Loop detector data were obtained for the upstream and downstream location of each crash, for a thirty minute time period before and after the reported time of the crash. The time of the crash was estimated on the basis of the traffic stream speed measured by the upstream and downstream loop detectors. The time of crash was determined as the time when the speed at the upstream detector station abruptly decreased due to the passage of a compression shockwave (Figure 1).

**Figure 1: Confirmation of Crash Time, *t***




This process resulted in the development of a crash database containing 299 crashes.

For each of the crashes in the crash database two control factor conditions were recorded, namely:

- Time of Day (Peak 6-10 AM and 4-7 PM; Off-Peak (7 PM – 6 AM and 11 AM – 4 PM); and
- Geometric Configuration (Straight and Merging/Diverging)

## C. Model Categorization

Model categorization, the fifth step in the safety model calibration process, is one of the most important. The control factors are already categorical (e.g. merge/diverge geometry versus straight sections). However, the precursors are continuous values and must be categorized. Categorization requires two decisions, namely (1) the number of categories into which the pre-cursor will be discretized and (2) specification of the pre-cursor value that defines the boundary between two categories (Figure 2). If three categories are chosen, then two boundary values must be specified. For four categories, three boundary values are required, etc.

**Figure 2: Categorization of Precursors**


After all precursor data were transformed into categorical references and values of exposure were calculated, log-linear analysis (Step 6) was performed using SPSS Version 13.0 (SPSS Inc., 2004) to calibrate the crash potential model. This procedure analyzes the frequency of samples in each cell of the contingency table to yield maximum likelihood estimates of the expected frequency of crashes under each possible condition. The analysis used an iterative fitting process until the difference between the current and previous estimates converged to 0.001. Because crashes are considered to be random events, the crash frequency was assumed to follow a Poisson distribution in the fitting process.

The performance of each categorization case was measured in terms of 1) overall model fit; 2) the statistical significance of individual coefficients; and 3) the consistency of coefficients with the order of levels of precursors (i.e. it is expected that "high" levels of precursors contribute more to crash potential than "low" level precursors).

The overall model fit was measured by a log-likelihood ratio $\chi^2$ test. This test measures the differences between the observed crash frequencies and expected crash frequencies for any combination of crash precursor categories and control factors. A low $\chi^2$ and a high p-value (greater than $1 - \alpha$) indicate that the distribution of the expected crash frequencies is not significantly different from the distribution of the observed crash frequencies – or in other words, the model fits well.

## D. Categorization Method 1: Ad hoc method

Allaby (2006) calibrated the crash potential model using a heuristic ad-hoc method in which he selected a number of categories for each of the three pre-cursors and then selected boundary values for each precursor. The number of categories is practically limited to between 2 and 4 as the total number of cells in the contingency table must be less than the number of crashes in the database. However, any set of boundary values is possible. Allaby constrained the number of options by assuming a proportion of the observed values would fall into each precursor category. For example, if 4 categories are assumed for precursor CVS, then one set of boundary values can be determined by assuming 20% of CVS observations in the first category, 30% in the second, 30% in the third, and 20% in the last category. Allaby evaluated a relatively small number (approximately 30) of categorizations and selected the one which provided the best model fit.

## E. Categorization Method 2: Near-Optimal Method

One criticism of the method adopted by Allaby is that there is no certainty that the selected model is the best. For example, a better model may have been obtained if some other categorization had been attempted. To address this concern, we developed software to automate the model calibration process (Steps 5 and 6) permitting the evaluation of a very large number (almost 5,000) of categorizations. The statistical fit of the model for each categorization was maintained in a separate database. Model suitability criteria, similar to those used by Allaby, were automatically applied to the database to identify the optimal model categorization.

## F. Categorization Results

Both Method 1 and Method 2 were applied to the same set of field data to obtain a calibrated crash potential model. The model characteristics are provided in Table 1.

**Table 1. Crash Model Parameter Estimates from Log-linear Analysis**

| Calibration Method | Crash Precursor | Categorization | Boundary Values $B_1$ | $B_2$ | $B_3$ |
|---|---|---|---|---|---|
| Method 1 | CVS | 20/30/30/20 | 0.062 | 0.089 | 0.139 |
| | Q | 20/30/30/20 | -9.19 | 0.09 | 8.77 |
| | COVV | 40/40/20 | 1.49 | 3.44 | -- |
| Method 2 | CVS | 37/33/30 | 0.079 | 0.126 | -- |
| | Q | 43/37/20 | -1.433 | 10.05 | -- |
| | COVV | 23/34/43 | 0.815 | 1.833 | -- |

For Method 1, the log-linear analysis resulted in a p-value close to 1.0 and a chi-squared likelihood ratio of 112.18 with degrees of freedom of 180. For Method 2, the log-linear analysis resulted in a p-value of 1.0 (closer to 1.0 is better) and a chi-squared likelihood ratio of 49.47 (smaller is better) with degrees of freedom of 98. The parameter estimates and statistical significance resulting from categorizing Methods 1 and 2 are provided in Tables 2 and 3, respectively.

The constant term in the model provides the crash frequency ($=e^{\theta}$) for base case factors (parameters with estimates equal to 0) for given values of exposure. Parameters for which estimates are negative indicate a declining contribution to crash potential, whereas positive estimates indicate an increasing contribution to crash potential.

**Table 2. Crash Model Parameter Estimates from Calibration Method 1**

| Parameter | Category Level | Estimate | Z-Value |
|---|---|---|---|
| **Constant ($\theta$)** | N/A | 1.518 | 9.78 |
| $\lambda_{COVV}$ | [1] Low | -1.300 | -5.32 |
| | [2] Intermediate | -0.884 | -3.70 |
| | [3] High | 0 | . |
| $\lambda_{CVS}$ | [1] Low | -0.914 | -5.17 |
| | [2] Intermediate A | -1.735 | -8.11 |
| | [3] Intermediate B | -1.496 | -7.32 |
| | [4] High | 0 | . |
| $\lambda_{Q}$ | [1] High Acceleration | -0.875 | -4.89 |
| | [2] Moderate Acc. | -1.738 | -8.12 |
| | [3] Mod Deceleration | -1.508 | -7.44 |
| | [4] High Deceleration | 0 | . |
| $\lambda_{R}$ (for Geometry) | Straight | -0.530 | -4.40 |
| | Merge Diverge | 0 | . |
| $\lambda_{P}$ (for Time) | Off-Peak | -1.254 | -8.16 |
| | Peak | 0 | . |
| $\beta$ (for Exposure) | N/A | 0.084 | 7.22 |

**Table 3. Crash Model Parameter Estimates from Calibration Method 2**

| Parameter | Category Level | Estimate | Z-Value |
|---|---|---|---|
| **Constant ($\theta$)** | N/A | 1.929 | 4.81 |
| $\lambda_{COVV}$ | [1] Low | -2.132 | -3.13 |
| | [2] Intermediate | -2.107 | -5.20 |
| | [3] High | 0 | . |
| $\lambda_{CVS}$ | [1] Low | -1.577 | -6.88 |
| | [2] Intermediate | -1.203 | -7.4 |
| | [3] High | 0 | . |
| $\lambda_{Q}$ | [1] High Acceleration | -2.452 | -4.14 |
| | [2] Moderate | -2.107 | -4.74 |
| | [3] High Deceleration | 0 | . |
| $\lambda_{R}$ (for Geometry) | Straight | -0.618 | -5.03 |
| | Merge Diverge | 0 | . |
| $\lambda_{P}$ (for Time) | Off-Peak | -1.544 | -6.41 |
| | Peak | 0 | . |
| $\beta$ (for Exposure) | N/A | 0.049 | 3.78 |

On the basis of the results provided in Tables 2 and 3, it is difficult to comment on the differences in the two crash precursor models. It is clear that the models have different number and magnitude of parameter coefficients, but it is not clear how these differences influence the safety estimates obtained by using these crash potential models. And it is the estimates of safety impacts that are ultimately of interest to traffic engineers. The next section describes the application of the two crash potential models presented in Tables 2 and 3 to estimate the safety impacts of a potential real-time control strategy, namely a variable speed limit sign system. In particular, the focus is on the degree of agreement between the safety impacts estimated by the two models.

# 4. IMPACT OF CALIBRATION ON SAFETY IMPACT ESTIMATES

The two calibrated crash potential models described in the previous section were used to estimate the safety impacts of a candidate variable speed limit control strategy. The focus on this application is the degree of consistency of the safety impacts obtained from the two calibrated safety models rather than on the variable speed limit sign system itself.

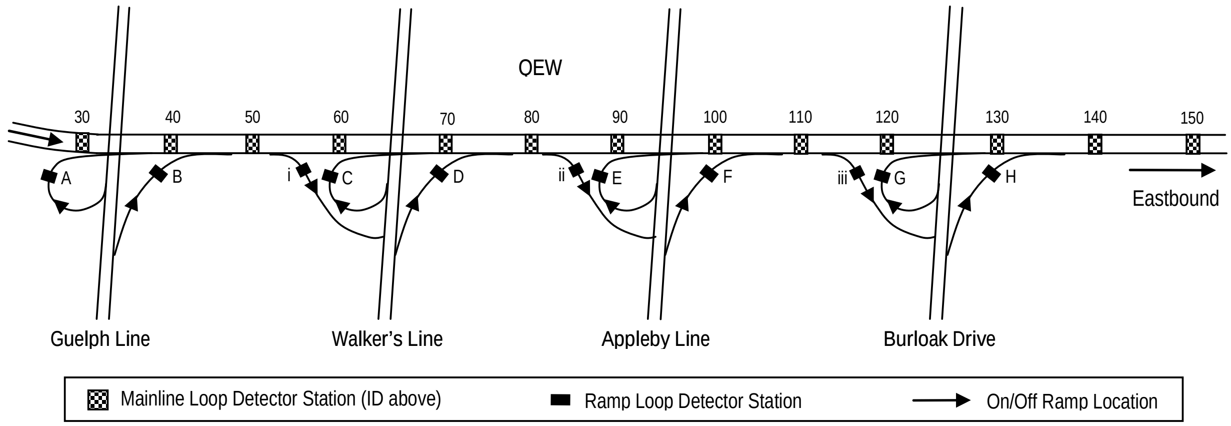
## A. Variable Speed Limit Sign (VSLS) Systems

Variable Speed Limit Sign (VSLS) systems consist of dynamic message signs (DMS) deployed along a roadway and connected via a communication system to a traffic management centre. The VSLS are used to display a regulatory or advisory speed limit. Unlike typical static speed signs, the VSLS system enables transportation system managers to dynamically post a speed limit that is appropriate for current traffic, weather, or other conditions. VSLS are thought to improve safety and reduce driver stress while improving traffic flow and travel times. However, few evaluation studies have been conducted to establish the safety impacts associated with VSLS. Allaby (2006), Allaby and Hellinga (2006) and Allaby et al. (2006a, 2006b) combined a crash potential model (i.e. categorization Method 1) with a micro-level traffic simulation model (PARAMICS) to estimate the safety and system delay impacts of a candidate VSLS system.

## B. Test Network

An 8 km section of the eastbound Queen Elizabeth Way (QEW) located near Toronto, Canada was selected as the test network. The QEW services a large volume of commuter traffic in the morning and evening peak periods, resulting in heavy congestion and a high frequency of crashes. The study segment features a posted speed limit of 100 km/hr, has three mainline lanes, contains four interchanges, and experiences a directional AADT of about 70 000 vehicles. The section is instrumented with dual loop detector stations in each mainline lane spaced at approximately 600 m and single loop stations on entrance and exit ramps (Figure 3). Every 20 seconds, speed, volume, and occupancy are recorded for all mainline stations, whereas volume is recorded for all ramp stations.

**Figure 3: Test Network**

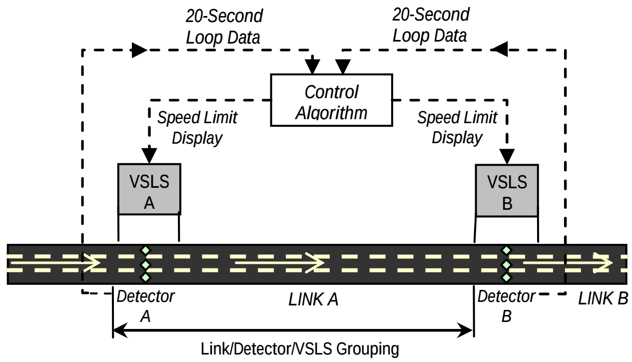



During the morning peak period (6:00 am to 10:00 am) this freeway section experiences high levels of recurrent congestion. This congestion is mainly caused by a bottleneck created at the most downstream interchange. At this location, a high volume of traffic (~1000 veh/hr) entering the already congested mainline results in reduced freeway speeds, queues, and an upstream moving shockwave that penetrates much of the section. Freeway speeds through the bottleneck during this period typically range from 30 km to 50 km, but at times traffic is observed to be at a standstill.

The simulation model was calibrated to observed network conditions for the morning peak period. Simulation parameters were adjusted until the speed profiles adequately matched the observed profiles (within confidence limits of +/- 2σ).

The VSLS system infrastructure was represented within PARAMICS by thirteen variable speed limit signs placed throughout the network. Each VSLS was placed next to a loop detector, spaced at approximately 500 m to 600 m. Since PARAMICS assigns speed limits by link, the mainline was coded as a series of links corresponding to each detector-VSLS pair. Each link / detector / VSLS set acted as its own entity – the detector gathered information about traffic conditions, the appropriate "condition based" speed was assigned to the link, and the VSLS displayed the current speed limit for the benefit of the user/observer. Figure 4 illustrates this layout. Based on traffic data received every 20 seconds from "loop detector A", a control algorithm determined the appropriate speed limit to be displayed at "VSLS A." This displayed speed limit governed until the end of "Link A", at which point a new displayed speed limit at "VSLS B" was determined by traffic data from "loop detector B."

**Figure 4: VSLS system configuration**




A candidate VSLS system control strategy was developed and is described elsewhere (Allaby and Hellinga, 2006; Allaby et al., 2006a, 2006b).

The VSLS impact analyses were performed on three traffic scenarios of varying levels of congestion – heavy, moderate, and light. These scenarios were termed Peak, Near-Peak, and Off-Peak, respectively. The validated simulation model from the observed morning peak period conditions represented the Peak traffic scenario. The Near-Peak and Off-Peak scenarios were represented by approximately 90% and 75%, respectively, of the peak volumes. These scenarios were not calibrated for existing conditions as their purpose was to investigate and understand the varying reaction of the VSLS system to changes in congestion, rather than to replicate real traffic conditions. The VSLS impact was quantified in terms of the relative changes in safety (crash potential) and vehicle travel times before and after the implementation of the VSLS control strategy. The results of the VSLS safety impacts are presented in the next section.

## C. Estimates of Safety Impacts

In this study, the safety impact of VSLS was measured by calculating the relative change in crash potential from the non-VSLS case to the VSLS case. Ten simulation runs were performed for the non-VSLS case and ten for the VSLS case. The same set of ten seed values was used for the VSLS and non-VSLS runs. For each simulation run, at each station, a value of crash potential (CP) was calculated from crash precursor values on 20-second intervals. Then, average values of station crash potential (SCP) were obtained for each run over the simulation period using Equation 6.

$$SCP_i = \frac{1}{n}\sum_{j=1}^{n} CP_{ij} \tag{6}$$

Where,

$SCP_i$ : Station Crash Potential for Station *i* (crashes/million veh-km);
$CP_{ij}$ : Crash Potential for Station *i* at 20-second interval *j* (crashes/million veh-km);
$n$ : Number of 20-second intervals in simulation period (720 for 4-hour period)

Since the non-VSLS and VSLS cases differed only by the introduction of the VSLS system, the SCP values could be paired by simulation run. A paired 2-tailed student t-test was used to test for the significance of the change in SCP (or VSLS impact) at the 95% level of confidence. If the difference was found to be significant, the relative safety benefit (RSB) was calculated using Equation 7. A positive relative safety benefit represented a decrease in crash potential.

$$RSB_i = \left( \frac{ASCP_i(non-VSLS) - ASCP_i(VSLS)}{ASCP_i(non-VSLS)} \right) \times 100 \tag{7}$$

Where,

*RSB*$_i$ : Relative Safety Benefit at Station *i* (%);

*ASCP*$_i$ : Average Station Crash Potential (average of SCP over *x* simulation runs) at Station *i* (crashes/million veh-km).

For example, the average relative safety benefit associated with the VSLS as estimated using categorization Method 2 for the Near-Peak scenario is 17% (Table 4).

**Table 4. Method 2 Categorization – Near Peak Scenario**

| Station | ACP | | Impact | | Significant at 95%? | Significant Impacts |
|---|---|---|---|---|---|---|
| | Non-VSLS | VSLS | | | | |
| 40 | 1.014 | 2.113 | -108% | NEGATIVE | YES | -108% |
| 50 | 0.218 | 0.226 | -4% | NEGATIVE | NO | |
| 60 | 0.208 | 0.062 | 70% | POSITIVE | NO | |
| 70 | 0.455 | 0.302 | 34% | POSITIVE | YES | 34% |
| 80 | 0.207 | 0.133 | 36% | POSITIVE | NO | |
| 90 | 0.520 | 0.429 | 18% | POSITIVE | NO | |
| 100 | 0.839 | 0.663 | 21% | POSITIVE | NO | |
| 110 | 0.816 | 0.643 | 21% | POSITIVE | YES | 21% |
| 120 | 0.573 | 0.314 | 45% | POSITIVE | YES | 45% |
| 130 | 0.653 | 0.310 | 53% | POSITIVE | YES | 53% |
| 140 | 0.332 | 0.131 | 61% | POSITIVE | YES | 61% |
| | | | | | Average | 17% |

However, in this paper we are specifically interested in a comparison of the average relative safety benefit obtained from the two categorization methods (Table 5).

For the Peak scenario, the use of categorization Method 2 suggested an average relative safety benefit (i.e. improvement in safety) of 44.3%. The use of categorization Method 1, with all other aspects of the analysis unchanged, estimates an average relative safety benefit of 40.1%. These estimates are relatively similar (with the absolute difference being only approximately 10.1% of the average of the relative safety benefits associated with Methods 1 and 2).

For the Near-Peak scenario, the average relative safety benefits obtained from the two categorization methods are again similar. Method 1 results in an estimate of 19.9% improvement in safety and Method 2 results in an improvement of 17.4%. The absolute difference between these two estimates is approximately 13% of the average of the estimates provided by Methods 1 and 2. However, for this scenario Method 2 provides a lower estimate of the safety benefits than Method 1. In the Peak scenario, Method 2 provides a higher estimate of the safety benefits.

For the Off-Peak scenario, the safety impacts obtained from the two categorization methods differ substantially. Method 1 suggests a 10.8% decrease in safety. Method 2 suggests a decrease of almost 54%. Examination of the results from Method 2 for the Off-Peak scenario revealed that the average relative safety dis-benefit was computed from only 3 of the 11 stations, as only these three stations had statistically significant results at the 95% level. When more of the stations are included in the calculation (i.e. using a 90% confidence limit) then the average relative safety impact changes to a 28% decrease in safety.

Despite the differences that are observed between the two categorization methods, the two methods do provide consistent estimates of the sign of the safety impact (i.e. increase in safety versus decrease in safety). Furthermore, for the Peak and Near-Peak scenarios, the magnitude of the safety impact estimated by the two categorization methods differ by less than approximately 13%. Given the level of uncertainty associated with other aspects of these types of safety impact studies (e.g. simulation calibration, driver behaviour assumptions, etc.) this level of consistency is likely adequate for decision making. Furthermore, the demonstrated level of consistency suggests that the safety model approach demonstrated in this paper is robust in that the final conclusions regarding the expected safety impacts of the candidate VSLS system are not highly sensitive to the degree of optimality of the safety model calibration.

**Table 5. Comparison of Estimated Safety Impacts**

| Scenario | Average Safety Benefit | | |
|---|---|---|---|
| | Method 1 | Method 2 | Change |
| Peak | 40.1% | 44.3% | 10.1% |
| Near-Peak | 19.9% | 17.4% | -13.0% |
| Off-Peak | -10.8% | -53.9% | 133.3% |

# 5. CONCLUSIONS

This study has demonstrated the following:

1. Different approaches to safety model calibration may result in different model structures (e.g. number of categories) with different coefficient values.

2. The importance of differences in safety model structure and/or coefficient values can not be determined solely on the basis of a comparison of the model characteristics. Rather it is necessary to apply the models to estimate the safety impacts of a particular intervention (in this case a candidate VSLS system). Comparing the safety impacts from the different candidate safety models provides an indication of the important of model differences.

3. The safety impact assessment method utilized in this paper (e.g. combination of the crash potential model with a micro-level simulation model) to assess the safety benefits of a

candidate VSLS system is robust in that the estimated benefits generally are not highly sensitive to the optimality of the safety model calibration.

This last observation also suggests that the crash potential model used in this study may be transferable to other freeway sections. If the safety impacts are not highly sensitive to the optimality of the model calibration to data obtained from a given freeway site, then it may also be true that the safety benefits obtained from a model calibrated to data from a particular freeway site may be relatively similar to the benefits obtained from a model calibrated to data from some other freeway site. Though this study has not directly tested this hypothesis, the results of this study do provide evidence that the safety impacts are not highly sensitive to the safety model calibration. This is encouraging, as it is desirable to be able to apply a calibrated safety model to various freeway sections.

# 6. REFERENCES

P. Allaby, B. Hellinga, and M. Bullock (2006a) *Variable Speed Limit Systems: Safety and Operational Impacts for Freeways Applications*. Presented at the 13th ITS World Congress on Intelligent Transportation Systems held October 8 – 12, 2006 in London, UK.

P. Allaby, B. Hellinga, and M. Bullock (2006b) *Variable Speed Limits: Safety and Operational Impacts of a Candidate Control Strategy for an Urban Freeway*. Presented at the 9th International IEEE Conference on Intelligent Transportation Systems held Sept. 17 – 20 in Toronto, Ontario Canada.

P. Allaby and B. Hellinga (2006) *Variable Speed Limits: Safety and Operational Impacts of a Candidate Control Strategy on a Modelled Urban Freeway*. Submitted to the Advanced Traffic Management Section of the Ministry of Transportation of Ontario (69 pages).

P. Allaby (2006) *An Evaluation of the Safety and Operational Impacts of a Candidate Variable Speed Limit Control Strategy on an Urban Freeway*. MASc Thesis submitted to the University of Waterloo

C. Lee, B. Hellinga, and F. Saccomanno (2006) *Evaluation of variable speed limits to improve traffic safety*. Transportation Research Part C - Emerging Technologies Vol. 13 Issue No. 3, pp. 213-228. Published by Elsevier Ltd.

C. Lee, B. Hellinga, and F. Saccomanno (2004) *Assessing Safety Benefits of Variable Speed Limits*. Transportation Research Record 1897 - Statistical Methods and Safety Data Analysis and Evaluation. Pages 183-190. Published by the Transportation Research Board, National Academy Press. Washington, D.C.

C. Lee, B. Hellinga, and F. F. Saccomanno (2004) *Real-Time Crash Prediction Model for the Application to Crash Prevention in Freeway Traffic*. Transportation Research Record 1840 - Journal of the Transportation Research Board. National Academy Press. Washington, D.C. pages 67 – 78.

C. Lee, B. Hellinga, and F. Saccomanno. (2003) *Proactive Freeway Crash Prevention Using Real-Time Traffic Control.* Canadian Journal of Civil Engineering - Special Issue on Innovations in Transportation Engineering. 30(6). Pages 1034-1041

C. Lee, F. Saccomanno, and B. Hellinga (2002) *Analysis of Crash Precursors on Instrumented Freeways*. Transportation Research Record 1784 - Journal of the Transportation Research Board. National Academy Press. Washington, D.C. pp. 1 - 8.

SPSS Inc. (2004) *SPSS for Windows Release 13.0.* Chicago, Illinois.

# AUTHORS' INFORMATION

Bruce Hellinga, PhD, PEng
Associate Professor
Department of Civil and Environmental Engineering,
University of Waterloo
Waterloo, Ontario, N2L 3G1, Canada
Tel: (519) 888-4567 ext. 32630
Fax: (519) 888-4349
E-mail: bhellinga@uwaterloo.ca

Amir Samimi
PhD Candidate
Department of Civil and Materials Engineering,
University of Illinois at Chicago
Chicago, Illinois, 60607-7023 USA
Email: asamim2@uic.edu