

1 **DEVELOPING MACRO-LEVEL COLLISION PREDICTION MODELS TO EVALUATE**  
2 **BICYCLE SAFETY IN THE CITY OF VANCOUVER**

3  
4  
5 **Bianca Popescu, M.A.Sc.**

6 Research Assistant, Department of Civil Engineering

7 The University of British Columbia

8 6250 Applied Science Lane

9 Vancouver, BC, Canada V6T 1Z4

10 Email: [Bianca.Popescu@civil.ubc.ca](mailto:Bianca.Popescu@civil.ubc.ca)

11  
12 **Tarek Sayed, Ph.D., P.Eng.**

13 Professor, Department of Civil Engineering

14 The University of British Columbia

15 6250 Applied Science Lane

16 Vancouver, BC, Canada V6T 1Z4

17 Email: [tsayed@civil.ubc.ca](mailto:tsayed@civil.ubc.ca)

18  
19  
20  
21  
22  
23 *Keywords:* Safety, Macro-Level Collision Prediction Models, Bicycle, Active Transportation,  
24 Bicycle Safety, Black Spot Programs, Road Safety Improvement Programs  
25  
26  
27

28 Word count: 5,192 words text + 3 tables/figures x 250 words (each) = 5,942 words

## ABSTRACT

To encourage greener cities while reducing transportation impacts, such as climate change, traffic congestion, and road safety issues, governments have been investing in sustainable modes of transportation such as cycling. A safe and comfortable cycling environment is critical to encourage bicycle trips, since cyclists are usually subject to greater safety risks. Traditionally, engineering approaches of road safety management have addressed road safety in reaction to existing collision records. For bicycle collisions, which are rare events, a proactive approach is more appropriate. This study describes the use of bicycle related macro-level (i.e. neighbourhood or zonal level) collision prediction models (CPMs) as empirical tools in road safety diagnosis and planning. The developed models incorporate an actual bike exposure indicator (bike kilometers travelled). The macro-level bicycle-vehicle collisions models were applied to a case study in the City of Vancouver at the zonal level. Collision prone zones in the City of Vancouver are identified and the highest-ranked zones are diagnosed to identify bicycle safety issues and to recommend potential safety countermeasures. The findings from this study suggest that the safety issues may be a result of the high density and commercial land use type, coupled with high traffic volume particularly on the arterial routes and high bicycle volumes on routes with mixed vehicle and bicycle traffic. The case study successfully demonstrated the use of the models to proactively enhance bicycle safety.

## 1 INTRODUCTION

2  
3 The transportation-related challenges of climate change, traffic congestion, public health  
4 and road safety are impacting cities worldwide. To encourage greener cities, governments are  
5 prepared to invest in sustainable transportation systems. In contrast to motor vehicle travel, the  
6 public health benefits of active transportation, such as cycling and walking, are significant. In the  
7 case of bicycle transport, some key advantages include: energy efficiency, low cost, health  
8 benefits and zero emissions, as well as an effective use of road space and parking. To encourage  
9 active transportation, policies need to seriously consider the safety of these road users.  
10 Sustainable transportation planning and design that enables less dependence on single occupancy  
11 vehicles and promotes the safe use of bicycles could be an effective way to achieve a sustained  
12 reduction in road collision risk, frequency, and severity for cyclists.

13 The social and economic burdens due to road collisions are recognized as a global  
14 problem. According to the World Health Organization's global status report on road safety, road  
15 traffic injuries are the eighth leading cause of death globally, remaining unacceptably high at  
16 1.24 million road traffic deaths per year [1]. Current global trends suggest that by 2030 road  
17 traffic deaths will become the fifth leading cause of death unless urgent action is taken [2]. In  
18 addition to the human cost, collisions put a large economic burden on society. Transport Canada  
19 estimates the annual cost of total collisions at \$62.7 billion for the Canadian economy [3].

20 Across Canada, cycling is growing in popularity as a daily commuting option, providing  
21 a convenient and affordable alternative to the congested roads and crowded transit systems in  
22 urban areas [4]. However, inadequate infrastructure and unsafe environments present a threat to  
23 meeting sustainable transportation system goals. Vulnerable road users (VRUs), such as  
24 pedestrians and cyclists, are subject to greater safety risks and represent the highest share of  
25 severe and fatal road collisions [5]. In British Columbia in 2013, 100% of the 1,500 reported  
26 automobile collisions involving bicycles resulted in an injury or fatality for the cyclist due to the  
27 severe nature of such collisions [6]. This study focuses on bicycle collisions with motor vehicles,  
28 since studies show that such collisions are more severe to the cyclist [7] [8].

29 The goal of this study is to investigate the use of macro-level collision prediction models  
30 for bicycle safety evaluation using a traditional Road Safety Improvement Program (RSIP)  
31 methodology. Collision prone zones in the City of Vancouver are identified and the  
32 highest-ranked zones are diagnosed to recognize bicycle safety issues and to recommend  
33 potential safety countermeasures. The models developed in this study are unique as they  
34 incorporate an actual bike exposure indicator (i.e. bike kilometers travelled) unlike similar  
35 studies, which rely on exposure proxies such as pollution or bike network length.

## 36 PREVIOUS WORK

### 37 Macro-Level CPMs

38  
39 Progress in road safety research has improved the empirical tools for proactive road  
40 safety management, including the development of macro-level CPMs. Due to the interest in  
41 community-wide proactive safety planning, there has been a growing body of research produced  
42 on the development of macro-level CPMs to date [9] [10] [11] [12] [13] [14]. Replacing the  
43 single-facility, individual link and node data used for micro-level CPM development and  
44 prediction, the macro-level CPM input data is aggregated to the traffic analysis zone (TAZ) level.  
45  
46

The resulting macro-level CPMs relate the zonal collisions to other variables such as zonal volume, infrastructure, and socio-demographic variables.

Most macro-level CPMs in previous studies have been developed for vehicle collisions only. More recently, there have been models taking account of transit characteristics and transit collisions [15] [16], as well as bicycle collisions. Due to the increase of bicycle volumes on-street and the potential effect on safety, researchers have recommended the importance of developing bicycle related macro-level CPMs [17] [18] [19]. There have been several studies that have developed macro-level bicycle-related CPMs. Wei and Lovegrove (2013) developed macro-level community-based CPMs for bicycle-vehicle collisions, using data from the Central Okanagan Regional District and using bicycle lane kilometers and total lane kilometers as lead exposure variables [18].

There have been a number of recent studies evaluating the macro-level effects on bicycle safety. Siddiqui et al. (2012) showed that employment, household income and population are related to cyclist crash frequency [20]. Chen (2015) showed that the parking sign density was positively associated with cyclist crash frequency that the installation of bike lanes did not lead to additional crashes, but a possible increase in the number of cyclists instead [21]. Prato et al. (2015) showed that bike traffic was associated with more cyclist crashes and this association was nonlinear and complied with the “safety in numbers” hypothesis [22]. Amoh-Gyimah et al. (2016) showed that an increase in residential area were positively associated with cyclist crashes [23].

For model development, generalized linear regression (GLM) assuming a non-normal error structure distribution, usually a Poisson or a Negative Binomial (NB) error structure, has become the norm in recent literature, overcoming the limitations of linear regression models and producing better fit to the observed collision data [13] [24]. Previous research [25] has noted that the model form should satisfy the following conditions:

- The model must produce logical results, for example at zero exposure collisions must equal to zero; and,
- In order to use GLM, there must be a recognized link function that can linearize the model form to estimate coefficients during the GLM process [16] [24] [25].

Assuming the error structure of the model follows the negative binomial distribution, the general model form used in previous research is shown in Equation 1.

$$E(\Lambda) = a_o Z^{a_1} e^{\sum b_i X_i} \quad (1)$$

Where:

$E(\Lambda)$  : predicted collision frequency (over 3 years)

$a_o, a_1, b_i$  : model parameters

$Z$  : external exposure variable

$X_i$  : explanatory variables

Following this methodology, the *Pearson*  $X^2$  and scaled deviance statistical measures were used to assess the goodness of fit of the models [24] [25]. To select the model variables, a forward stepwise procedure was used in which the variables were added to the model one by one [26]. A variable was retained in the model based on three criteria: a significant parameter t-statistic, a significant improvement in model fit, and low correlation with other independent variables. The models must meet all goodness-of-fit measures at the 95% confidence level to be used to predict the level of safety in each TAZ. This previous work by Lovegrove and Sayed

(2006) [13] demonstrated the potential use of the developed macro-level CPMs for planners and engineers in proactive road safety applications. In addition, Lovegrove and Sayed (2007) [27] provided a case study on how to use macro-level CPMs in macro-reactive RSIP applications such as black spots programs.

## Black Spot Programs

The objective of reactive Road Safety Improvement Programs (RSIPs), or black spot programs, is to identify and treat locations that are considered hazardous based on the analysis of collision, traffic and roadway data. RSIPs involve the following purposes:

- To identify hazardous *locations* and detect black spots;
- To identify *problems* through location diagnosis; and,
- To identify *solutions* and remedies by finding countermeasures to solve the problems [28].

The assumption of RSIPs is that road design typically plays a significant role in contributing to collision frequency. Sayed et al. (1995) found that road-related factors have caused about 32% of collisions based on analysis of collision data in British Columbia [29]. For that reason, improving the transportation engineering elements of black spots can significantly decrease a proportion of collisions. Properly identifying and ranking black spots for diagnosis and treatment is important to ensure that resources are only spent on areas with highest potential collision reduction (PCR).

A black spot is defined as a location or area that are found to have a significantly high collision potential compared to a group of similar locations, typically through the measure of collision frequency. CPMs' use of an exposure variable corrects for possible frequency bias when comparing different locations with different traffic volumes (exposure levels). To ensure that only truly hazardous locations are identified as black spots, a popular statistical technique is the Empirical Bayes (EB) technique. The EB technique defines the collision probability of the mean collision frequency of a given area as dependent on the observed mean collision frequency of the location and an objective prior distribution based on empirical data from a reference population or from using CPMs. CPMs have the advantage of analyzing collision frequency instead of collision rates and eliminate the need for a very large reference population [24]. In addition, using CPMs allows for location-specific prior distribution to be derived for each location from an imaginary reference population, of which the estimates of mean and variance have shown in previous research to be more reliable than a sample from real reference population [26]. It is therefore vital that the appropriate CPM is selected on the basis of traits examined and type of safety estimate [27].

After the development of CPMs and the selection of the appropriate model the first step is to estimate the safety of the prior distribution,  $E(\Lambda)$ , by using the models to make predictions for each location. The prior distribution variance is shown in Equation 2, and is calculated assuming it follows a gamma distribution.

$$Var E(\Lambda) = \frac{[E(\Lambda)]^2}{\kappa} \quad (2)$$

Where:

$E(\Lambda)$  : predicted collision frequency

$\kappa$  : overdispersion parameter

This model outputs allow for the development of location-specific estimates for  $E(\Lambda)$  and  $Var E(\Lambda)$ .

The following step is to gather local observed collision history data to refine the prior estimate provided by the CPM to define the posterior distribution. This posterior distribution represents how the mean collision frequency varies in a subpopulation of variables having similar traits in terms of traffic, geometry and collision history, and is gamma distributed. The mean of the posterior distribution is in other words the EB safety estimate for location  $i$ ,  $EB_i$ . The mean,  $EB_i$ , and the variance,  $Var(EB_i)$ , of the posterior distribution are:

$$EB_i = E(\Lambda|Y = count) = \left[ \frac{E(\Lambda_i)}{\kappa + E(\Lambda_i)} \right] (\kappa + count) \quad (3)$$

$$Var(EB_i) = Var(\Lambda|Y = count) = \left[ \frac{E(\Lambda_i)}{\kappa + E(\Lambda_i)} \right]^2 (\kappa + count) \quad (4)$$

Where:

$EB_i$  : EB safety estimate for location,  $i$

The final step to identify black spots is to compare the results of the value of each location's safety,  $EB_i$  to the regional average or norm for locations with similar traits. Each location would be considered collision prone if there is a significant probability,  $\sigma$ , usually 0.95 or 0.99, that the EB safety estimate,  $EB_i$  exceeds the specified standard. The location is identified collision prone if the following condition is met:

$$1 - \int_0^{E(\Lambda)} f_{EB}(\lambda) d\lambda = \left[ 1 - \int_0^{E(\Lambda)} \frac{[\kappa/E(\Lambda) + 1]^{(\kappa+count)} \lambda^{(\kappa+count-1)} e^{-[\kappa/E(\Lambda)+1]\lambda}}{\Gamma(\kappa + count)} d\lambda \right] \geq \sigma \quad (5)$$

Where:

$\sigma$  : probability that the EB safety estimates exceed a specific value (usually 0.95 or 0.99)

Once a location has been determined as collision prone, ranking must occur to ensure the locations in most need of treatment are looked at first. Ranking criteria has typically used the potential collision reduction (PCR), based on the difference between expected and observed collision frequency:

$$PCR_i = EB_i - E(\Lambda)_i \quad (6)$$

Following black spot identification and ranking for treatment, a safety diagnosis is performed. First, collision history is analyzed to identify an overrepresentation of clusters of specific collision types by comparing percentages of specific collision types to other similar locations. Second, location specific traits are identified and analyzed to identify potential causes of overrepresented collision types. Once the safety issue has been identified, the next step is to generate a list of potential remedies to decide on the remedy for the specific location. The final choice of remedies to implement will involve engineering judgement and experience [24] [27].

## GUIDELINES FOR MACRO-REACTIVE USE

Macro-reactive black spot analysis often uses the individual TAZ as a unit of analysis instead of an intersection or road segment as for micro-reactive analyses. Macro-reactive guidelines generally follow the traditional micro-reactive methods, however there are some differences in methodology:

- Macro-reactive black spot analysis only requires the aggregate collision history for the zone instead of a collision history for each intersection and road segment; and,
- The macro-reactive analytical unit is the entire TAZ.

The CPMs were developed in this study with variables chosen to meet the needs of the macro-reactive black spot safety evaluation. The scope of this citywide bicycle safety study was to identify Collision Prone Zones (CPZs) and variables that may be affecting bicycle safety.

### Selection of Appropriate Macro-Level CPMs and Variables

To meet the needs of the macro-level bicycle safety evaluation, the minimum number of models was chosen to minimize data extraction effort and computational resources in the analysis while still providing accurate and reliable results. In the development of the models, this study was unique in its use of the leading exposure variable BKT, for which the volume was represented by the bicycle Annual Average Daily Traffic (AADT). In addition, the potential explanatory variables were then grouped into themes of:

- Exposure variables related to bicycle and vehicle kilometers travelled, and therefore collision probability;
- Network variables related to the transportation road network; and,
- Socio-economic and commute (TDM) preference variables.

### Identification and Ranking

The enhanced EB method using CPMs to identify black spots is based on research set out in Sawalha & Sayed (1999) [25], generally following the EB method described in Higle & Witkowski (1988) [30], with modifications to use CPMs. To identify and rank black spots with macro-level CPMs, there must be some adjustments made to the conventional reactive method.

1. The observed local collision history (count) is based on the zonal aggregate, providing the first observational clue on safety.
2. The zonal  $E(\Lambda)$  and  $Var E(\Lambda)$  are calculated to provide the location-specific prior clue to calculated the zonal EB safety estimate.
3. Due to multiple CPMs, there will be several EB safety estimates for each zone resulting in a majority rule when determining a CPZ.

Finally, the calculated zonal  $E(\Lambda)$  and EB safety estimates would differ for each zone resulting in differences in zonal rankings. This can be resolved by using a modified ranking approach, by summing each zone's PCR rankings across all macro-level CPMs to develop a total ranking score for each CPZ. This score will denote which CPZs are most frequently ranked as the least safe and are in need for a diagnosis [24] [27].

## Diagnosis

Once the CPZs have been identified and ranked for treatment, the diagnosis stage to find the cause of the safety problem is begun. Safety issues for CPZs can be diagnosed using a methodology similar to the conventional approach. As with this approach, diagnosis begins by first looking at an overrepresentation of collision patterns: clusters of particular collision types. Using macro-level CPMs method to determine CPZs, an additional indicator can be used: trigger variables from each model that are hypothesized to contribute to the identification of the zone as a CPZ. To identify trigger variables, the value of each variable in the top ranked CPZs is compared with the value of regional averages to understand which variables are triggering CPZ identification. A regional average is the mean of the specific variable value for all zones in the study area. The regional City of Vancouver (COV) statistics (average and standard deviation) for variables used in this study are shown in Table 1. The variable values that are found to be significantly different than the regional statistics are identified as trigger variables. This indicator can be used together with observations of collision patterns and site visits to understand the overall safety issues in each CPZ [27]. The identification of the zonal safety problem is an important step to realize potential suitable remedies.

**TABLE 1: Regional statistics in the City of Vancouver for included variables**

Included Variables	Variable Symbol	Zonal Regional Statistics			
		Average	Standard Deviation	Min	Max
Exposure					
Bicycle Kilometers Traveled (km)	BKT	1047.77	2109.96	0	21462.8
Total "AAA" Bicycle Kilometers (km)	AAAKM	0.59	1.26	0	10.0
Collisions					
Total Bicycle Collisions over 3 years	TB3	12.72	13.49	0	78
Socio-demographic and Commute					
Total Population	POP	4678.11	2648.33	0	10668.0
Total Employment	EMP	3031.23	3337.34	65.0	21189.0
Household Density (household/hectare)	HHD	2058.95	1225.99	0	6163.0
Commercial Land Use (m <sup>2</sup> )	COM	36538.10	45152.60	0	230721
Total Commuters	TCM	921.74	413.46	78.1	2195.8
Network					
Signal Density (signals/hectare)	SIGD	0.143	0.185	0	1.1
Intersection Density (intersections/hectare)	INTD	0.744	0.337	0.1	2.4
Total Bus Stops	BS	14.127	8.159	0	37
Percentage of Residential Street Kilometers (%)	RPKM	56.444	30.635	0	232.2
Total Arterial Street Kilometers (km)	AKM	2.129	1.395	0	7.6
Percentage of Park Area (%)	PMP	7.695	12.750	0	97.3

## Remedy

To match the safety issue with a suitable zone-wide safety remedy, a strategic zonal safety analysis must be conducted when using macro-level CPMs. Examples of possible



remedies could be generated from considering variable theme and diagnosis results, whether the variable theme is network (i.e. number of signalized intersections could be associated with increased collisions), TDM (i.e. increased vehicle commuters could be associated with increased collisions), or socio-demographic (i.e. increased populations is associated with decreased collisions) [27]. After identifying the zonal safety problem using the methodology above, zonal characteristics were analyzed at a micro scale to identify possible remedies. This included factors such as the number of collisions, the quality of the bicycle infrastructure, the topography and the amount of arterials in each zone, for example. Finally, a list of potential bicycle safety countermeasures were analyzed and applied to each zone using engineering judgement.

## CASE STUDY

The macro-reactive guidelines described above were use to conduct a black spot case study to identify, diagnosis and recommend remedies for the most dangerous CPZs in the COV. The land area of the city is about 115 square kilometers. In 2011, the year of this study's focus, the COV population totaled about 600,000 residents, dwelling in almost 265,000 households. Land area in the COV is dense and urban, with an average population density of 5,249 persons per square kilometer [31]. The COV's cycling network in 2011 was composed of about 240 kilometers of separated bike lanes, painted bike lanes, local street bikeways and shared use lanes [32]. In 2011, cycling accounted for approximately 4.4% of trips to work in the COV [33].

The data for this case study involved 3 years of collision data and zonal infrastructure, socio-demographic and commuting trait data:

- The Metro Vancouver regional transportation authority, TransLink, provided geocoded files of TAZs, as well as road, transit and bicycle networks from the year 2013. In addition, TransLink provided EMME/2 transportation model outputs consisting of travel demand (vehicle and transit kilometers travelled, average zonal speed, average, zonal congestion and transportation mode split), as well as socio-demographic and land use data for the base year 2011.
- City of Vancouver (COV) provided exposure and infrastructure variables such as number of signals from their most recent Open Data catalogue, accessed in late 2015. The exposure variable total "AAA" bicycle kilometers refers to the total kilometers of bicycle network that is classified as comfortable and safe for users of All Ages and Abilities (AAA). This classification is typically found for bicycle facilities that are classified as separated bike lanes or off-street paths.
- The Digital Road Atlas (DRA) was used for intersection data for the base year 2011.
- AADT bicycle exposure data was acquired from the COV's bicycle count volume data for the years 2010 and 2011, which was a previously developed comprehensive database of bicycle volume data obtained from the expansion of temporary and permanent bicycle counts [34].
- The Insurance Corporation of British Columbia (ICBC), the province's public automobile insurance company delivered geocoded files of bicycle collision claims in the COV for the years of 2009, 2010 and 2011. Three years of collision data was used to decrease the randomness bias and quantify the relatively uncommon bicycle collision data [13]. The availability of geocoded insurance claim data centralized from ICBC was considered a

great advantage to overcoming any potential incomplete and unreported collision data problems [24].

The aggregation units used for this case study were based on 134 traffic analysis zones (TAZs) used in TransLink's EMME/2 transportation planning model. This level of aggregation was chosen because the TAZ boundaries overlap with both census tracks and municipal boundaries, allowing for easy data integration of current and future demographic data and transportation demand [27]. When analyzing urban form, using aggregation to the traffic analysis zone can be problematic because boundaries tend to be on major transportation corridors and intersections. This thesis assumed that collision data geo-coded near zone boundaries has an influence on adjacent zones and proceeded with a method to assure that collisions are representative for each TAZ. The methodology to deal with the collision boundary effects used exposure ratio methods to aggregate boundary data into adjacent zones. Collisions were aggregated to each zone according to the BKT (leading exposure variable) ratio between adjacent zones.

## Approach and Results

### *Selection of Models*

There were 10 macro-level CPMs developed for the purpose of conducting a black spot case study for the City of Vancouver. The intention was to identify and then rank CPZs, diagnose safety issues and recommend remedies for CPZs in the COV. The collision prediction models developed used exposure, socio-demographic, TDM and network variables as indicators, along with Bicycle Kilometers Travelled (BKT) as the leading exposure variable. Table 2 presents the models that predict total bicycle collisions and their goodness of fit summary statistics. Most models presented showed explanatory variables as significant at a 95% confidence level, except for the explanatory variables of total All Ages and Abilities (AAA) bicycle kilometers (AAAKM), total commuters (TCM) and percentage of residential street kilometers (RPKM), which were significant at a 90% confidence level. As expected, the models showed that increased collisions were positively associated with increased exposure variable BKT. This confirms the intuitive expectation that more bicycle exposure contributes to bicycle-vehicle collisions. However, the exponent of BKT is less than 1.0, indicating that the rate of increase of bicycle collisions reduces as more cyclists use the network. This confirms the safety in numbers concept, which states that an increase in people cycling will result in an increase in safety.

Using BKT as the leading exposure variable, the models were developed among the four themes of exposure, TDM, socio-demographic and network. The exposure variable total All Ages and Abilities (AAA) kilometers (AAAKM) was positively associated with collisions. The TDM and socio-demographic variables total commuters (TCM), total employment (EMP), household density (HHD), commercial land use (COM), and population (POP) were also found to be positively associated with collisions. The network variables signal density (SIGD), intersection density (INTD), bus stops (BS), percentage of residential street kilometers (RPKM), and arterial street kilometers (AKM) were all found to be positively associated with collisions, while percentage of park area (PMP) was found to be negatively associated with collisions.

1 **TABLE 2: Collision Prediction Models and their goodness of fit summary statistics**

Model Form	k	df	SD	Pear- son $\chi^2$	x2 0.05, df	p-value
<b>Exposure</b>						
$0.1843BKT^{0.6438}$	2.25	131	143	125	159	BKT <.0001
$0.1425BKT^{0.7167} \cdot \exp(-0.1213AAA$ $KM + -0.0224PMP)$	2.80	129	143	146	157	BKT <.0001; AAAKM 0.0642; PMP 0.0007
<b>Transportation Demand Management</b>						
$0.1566BKT^{0.6248} \cdot \exp(0.0003TCM)$	2.34	130	143	127	158	BKT <.0001; TCM 0.0672
<b>Socio-Demographic</b>						
$0.2745BKT^{0.5099} \cdot \exp(0.0001COM$ $+ 0.0001EMP)$	2.99	129	144	121	157	BKT <.0001; COM <0.0001; EMP 0.0005
$0.1461BKT^{0.5358} \cdot \exp(0.0001EMP$ $+ 0.0003HHD)$	3.40	129	144	126	157	BKT <.0001; EMP <0.0001; HHD <0.0001
$0.0816BKT^{0.6336} \cdot \exp(0.0001POP$ $+ 0.6203INTD)$	2.57	129	141	124	157	BKT <.0001; INTD 0.0031; POP 0.001
<b>Network</b>						
$0.1068BKT^{0.5900} \cdot \exp(1.6143SIGD$ $+ 0.0409BS)$	3.12	129	142	136	157	BKT <.0001 SIGD <0.0001; BS <0.0001
$0.1776BKT^{0.6069} \cdot \exp(-0.0263PMP$ $+ 0.1851AKM)$	3.24	129	143	139	157	BKT <.0001; PMP <0.0001; AKM <0.0001
$0.083BKT^{0.6094} \cdot \exp(0.6443INTD$ $+ 0.034BS)$	2.74	129	142	118	157	BKT <.0001; INTD 0.0018; BS <0.0001
$0.1835BKT^{0.6342} \cdot \exp(-0.0040RPK$ $M + 0.3645INTD)$	2.40	129	142	121	157	BKT <.0001; RPKM 0.0841; INTD 0.0879

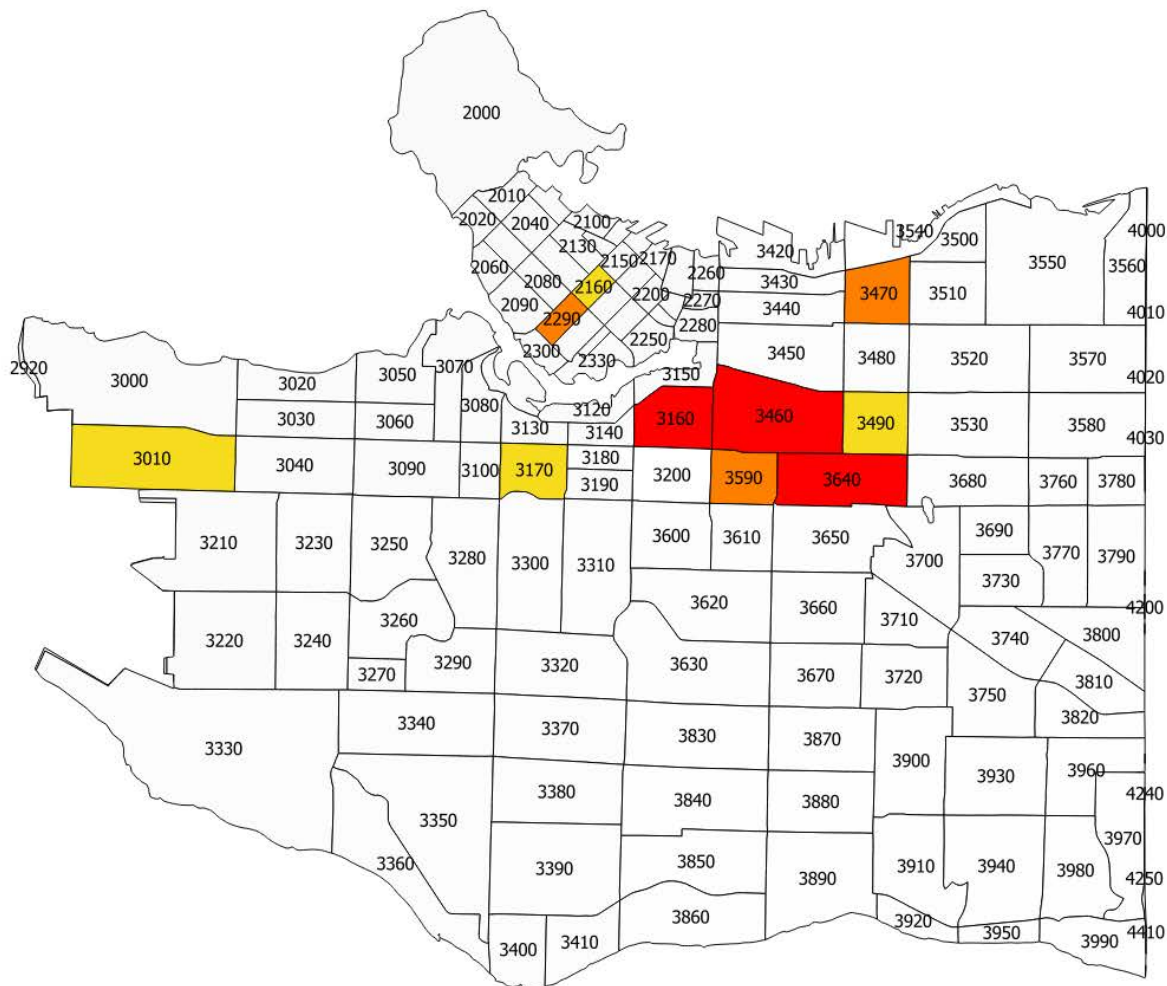
Macro-reactive guidelines were followed for the black spot analysis case study due to the citywide scope of this study to evaluate 134 TAZs at a planning-level safety analysis. All variables that were chosen for the models were considered potential trigger variables for bicycle collisions from the four variable themes. Since the task was in part to identify and rank CPZs, all model groups that were found to be significant using multiple variables were considered.

### Identification and Ranking

The CPMs were used to first estimate the expected location specific collisions for each TAZ,  $E(\Lambda)$ . This clue, together with the observed zonal collision count over three years resulted in 10 EB safety estimates for each zone (one from each of the 10 models). Each of the 10 CPMs identified a range from 7 to 20 collision prone zones at the 99% confidence level. The  $E(\Lambda)$  was used as the reference group norm for comparison to the EB safety estimate to identify the Collision Prone Zones (CPZs) that have the highest potential collision reduction (PCR). The ranking of PCR was based on the difference between expected and observed collision frequency. Using a modified ranking technique, which considered the zonal ranking of all CPMs, the zones most frequently ranked with high PCR were further analyzed for diagnosis and remedy.

The geographic locations of the top 10 ranked CPZs were ranked by a gradient colour scheme: red for most severe top three zones, orange for the fourth to sixth severe zones, and yellow for the seventh to tenth severe zones. An example of the base reference model, which

used only BKT as the exposure explanatory variable, is shown in Figure 1.



**FIGURE 1: Top ten CPZs for the exposure model using BKT**

The ten model groups showed consistency by providing relatively similar top CPZs identification and ranking, with a small variability. With this ranking technique, the three areas with CPZs that scored the worst were then carried forward for diagnosis to identify safety problems and potential solutions.

### *Diagnosis and Remedy*

The top ten zones that were ranked collision prone were analyzed to identify safety problems using two indicator techniques. The values of each CPM's top ranked CPZ variables were compared with their corresponding regional averages in Table 1. These trigger variables, along with the collision frequencies identified the zonal safety issues. Across the models, the results show that collisions are associated with a large variability (defined as either low or high compared to the regional average) of the total All Ages and Abilities (AAA) kilometers (AAAKM), however the top most severe CPZs are associated with low AAAKM. Across the models, collisions are associated with a high variability of the TDM variable total commuters

(TCM), as well as the socio-demographic variables of total employment (EMP), household density (HHD), percentage of commercial land use (COM), and population (POP). The most severe CPZs are found to be associated with high COM and EMP. For infrastructure variables, collisions are associated with a large variability of intersection density (INTD), bus stops (BS), arterial street kilometers (AKM), and percentage of park area (PMP), but did not have an association with variability from the regional averages of the variables signal density (SIGD) and percentage of residential street kilometers (RPKM). The most severe CPZs found collisions to be associated with a low park area percentage, a high number of bus stops, a large amount of arterial streets and a high intersection density.

Subsequently, the top three collision prone areas were carried forward for diagnosis of safety problems and analysis of potential remedies. The three areas under analysis consisted of five CPZs: 3160, 3200, 3460, 3640 and 2290 (in the Mount Pleasant and Downtown neighbourhoods of the City of Vancouver). All areas examined generally showed a higher than average bicycle kilometer travelled, total commuters, commercial land use area, total employment, household density signal and intersection density, bus stop density and total arterial street kilometers; and showed a lower than average percentage of park area and total “AAA” bicycle kilometers.

These clues suggest that the safety issue may be a result of the high density and commercial land use type, coupled with high traffic volume particularly on the arterial routes and high bicycle volume on bicycle routes that are classified as painted bike lanes, shared lanes and local street bikeways. These types of bicycle routes are typically less safe for cyclists and they are not classified as “AAA”. There were many observed collisions along arterial routes or bicycle routes, areas that had the highest bicycle and vehicle volumes. Possible remedies suggested to solve the safety issues in these zones include:

- Update existing bicycle routes to be ranked “AAA,” by installing separated bike lanes or installing traffic calming measures where applicable, depending on vehicle volumes on the route. This includes installing future separated bicycle lanes on the busy arterial routes where cyclists want to travel to access shopping, jobs and residences.
- Increase signage along all routes, and install more local street bikeways onto low roads with low vehicle volumes to raise drivers’ attention to cyclists.
- Continue to address intersection safety through bicycle infrastructure spot improvements such as curb bulges, bicycle refuges, elephant feet, bike boxes and bike signals, especially at high volume intersections. Protected intersections should be prioritized at high vehicle volume locations that also have high bicycle volumes.
- Monitor bicycle infrastructure improvements in the areas, to see whether there is an observed improvement in bicycle safety through a reduction of collisions.

## CONCLUSIONS

In this study, macro-level CPMs were developed for bicycle-vehicle collisions and applied to a case study of the City of Vancouver (COV) at the zonal level, with 134 traffic analysis zones. The purpose of this research has been to develop macro-level CPMs and to use the models as empirical tools for bicycle road safety evaluation and planning. The main motivations of this study came from environmental, public health and safety concerns. To encourage the sustainable transportation mode of cycling, it is fundamental to build a safe and comfortable cycling environment. Addressing the safety of cyclists, there is a need for empirical

tools to evaluate bicycle safety proactively before collisions occur. Bicycle safety research is an emerging field of study, which has yet to reach the detailed extent of vehicular traffic research.

The two main contributions of this study relate directly to the demonstration of the validity of the use of the macro-level CPMs to enhance traditional safety initiatives, through model use in macro-reactive road safety programs. Traditionally, the use of CPMs in road safety improvement programs has been focused on micro-level CPMs, at an intersection or single facility scale. The macro-level models were applied to a case study of the COV to identify, rank, diagnose and remedy CPZs with respect to bicycle safety. Using the information provided by the models, potential safety countermeasures were brought forward for the top three collision prone areas in the city. This case study effectively demonstrates the use of the models to enhance bicycle safety using this valuable safety tool.

For future research the methodology described in this paper could be further improved with research in the following topics:

- The reduction of data needs for the development of macro-level CPMs. Data assembly for reliable statistical model development is intensive. For planners and engineers with a lack of good quality of data, the models could be refined to require fewer variables while retaining model goodness-of-fit. The timeliness of the data is important, as infrastructure planning and design decisions should to be made with up-to-date data. Simplifying data needs would encourage practitioners to develop and apply CPMs and advance road safety in the transportation planning process.
- The macro-reactive black spot steps of (1) diagnosis and (2) remedy were conducted at a preliminary level due to a lack of collision detailed information and the large TAZ size. Improved data quality and smaller aggregation for zones is recommended for future studies in high-density areas.

Continuing to focus research on developing models for vulnerable road users' safety will advance sustainable transportation planning. The safety of cyclists is critical, since bicycle travel has a higher per-mile casualty rate than car travel, yet poses minimal risk to other road users. The study and monitoring of sustainable transportation modes is vital to encourage environmentally friendly transportation options. Important considerations include collecting cycling and pedestrian data at the same level as vehicle data, to be able to accurately monitor volumes and safety. Currently, the largest barrier to increasing bicycle mode split is improving the comfort and safety of riding for cyclists.

## REFERENCES

- [1] World Health Organization, "Global status report on road safety 2013," 2013.
- [2] World Health Organization, "The Global Burden of Disease, 2008," 2011.
- [3] Transport Canada, "Transport Canada Releases First, Systematic Analysis of Cost of Urban Traffic Congestion in Canada," 2006.
- [4] Vijayakumar, N., & Burda, C., "Cycle Cities: Supporting cycling in Canadian cities," Pembina Institute, 2015.
- [5] City of Vancouver, "Cycling Safety Study," 2015.
- [6] ICBC, "Quick Statistics," 2015.
- [7] Chong S, Poulos R, Olivier J, et al., "Relative injury severity among vulnerable non-motorised road users: comparative analysis of injury arising from bicycle-motor vehicle and bicycle-pedestrian collisions," *Accid Anal Prev.*, vol. 42, no. 1, pp. 290-296, 2010.
- [8] Cripton, P., et al. , "Severity of urban cycling injuries and the relationship with personal, trip, route

and crash characteristics: analyses using four severity metrics," *BMJ Open*, 2014.

- [9] Ho, G., & Guarnaschelli, M., "Developing a road safety module for the regional transportation model, technical memorandum one framework," *ICBC*, 1998.
- [10] Lord, D., Persaud, B., & Palmisano, J., "Calibration and transferability of accident prediction models for urban intersections," pp. 57-64, 2002.
- [11] Hadayeghi, A., Shalby, A., Persuad, B. , "Macro-level accident prediction models for evaluating the safety of urban transportation systems," *Transportation Research Board*, pp. 48-61, 2003.
- [12] Ladron de Guevara, F., Washington, S., & Oh, J., "Forecasting crashes at the planning level: simultaneous negative binomial crash model applied in Tucson, Arizona," *Transportation Research Record*, vol. 1897, no. 1897, 2004.
- [13] Lovegrove, G., & Sayed, T., "Using macro-level collision prediction models in road safety planning applications," *Transportation Research Record. 1950*, pp. 73-82, 2006.
- [14] Lovegrove, G., Lim, C., & Sayed, T., "Using Macro-Level Collisions Prediction Models to Conduct a Road Safety Evaluation of a Regional Transportation Plan," *Transportation Research Board 87th Annual Meeting* , 2008.
- [15] Cheung, C., Shalaby, A., Persaud, B., & Hadayegy, A , "Models for safety analysis road surface transit ," *Transportation Research Record 2063* , pp. 168-175, 2008.
- [16] Quintero, L., Sayed, T., & Wahba, M., "Safety models incorporating graph theory based transit indicators," *Accident Analysis and Prevention*, vol. 50, pp. 635-644, 2013.
- [17] Sun, J., & Lovegrove, G., "Comparing the road safety of neighbourhood development patterns: traditional versus sustainable communities ," *Canadian Journal of Civil Engineering* , vol. 40, no. 1, pp. 35-45, 2013.
- [18] Wei, F. & Lovegrove, G., "An empirical tool to evaluate the safety of cyclists: Community based, macro-level collision predicition models using negative binomial regression," *Accident Analysis & Prevention*, vol. 61, pp. 129-137, 2013.
- [19] Wei, F., Alam, A., & Lovegrove, G., "Macro-Level Collision Prediction Models Related to Bicycle Use," in *ICTIS 2011.*, 2011.
- [20] Siddiqui, Chowdhury, Mohamed Abdel-Aty, and Keechoo Choi, "Macroscopic spatial analysis of pedestrian and bicycle crashes," *Accident Analysis & Prevention*, vol. 45, pp. 382-391, 2012.
- [21] Chen, Peng, "Built environment factors in explaining the automobile-involved bicycle crash frequencies: a spatial statistic approach," *Safety Science*, vol. 79, pp. 336-343, 2015.
- [22] Prato, Carlo Giacomo, et al. urnal of Transportation Safety & Security just-accepted, 2015, pp. 00-00., "Infrastructure and spatial effects on the frequency of cyclist-motorist crashes in the Copenhagen Region," *Journal of Transportation Safety & Security*, vol. 8, no. 4, pp. 346-360, 2016.
- [23] Amoh-Gyimah, Richard, Majid Sarvi, and Meead Saberi, "Investigating the Effects of Traffic, Socioeconomic, and Land Use Characteristics on Pedestrian and Bicycle Crashes: A Case Study of Melbourne, Australia," *Transportation Research Board 95th Annual Meeting. No. 16-1931*, 2016.
- [24] Lovegrove, G., *Road Safety Planning: New Tools for Sustainable Road Safety and Community Development*. Germany: Verlag Dr. Muller, 2007.
- [25] Sawalha, Z., & Sayed, T. , "Accident Prediction and Safety Planning Models for Urban Arterial Roadways," 1999.
- [26] Sawalha, Z., & Sayed, T., "Traffic accident modeling: some statistical issues," *Canadian Journal of Civil Engineering*, vol. 33, no. 9, p. 1115, 2006.
- [27] Lovegrove, G. & Sayed, T., "Macrolevel Collision Prediction Models to Enhance Traditional Reactive Road Safety Improvement Programs," *Transportation Research Record: Journal of the Transportation Research Board, No. 2019*, pp. 65-73, 2007.
- [28] Sayed, T., "New and Essential Tools, Part 1: Selecting Sites for Treatment," *Proceedings from*

*Seminar on New Tools for Traffic Safety*, pp. 1-13, 1998.

- [29] Sayed, T., W. Abdelwahab, and F. Navin, "Application of Fuzzy Pattern Recognition to the Identification of Accident Prone Locations," *Journal of Transportation Engineering*, vol. 121, no. 3, pp. 352-358, 1995.
- [30] Hagle, J. & Witkowski, J., "Bayesian Identification of Hazardous Locations," *Transportation Research Record 1185, Journal of the TRB*, pp. 24-36, 1988.
- [31] Statistics Canada, "Focus on Geography Series, 2011 Census," Statistics Canada Catalogue no. 98-310-XWE2011004, Ottawa, ON, 2012.
- [32] City of Vancouver, "Cycling in Vancouver," 2015.
- [33] Statistics Canada, "Canadian National Household Survey," 2011.
- [34] Esawey, M., Lim, C., Sayed, T., & Mosa, A., "Development of Daily Adjustment Factors for Bicycle Traffic," *Journal of Transportation Engineering*, vol. 139, pp. 859-871, 2013.