Uncertain Data and Confident Decisions: Winnipeg's Safety Performance Functions and Network Screening Project

Bartholomew Ghanney, Craig Milligan, Rebecca Peterniak, Mark Hearson

Abstract

This paper presents modelling techniques for addressing three data reliability issues encountered in the City of Winnipeg safety performance functions (SPF) and network screening project: (A) uncertain traffic volume data, (B) non-uniform collision under-reporting linked to segment length, and, (C) missing traffic volume data for minor roads of an intersection. The first issue relates to uncertain traffic traffic volume data. The City of Winnipeg uses short-term count stations and has a form of traffic data known as the weekday average daily traffic (WADT). The issue of data uncertainty is addressed by estimating the amount of error in the volume data that goes into the SPF and accounting for this error in the SPF development using a Monte Carlo-based modelling approach. The simulation approach maps traffic volume uncertainty to SPF parameter uncertainty. The second issue relates to non-uniform under-reporting linked to segment length. The non-uniform under-reporting of segment collisions resulted in a global and localized model bias. We introduce a new technique to detect, quantify, and correct this bias by using residuals analysis to stratify the population. The third issue relates to modelling intersections with missing minor street flow volumes. We apply an approach that uses the functional class of the intersections as proxies for the missing flow volumes. For each issue, we demonstrate quantitatively that good modelling results can be obtained despite input data limitations. Key conclusions are: (A) for traffic volume measurement errors of up to 30%, Monte Carlo analysis shows that the ability to create reliable SPFs is not affected; (B) residuals analysis to stratify a population according to non-uniform under-reporting can essentially eliminate global and local model bias, and (C) using a readily available proxy for a missing predictor variable can improve predictive ability by almost 50% (measured by mean absolute deviation) when compared to omitting that predictor variable entirely.

1. Introduction

Safety performance functions (SPFs) are mathematical models that predict the expected frequency of collisions together with its variance occurring for a given road facility as a function of traffic volumes and other relevant potential road attributes. SPFs primarily serve a network screening purpose to aid in the detection of high risk collision prone locations in a road network (Hauer, Kononov, Allery, & Griffith, 2002). SPFs have many other functions such as interactive design support (Ng & Sayed, 2004), countermeasure selection analysis (Li, Carriquiry, Pawlovich, & Welch, 2008), and before-and-after treatment evaluations (Persaud & Lyon, 2007). The use of SPFs is a core component of *Advanced General Purpose Network Screening* as defined by the Transportation Association of Canada, which many urban and provincial jurisdictions are currently adopting or have already adopted. SPFs are also foundational to the AASHTO *Highway Safety Manual*.

Despite the significant role of SPFs in network screening, questions regarding data reliability and data limitation issues can cause professionals to question screening results or the merit of developing an SPF-based screening program. To our knowledge, research on the impact that data limitations bring to SPFs and network screening results has been very limited. Real traffic volume data most often has to be estimated from low quality data, for this reason, the precision

of the estimated collisions are affected (Davis, 2000). El-Basyouny and Sayed (El-Basyouny & Sayed, 2010) proposed a measurement error negative binomial approach which reduces the bias in predicted crashes when traffic volumes have measurement uncertainties, however, the extent of influence of measurement error in traffic flow volumes on parameter robustness and reliability have rarely been discussed in the literature.

Under-reporting bias is an issue that is commonly discussed but not significantly explored in terms of SPF impact. If the bias is uniform, it can be either ignored in screening or accounted for with a scaling parameter. However, if the bias is non-uniform across time or the network, the modelling response is more complicated. In capturing the segment data for the City of Winnipeg SPF project, segment collisions were under reported to a higher degree for relatively shorter segments. This unique data recording challenge introduced model bias as collision count under reporting varied with respect to segment length.

The standard practice in modelling intersection collisions has been to use both the minor and major traffic flow volumes as explanatory variables, usually the model form usually expressing them as some kind of product to account for the quantity of potential conflicts. In Winnipeg, as is the case with many jurisdictions, the traffic count information for the minor road intersection is often not available. Information on how to model intersections with complete flow volume information abound in the literature (Dixon et al., 2015; Lyon, Haq, Persaud, & Kodama, 2005). Dixon et al. (2015) present approaches to estimate the minor street volume using other auxilliary information based on data collected as part of the SPF development. However, in the absence of the relevant data required for estimating the minor street volume, there has not been any extensive documentation in the literature that we know of on how to treat intersections with only one street volume.

This study proposes a methodology that addresses data reliability issues by quantifying the amount of error in the volume data that goes into the SPF and accounting for this error in the SPF development using a Monte Carlo-based modelling approach. The simulation based approach maps traffic volume uncertainty to SPF parameter reliability. In the situation of non-uniform under-reporting, we propose a method that quantifies the model bias present and introduce a new modelling strategy that accounts for the range of segment length in the model structure. The cumulative residual plot is used to monitor the bias correction. For intersections with missing street volumes, we propose a functional class modelling approach which uses the class of the intersection as a proxy for the unknown traffic flow volume information. The impact of the functional class model approach on predictive ability is assessed by comparing the proposed approach with the standard approach by emulating the statistical tests for model validation in (Young, Park, & Eng, 2012).

2. Objectives, Scope, and Organization of Paper

The geographic scope of the paper consists of SPFs developed for the City of Winnipeg regional street network. The temporal scope consists of collision data from 2012 and 2013. The modelling covered 2374 segments, and 580 intersections with only one street volume. This paper is meant to provide a high level overview for practitioners of the three techniques to deal with uncertainty and not an in-depth proof of the techniques for the academic community, so complex modelling jargon is kept to a minimum. The overall objective is to describe how, with the right techniques, confident decisions can be made even when the input data has serious limitations. The rest of the paper is organized according to the three data issues considered. Each issue has its own

methodology, results, and sub-conclusion. After the three sections dealing with individual data issues, and concluding discussion is provided.

3. Data Issue A - Uncertainty in Traffic Volume Data

3.1. Methodology

3.1.1. Methodology for Model Form

A typical intersection SPF model form is given below:

$$N_{intersection} = \beta_0 \times V_1^{\beta_1} \times V_2^{\beta_2} + \varepsilon$$

where $N_{intersection}$ denotes the predicted counts of collisions, V_1 the minor street volume for the intersection under consideration and V_2 the major street volume. The estimable parameters of the model are β_0, β_1 and β_2, and ε denotes the error associated with the model which is assumed to have a negative binomial error structure. We used this model form for this study.

3.1.2. Methodology for Introduction of measurement error

Monte Carlo (MC) simulation is a statistical routine that generates artificial samples in a random manner using a set of parameter specifications. The statistical technique was used by (Torres, Lechón, & Soto, 2012) to toggle between different scenarios of traffic accident paths and public strategies to understand how traffic accidents evolve with time. (Miranda-Moreno, Lord, & Fu, 2008) also used MC to determine the best way of incorporating prior information in road safety analysis. The original data considered for the study was assumed to be the ideal data with error free traffic flow volumes. An established level of measurement error was introduced in the flow volumes of the original data to create a new data with known level of induced measurement error using the random truncated normal distribution.

The data perturbation was done using the random truncated normal distribution function *rtnorm()* in the *msm* package in *R* (Jackson & Jackson, 2015). The truncated normal distribution with a minimum truncation point at zero was used to describe the random variation in the traffic flow volumes (Burkardt, 2014). The *rtnorm()* function requires three arguments, namely: the sample size, the mean and the standard deviation. We supply the mean to be the original traffic volume and the standard deviation to be a proportion (ε) of the original traffic volume. The ε considered ranged from 5% to 50% of the original traffic volume in steps of 5% making ten in all for the different data scenarios considered. For any original traffic volume, we generate a new traffic volume from the truncated normal distribution with a sample size of one, and with mean that is equal to the original traffic volume that we seek to perturb, and standard deviation that is equal to a proportion of the original traffic volume. This perturbation of the original traffic volumes was carried out for each traffic volume in the original data whilst maintaining the number of collision counts. A model fit was then obtained based on the new sample with perturbed traffic volumes and the results stored in R. This simulation process was repeated 500 times with all the estimates of the parameters obtained and stored in R. Different cases of the sample size were considered to determine the effect of sample size on parameter reliability. In the first simulation process we considered the whole sample size, then only 75%, 50%, 25%, 20%, 15%, 10% and 5%. In all, for each reference SPF, we re-estimated that SPF 40,000 times (10 measurement error levels)x(8 sample size levels)X(500 MC simulations), recording the change in SPF parameters each time.

3.1.3. Methodology for measurement error evaluation

To assess the impact of the introduced traffic volume measurement errors on the estimated parameters in the SPF, we define the traffic volume uncertainty multiple, which is the ratio of the sample standard deviation in the estimates generated from the *Measurement Error Monte Carlo* ($MEMC$) samples to the standard error inherent in the corresponding estimate for the *reference single model* (SIM).

$$\text{Traffic uncertainty multiple(UM)} \; = \; \frac{S.D(estimates)_{MEMC}}{S.E(estimate)_{SIM}}$$

When $UM < 1$ it suggest the mapped uncertainty from flow volumes is less than the uncertainty already existing in the model with reference (perfect) volumes.

3.1.5. Methodology for Assessing Precision

To assess the precision of estimates when traffic volume measurement error is introduced we use the root mean square error($RMSE$). We compare the ($RMSE$) based on the model using the original data with perfect volumes and the average $RMSE$ based on the Monte Carlo induced measurement error samples. We denote the average $RMSE$ as $ARMSE$. The $RMSE$ and the $ARMSE$ are defined as $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n}\left(\widehat{N_i} - N_i\right)^2}$ and $ARMSE = \frac{\sum_{i=1}^{500} RMSE_i}{500}$ respectively where N_i and $\widehat{N_i}$ denotes the observed and predicted collision counts respectively, n represents the sample size and $RMSE_i$ is the root mean square error obtained from the $i-th$ simulation.

3.2. Results for addressing uncertainty in traffic volume data with Monte Carlo simulations

The results of the traffic volume uncertainty assessment are presented in two parts. First, a set of uncertainty multiple graphs are presented (graphs for the other sample scenarios are presented in the appendix). Second, two tables that assess the effect of sample size on uncertainty are also presented. The graphs show how for lower sample sizes or for higher traffic volume uncertainty, the degree of uncertainty caused by traffic volume errors goes up but is always much less than the degree of uncertainty that is already inherent in the model.

Figure 1: Traffic uncertainty multiple (UM) against traffic volume measurement error (ε) for intersections.
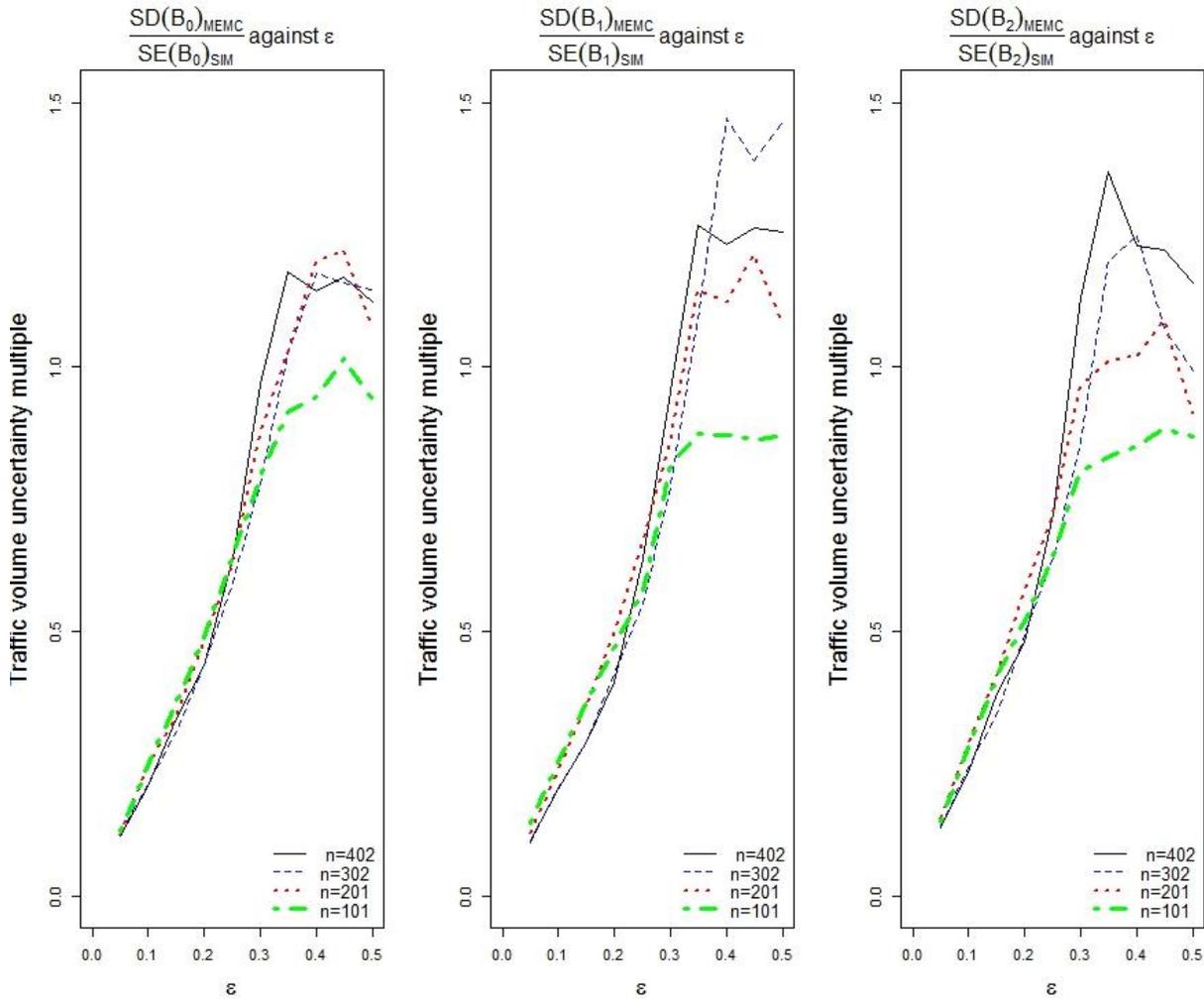
Notes: n denotes the sample size and $\dfrac{S.D(estimates)_{MEMC}}{S.E(estimate)_{SIM}}$ is the traffic volume uncertainty multiple. $MEMC$ indicates measurement error Monte Carlo samples. SIM indicates the single model based on original data and ε is the traffic volume measurement error. The intersection model has the form $N_{intersection} = \beta_0 \times V_1^{\beta_1} \times V_2^{\beta_2} + \varepsilon_m$, where $N_{intersection}$ denotes the predicted counts of collisions, V_1 the minor street volume for the intersection under consideration, V_2 the major street volume.

Table 1: The parameter estimates with their standard errors and root mean square error based on the single data model.

Sample size (n)	$RMSE$	$\widehat{\beta_0}$	$SE(\widehat{\beta_0})$	$\widehat{\beta_1}$	$SE(\widehat{\beta_1})$	$\widehat{\beta_2}$	$SE(\widehat{\beta_2})$
402	0.6623	-9.819	0.6650	0.5484	0.0409	0.8216	0.0649
302	0.6336	-9.430	0.7112	0.5829	0.0447	0.7574	0.0714
201	0.5498	-9.895	0.8873	0.6286	0.0744	0.7649	0.0957
101	0.5228	-9.428	1.3190	0.6616	0.1156	0.6957	0.1230
80	0.5262	-9.567	1.4236	0.6522	0.1230	0.7219	0.1376
60	0.5219	-10.743	1.5229	0.6505	0.1350	0.8454	0.1601
40	0.5848	-10.936	1.9648	0.6055	0.1826	0.9026	0.2150
20	0.4799	-12.428	2.2434	0.3164	0.2141	1.3257	0.2838

Notes: RMSE=Root mean square error, $\widehat{\beta_0}$= estimate of parameter β_0, SE=Standard error

Table 2: The average root mean square errors for the different error levels and for each sample size.

Traffic volume measurement error (ε) in %	RMSE by Sample size (n)							
	402	302	201	101	80	60	40	20
5	0.6612	0.6325	0.5526	0.5248	0.5284	0.5250	0.5863	0.4825
10	0.6582	0.6325	0.5614	0.5312	0.5361	0.5318	0.5943	0.4904
15	0.6566	0.6305	0.5762	0.5410	0.5473	0.5464	0.6044	0.5057
20	0.6646	0.6377	0.5972	0.5587	0.5613	0.5611	0.6220	0.5243
25	0.6849	0.6781	0.6298	0.5802	0.5824	0.5799	0.6425	0.5458
30	0.7262	0.6946	0.6743	0.6058	0.6085	0.6124	0.6681	0.5699
35	0.7817	0.7552	0.7138	0.6296	0.6353	0.6427	0.6941	0.6026
40	0.8203	0.8080	0.7554	0.6526	0.6572	0.6652	0.7156	0.6225
45	0.8576	0.8425	0.7782	0.6652	0.6718	0.6730	0.7266	0.6378
50	0.8736	0.8588	0.8022	0.6758	0.6816	0.6947	0.7418	0.6619

Note: RMSE=Root mean square error

3.3. Sub conclusion for addressing uncertainty in traffic volume data with Monte Carlo simulations

Figure 1 shows a graph of mapped uncertainty multiple against input measurement error levels. The three graphs represent the SPF parameters β_0, β_1 and β_2. The response of each of the parameters to uncertainty in flow volumes is the same. The graph shows the uncertainty multiple for is less than one for traffic data errors of up to about 30%, meaning that up to this point, the uncertainty inherent in the parameters themselves is bigger than the possible uncertainty that may arise as a result of measurement errors in flow volumes. Table 1 reports RMSE for the model with reference flow volumes and Table 2 report the ARMSE for the different input measurement

error levels for each sample size considered. Even with an input measurement error level as high as 30%, the prediction precision of the MEMC model is not very different from the reference model. From this analysis, we can conclude that existence of significant measurement errors in traffic flow volumes does not necessarily impede the development of reliable SPFs.

4. Data Issue B – Non-Uniform Under-Reporting of Segment Collisions

Collision data locations in Winnipeg are coded manually by data technicians working for a public insurance agency based on reports given by either the police or the parties involved. City of Winnipeg staff subsequently accept or reject each record based on the quality of the location information. This process is inherently subject to error through no fault of either agency or technicians in the agencies. As a result of this process, collision data in Winnipeg is subject to three kinds of under-reporting:

- Under-reporting of collisions where neither the police nor the public insurance company was notified. This under-reporting is assumed to be uniform throughout the network.
- Under-reporting of collisions arising from rejection of records due to location information upon transfer from the public insurer to the City of Winnipeg. This under-reporting is assumed to be uniform throughout the network.
- Under-reporting of collisions at specific locations where, due to artifacts of the network segmentation and the nature of the interaction between data analysts and involved parties when describing a collision location. This under-reporting is non-uniform throughout the network and depends on segment length. We found a tendency for collisions to be assigned to a neighbouring segment or intersection that increases as segment length decreased.

The third type of under-reporting introduces a level of non-homogeneity in the data. As a result, when the population is analyzed as a whole, various forms of model bias are manifested. The techniques presented here show a way to detect and reduce this bias.

4.1. Methodology

The basic functional form for predicting segment collisions has the form,

$$N_i = \beta_0 \times (WADT_i)^{\beta_1} \times L_i + \varepsilon_i$$

where ε_i denotes the errors of the predictive model (SPF), $WADT$ the weekly average daily flow of traffic, L_i the segment length, N_i collision count with β_0 and β_1 being the estimable parameters of the model. The ε_i's are assumed to have a negative binomial error structure. In calculating the residuals ε_i of the model, we correct for possible systematic bias that may arise as a result of non-homogeneity of the data by a positive term CF. The CF may also be viewed as a calibration factor that becomes necessary for the functional form of the model adopted. We define CF as

$$CF = \frac{\sum_{i=1}^{n} N_i}{\sum_{i=1}^{n} E(N_i)}$$

The CF can be used to adjust the predicted counts up when there is under prediction and down when there is over prediction respectively. This regulatory behavior of the CF stabilizes the residuals. The CF is a measure of the global systematic bias in a model in that if $CF < 1$ then global model over prediction bias is present and when $CF > 1$ model under prediction bias is present. Thus the extent of departure of the CF from 1 is a measure of bias introduced in the

model; we have found that large departures from 1 (e.g. a CF of 0.8) obtained in exploratory modelling are often an indicator of population non-homogeneity and the departures can be reduced by identifying sub-populations and modelling accordingly.

We adopt a modelling strategy that takes into account the range of the segment length being considered by finding population breakpoints using cumulative residual (CURE) plots. The ranges of segment length were chosen such that satisfactory CURE plots were obtained against $WADT$ and segment length L. Hauer and Bamfo (Hauer & Bamfo, 1997) demonstrate how CURE plots can be applied to safety performance functions, and a brief summary relevant to the current paper is provided here. Every data point used to fit the model has a corresponding residual equal to the observed value minus the fitted value. If the observation is less than the fitted value, the model over-predicted for that data point and the residual will be negative. In the same way, a positive residual indicates a case of under-prediction. To develop a CURE plot, we sort all data points and their residuals in the order of any given predictor variable (e.g. length or volume) and for each data point, we calculate the cumulative sum of all residuals up to that point. The cumulative sum of residuals plotted against the predictor variable is the CURE plot. If the model is unbiased, the CURE plot will oscillate around zero as over-predictions are typically followed by under-predictions of a similar magnitude on a random basis. When the CURE plot trends upwards or downwards, it signifies bias towards under prediction or over-prediction, respectively, along the range of the explanatory variable for which the trend is demonstrated. Figure 2 shows an example of a bad CURE plot that reveals a severe global over prediction bias as well as a good CURE plot that, despite some mild trends, shows a relatively bias-free model.

Figure 2: Example of a good and bad CURE plot
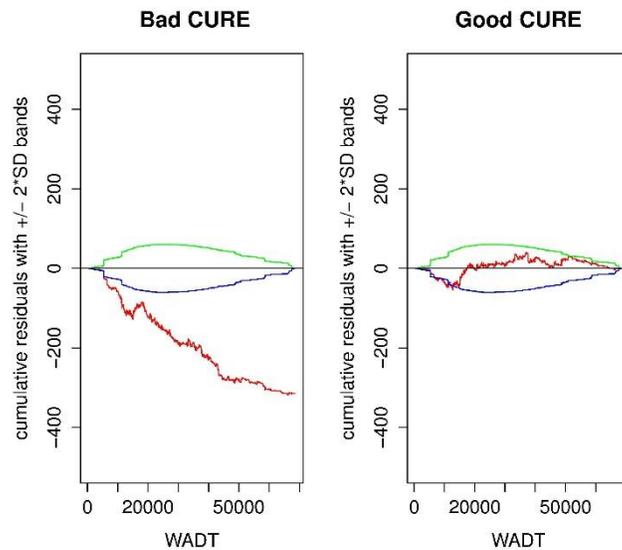
Our initial modelling for segment data showed areas of global and localized bias. We divided all the all the road segments in the street network into mutually exclusive categories such that each category considered on its own with its corresponding collision data provided a satisfactory CURE plot against $WADT$ and L. Generally, suppose $R_1, R_2, R_3, \dots, R_n$ are the ranges of segment lengths adopted, then the desired modelling procedure suggests,

$$logN_i = \beta_0 + \beta_1 R_1 + \beta_2 R_2 +, \dots\dots + \beta_{n-1} R_{n-1} + \log(WADT_i) + offset(L_i) + \varepsilon$$

$$\widehat{N_i} = \widetilde{\beta_0} \times WADT_i{}^{\beta_n} \times L_i \times \exp(\beta_1 R_1 + \beta_2 R_2 +, \ldots\ldots +\beta_{n-1} R_{n-1})$$

where $\widetilde{\beta_0} = \exp(\beta_0)$. This model form is the basic segment model structure but now augmented with an additional multiplicative term that adjusts the estimated number of collisions accordingly depending on the range of segment length being considered. The condition $R_1 + R_2 + \cdots + R_n = 1$ is always satisfied and the estimable parameters of the model are $\beta_0, \beta_1, \beta_2, \ldots\ldots, \beta_{n-1}$ and β_n. Eight ranges $R_1, R_2, \ldots\ldots, R_8$ were considered for the segments data in the Winnipeg SPF project. In the simple case where only eight ranges are considered the general model form reduced to the simple case:

$$\widehat{N_i} = \widetilde{\beta_0} \times WADT_i{}^{\beta_8} \times L_i \times \exp(\beta_1 R_1 + \beta_2 R_2 +, \ldots\ldots +\beta_7 R_7)$$

4.2. Results for Non-Uniform Under-Reporting of Segment Collisions

We obtained a model based on the basic functional form for segment collisions. The modelling was repeated using the segment model with population stratification. Table 3 shows the parameter estimates for the two modelling approaches, and Figures 3 and 4 show the resulting CURE plots. The CF is .7840 in the un-stratified model, which shows a large global over prediction bias, and is .9996 in the stratified model, which shows almost no global bias. The CURE plots likewise show a major reduction of bias by stratifying the models according to segment length.

Table 3. Estimated parameters for the un-stratified model with the CF value

Nature of population	$\widehat{\beta_0}$	$\widehat{\beta_1}$	CF	Sample size
Un-stratified	-4.4427	0.5668	0.7840	2374

Notes: CF= Calibration factor

Table 4. Estimated parameters for the stratified model with the CF value

Parameter	Estimate	CF
$\widehat{\beta_0}$	-5.5686	0.9996
$\widehat{\beta_1}$	1.7563	
$\widehat{\beta_2}$	1.3282	
$\widehat{\beta_3}$	1.1646	
$\widehat{\beta_4}$	0.0983	
$\widehat{\beta_5}$	0.7851	
$\widehat{\beta_6}$	0.2840	
$\widehat{\beta_7}$	0.7216	
$\widehat{\beta_8}$	0.5759	

Notes: CF= Calibration factor

Figure 3: Cure plot against segment length before stratification

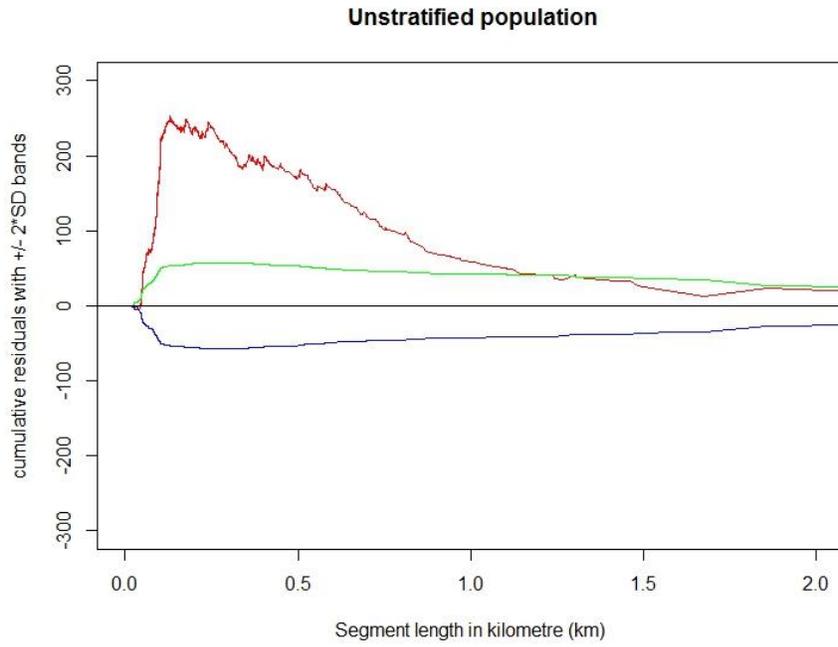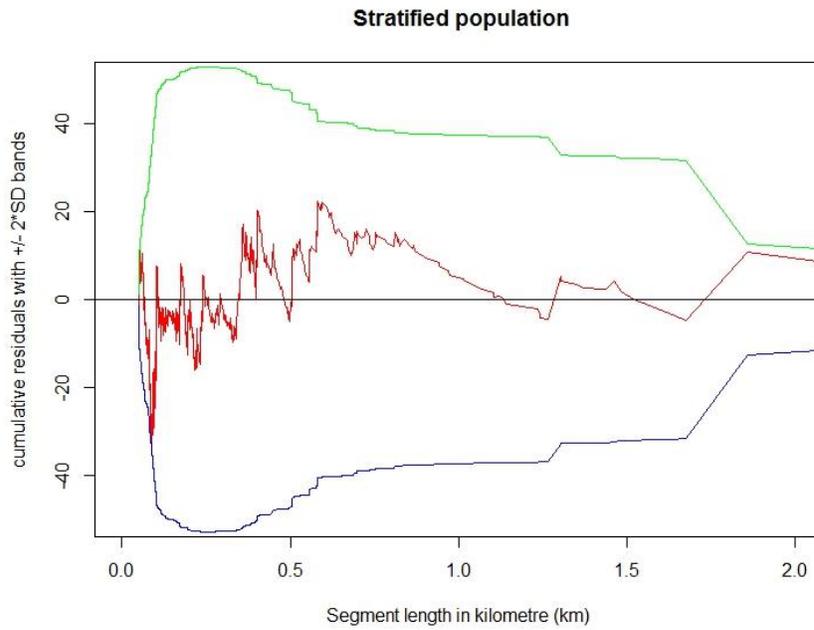### Unstratified population

Stratified population

Figure 4: Cure plot against segment length after stratification

4.3. Sub conclusion for addressing non-uniform underreporting of segment collisions

The results show that the proposed modelling approach is effective in correcting for possible systematic model bias. Despite the non-uniform under-reporting for segment collisions, when the population is stratified such that a satisfactory CURE with respect to the predictor variable linked to under-reporting is obtained, the bias due to the non-uniform under reporting is corrected.

5. Data Issue C: Missing Minor Road Volumes and Use of Class as a Proxy Variable

In modelling intersection collision counts in a street network, the major and minor traffic flow volumes constitute the primary input for the standard intersection model form. However, in the Winnipeg case, many intersections had only a major street volume with the minor street volumes missing or unavailable. This typically leaves two options: these intersections are either rejected for modelling and screening or they are modelled together as one population now with only one street volume as the predictor variable. Simply removing the minor road volume as a predictor variable, however, results in high prediction errors which are largely attributable to the population non-homogeneity that arises due to collapsing together intersections of different functional classes into a single population for modelling. We introduce a new modelling approach that uses the functional class of the intersections as proxy for the missing street flow volumes. We compare the impact on predictive ability of this approach to the standard model approach.

5.1. Methodology

Let N_i $(i = 1,2,3 \ldots \ldots, n)$ denote the predicted count of collision, V_1 the minor street traffic flow volume, V_2 the major street traffic flow volume, and ε_i the error associated with the model. The errors of the model are assumed to have a negative binomial error structure. The form of the basic intersection model is shown below:

$$N_i = \beta_0 \times V_{1i}^{\beta_1} \times V_{2i}^{\beta_2} + \varepsilon_i,$$

The estimable parameters of the model are β_0, β_1, and β_2. The functional form of the model can be assessed using the cumulative residual plots developed by (Hauer & Bamfo, 1997). Other forms of validation checks are: (1) statistical significance of the estimated parameters, (2) parameter logic, (3) model deviance, and (5) model over dispersion. The application of this model is valid if information for both minor and major street volumes are available. When V_1 is unavailable, a single model using only V_2 for all intersections is sometimes used.

The SPF can also be interpreted as an estimate of a best fit function that can be used to predict collisions for known traffic flow volumes. The precision of the estimated collision counts improves with increasing homogeneity in the traffic flow volumes. This level of homogeneity is reduced when an aggregate global model is considered for all intersections with only one street volume. Figure five shows an SPF function fit through a plot of all the intersection collisions. Residuals obtained are larger when considering the intersections together. The functional class approach disaggregates the model into sub populations, where the functional class is used as a proxy for the unknown traffic volume.

Figure 5: Plot of all intersections against traffic flow volumes

All intersersections
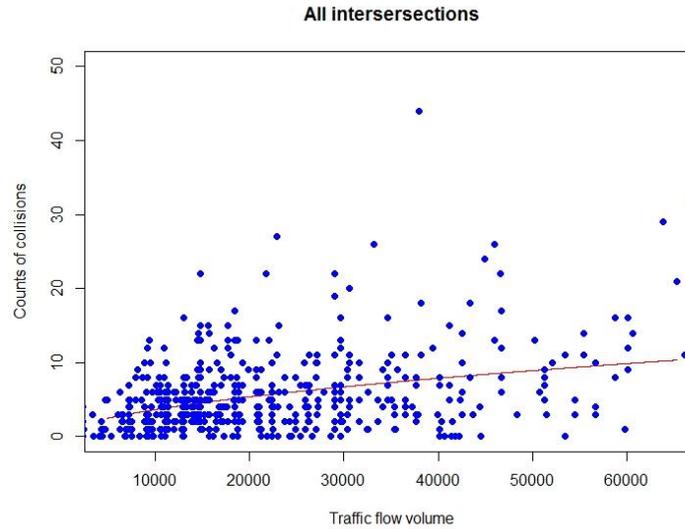
Traffic flow volumes are a function of the intersection class. For instance, collisions on Arterial-Local (AL) intersections are generally higher than collisions on Local-Local (LL) intersections largely due to the fact that the traffic flow volumes at AL intersections are higher. In the functional class modelling approach the population of intersections with only one street volume are stratified according to their class, thereby inducing homogenous sub-populations for modelling. The model form is shown below,

$$N_i = \beta_0 \times V_i^{\beta_1} \times \left((\beta_{LL} \times LL) + (\beta_{AC} \times AC) + (\beta_{AL} \times AL) + (\beta_{CL} \times CL)\right) + \varepsilon_i,$$

where the condition $AL + AC + CL + LL = 1$ is always satisfied, and A denotes arterial, C denotes collector, and L denotes Local. ε_i is the error of the model which assumes a negative binomial error structure. The functional classes are either zero or one in the model structure. In the simple case where the intersection being considered is (AL), the functional class model reduces to

$$N_i = \beta_0 \times V_i^{\beta_1} \times \left((\beta_{AL} \times AL)\right) + \varepsilon_i,$$

where $AL = 1$ with $AC = 0, CL = 0$ and $LL = 0$. There is no limit to the number of functional classes that can be considered, so long as the form of the model is maintained the functional class model is still applicable. With the functional class approach the residuals decrease considerably. If we think of the SPF as a best fit curve, then the functional class model considers four best curves instead of one for all the intersection collisions. Ideally, the residual when only one best function is used is larger compared to when four different SPFs developed based on functional classes of intersections are used. Figure six below shows four different SPF graphs for the different functional classes considered. The implication of the figure is an improved fit.

Figure 6: Plot of all intersections against traffic flow volumes
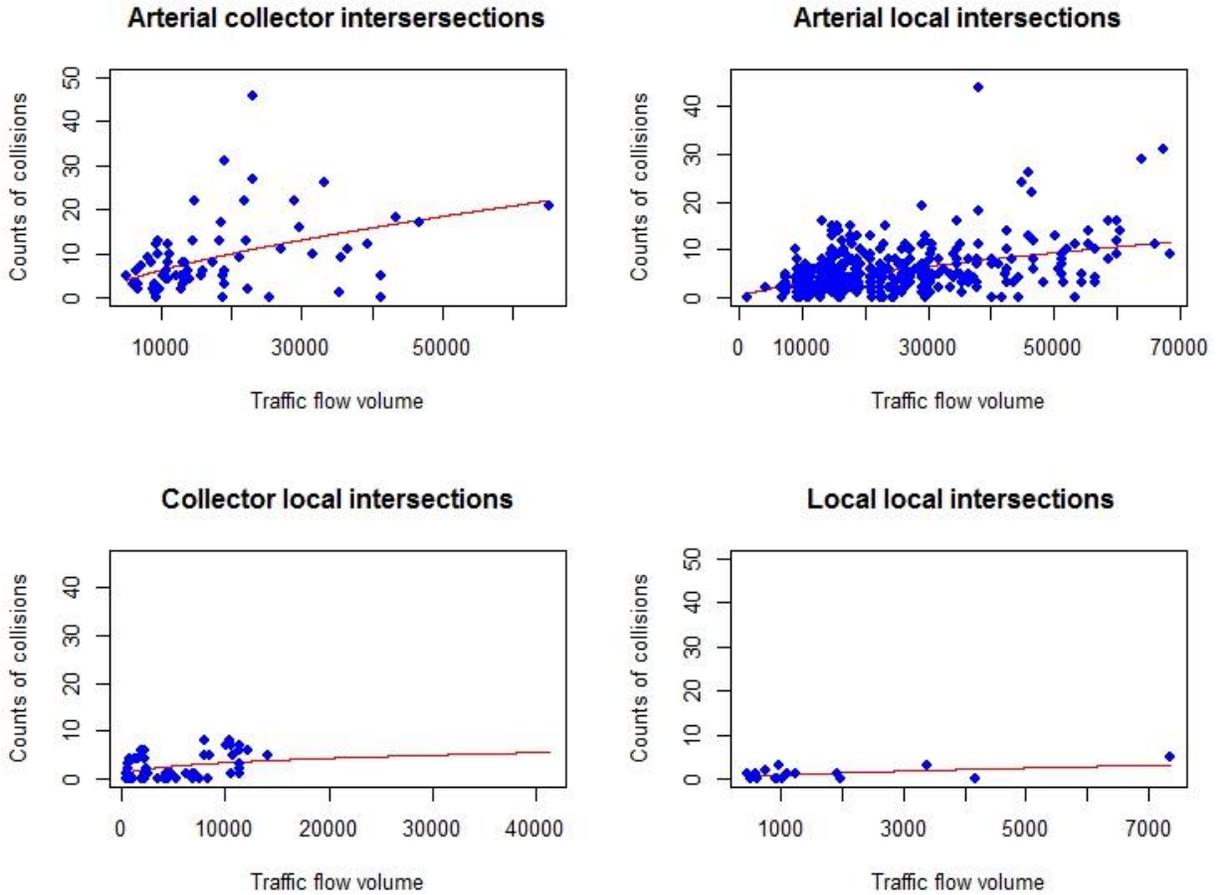
Young, Park and Eng (2012) uses the MPB, MAD and MSPE as criteria for evaluating the fit of an SPF. The baseline for objectively assessing the model performance is by comparing the predicted counts with the corresponding observed counts. Since the errors increase with the mean, relying on just this three criteria and the RMSE for model assessment may obscure our ability to ascertain the true magnitude of the model performance. As an additional criterion, we consider the mean absolute percentage error. Overall, five criteria were used to assess the model performance and determine the impact on predictive ability when the functional class model was introduced.

1. The mean square prediction error, $MSPE = \frac{1}{n} \sum_{i=1}^{n} \left(\widehat{N_i} - N_i \right)^2$

2. The mean absolute deviation, $MAD = \sum_{i=1}^{n} \frac{|\widehat{N_i} - N_i|}{n}$

3. The root mean square error, $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(\widehat{N_i} - N_i \right)^2}$

4. The mean prediction bias , $MPB = \frac{1}{n} \sum_{i=1}^{n} \left(\widehat{N_i} - N_i \right)$

5. The mean absolute percentage error, $MAPE = \sum_{i=1}^{n} \frac{|\widehat{N_i} - N_i|}{N_i}$

5.2. Results *for* Missing Minor Road Volumes and Use of Class as a Proxy Variable

Table 5: Results of SPF fit comparing the aggregate model and functional class model

Statistical test	Basic model approach	Functional class model approach			
	ALL	*AC*	*AL*	*CL*	*LL*
MSPE	60.4681	59.7035	18.5869	5.7072	1.2916
MAD	5.2402	5.2913	2.9358	2.1288	0.9272
RMSE	7.7761	7.7268	4.3112	2.3890	1.1365
MPB	-0.0675	-0.0717	-0.0089	0.0166	0.0090
MAPE	1.0578	0.8850	0.7089	0.5857	0.2162

Notes: MSPE=Mean square prediction error, MAD=Mean absolute deviation, RMSE=Root mean square error, MPB=Mean prediction bias, MAPE= Mean absolute prediction error, CL=Collector-local, AL=Arterial local, CL=Collector-local, LL=Local-local

5.3. Sub-conclusions for Missing Minor Road Volumes and Use of Class as a Proxy Variable

Table 4 summarizes the estimates of MAD, MPB, MSPE, RMSE and MAPE for both the basic model approach and the functional class model approach. Being a measure of model performance, the smaller the value of each of the evaluation criterion the better the model performance. Comparing the MAD, MPB, MSPE, RMSE and MAPE for the basic model and the functional class model, the table shows that for all the evaluation criteria considered, the functional class models demonstrated a superior performance over the basic model approach. This means, relying on an aggregate model which does not account for the class of the intersection to predict intersection collisions results in low precision predictions. When the functional class model is considered, the table shows very low values comparatively for each of the evaluation criterion suggesting the localized models for each class perform better. Overall, using the functional class of intersections as proxy for missing street volumes provides better and superior localized models for each functional class than a single aggregate model which collapses all intersections regardless of their functional class into one population.

6. Discussion and Conclusion

This paper has demonstrated some ways that reliable models for confident decision-making can be built even when the input data is less than ideal. In the Winnipeg case, three data issues were encountered: uncertain traffic volumes, non-uniform collision under-reporting, and missing traffic volumes. Three corresponding techniques were used to address these issues: Monte Carlo simulations to check uncertainty multiples, breaking populations with cumulative residual plots to reduce model bias, and the use of functional class as a proxy variable in lieu of missing minor road volumes to reduce model error. In each case quantitative results demonstrate the improved confidence that can be placed in the decision-support models as a result of the analysis. For the uncertain traffic volumes, the Monte Carlo Analysis showed that model parameter uncertainty attributable to volume measurement error is less than inherent paramater uncertainty for volume measurement erros of up to 30%. For the non-uniform under-reporting, model stratification using the CURE plots to discover population breakpoints reduced global model bias from 22% to 0% and almost eliminated local model bias. For missing minor road volumes, the technique of using functional class as a proxy variable reduced mean absolute deviations by almost 50% from 5.2 with a model omitting the minor road variable to an average of 2.8 with a model incorporating class as proxy.

With these and similar techniques depending on the data issue encountered, we suggest that confident decisions can be made with the best analytical tools despite input data limitations. Practitioners may often be faced with a decision about the maturity level of analytical tools they will employ to allocate road safety budgets. The highest maturity tools such as network screening using SPFs and Empirical Bayes techniques bring the promise of better road safety budget allocations and more lives saved, however, the tools have demanding input data requirements. Before deciding on a lower maturity level analytical tool – and correspondingly lower budget allocation effectiveness – due to input data concerns, we hope that practitioners can consider using these techniques to leverage the input data they have despite its limitations.

References

Burkardt, J. (2014). The Truncated Normal Distribution. *Department of Scientific Computing Website, Florida State University*.

Davis, G. (2000). Estimating traffic accident rates while accounting for traffic-volume estimation error: A Gibbs sampling approach. *Transportation Research Record: Journal of the Transportation Research Board*(1717), 94-101.

Dixon, K., Monsere, C., Avelar, R., Barnett, J., Escobar, P., Kothuri, S., & Wang, Y. (2015). Improved Safety Performance Functions for Signalized Intersections.

El-Basyouny, K., & Sayed, T. (2010). Safety performance functions with measurement errors in traffic volume. *Safety science, 48*(10), 1339-1344.

Hauer, E., & Bamfo, J. (1997). *Two tools for finding what function links the dependent variable to the explanatory variables.* Paper presented at the Proceedings of the ICTCT 1997 Conference, Lund, Sweden.

Hauer, E., Kononov, J., Allery, B., & Griffith, M. (2002). Screening the road network for sites with promise. *Transportation Research Record: Journal of the Transportation Research Board*(1784), 27-32.

Jackson, C., & Jackson, M. C. (2015). Package 'msm'.

Li, W., Carriquiry, A., Pawlovich, M., & Welch, T. (2008). The choice of statistical models in road safety countermeasure effectiveness studies in Iowa. *Accident Analysis & Prevention, 40*(4), 1531-1542.

Lyon, C., Haq, A., Persaud, B., & Kodama, S. (2005). Safety performance functions for signalized intersections in large urban areas: Development and application to evaluation of left-turn priority treatment. *Transportation Research Record: Journal of the Transportation Research Board*(1908), 165-171.

Miranda-Moreno, L. F., Lord, D., & Fu, L. (2008). *Bayesian road safety analysis: incorporation of past experiences and effect of hyper-prior choice.* Paper presented at the Proceedings of TRB 87th Annual Meeting Compendium of Papers DVD, Transportation Research Board, Washington, DC.

Ng, J. C., & Sayed, T. (2004). Effect of geometric design consistency on road safety. *Canadian Journal of Civil Engineering, 31*(2), 218-227.

Persaud, B., & Lyon, C. (2007). Empirical Bayes before–after safety studies: lessons learned from two decades of experience and future directions. *Accident Analysis & Prevention, 39*(3), 546-555.

Torres, M. D. S., Lechón, R. F., & Soto, P. F. (2012). *Road safety strategies: An analysis with system dynamics.*

Young, J., Park, P., & Eng, P. (2012). *Comparing the Highway Safety Manual's Safety Performance Functions with Jurisdiction-Specific Functions for Intersections in Regina.* Paper presented at the Implementation of Safety Prediction Methods in the Processes

and Procedures of a Transportation Agency Session of the 2012 Annual Conference of the Transportation Association of Canada, Fredericton, New Brunswick.