Quantifying safety benefit of winter road maintenance: Accident frequency modeling

Taimur Usman [a,*], Liping Fu [a,1], Luis F. Miranda-Moreno [b,2]

[a] Department of Civil & Environmental Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada
[b] Department of Civil Engineering & Applied Mechanics, McGill University, Montreal, Quebec H3A 2K6, Canada

ARTICLE INFO

ABSTRACT

This research presents a modeling approach to investigate the association of the accident frequency during a snow storm event with road surface conditions, visibility and other influencing factors controlling for traffic exposure. The results have the premise to be applied for evaluating different maintenance strategies using safety as a performance measure. As part of this approach, this research introduces a road surface condition index as a surrogate measure of the commonly used friction measure to capture different road surface conditions. Data from various data sources, such as weather, road condition observations, traffic counts and accidents, are integrated and used to test three event-based models including the Negative Binomial model, the generalized NB model and the zero inflated NB model. These models are compared for their capability to explain differences in accident frequencies between individual snow storms. It was found that the generalized NB model best fits the data, and is most capable of capturing heterogeneity other than excess zeros. Among the main results, it was found that the road surface condition index was statistically significant influencing the accident occurrence. This research is the first showing the empirical relationship between safety and road surface conditions at a disaggregate level (event-based), making it feasible to quantify the safety benefits of alternative maintenance goals and methods.

1. Introduction

In countries with severe winters such like Canada, winter road safety is a source of concern for transportation officials. Driving conditions in winter can deteriorate and vary dramatically due to snowfall and ice formation, causing significant reduction in pavement friction and increasing the risk of accidents (Andrey et al., 2001; Eisenberg and Warner, 2005; Velavan, 2006; Knapp et al., 2002; Nixon and Qiu, 2008; Roskam et al., 2002). The total cost due to weather-related injuries and property damages is estimated to be in the range of $1 billion per year in Canada (Andrey et al., 2001).

Winter road maintenance (WRM) such as plowing and salting plays an indispensable role in maintaining good road surface conditions and keeping roads safe (Fu et al., 2006; Hanbali and Kuemmel, 1992; Norrman et al., 2000). However, winter maintenance operations also incur significant monetary costs and negative environmental effects. For example, the direct cost of winter main-

tenance programs in Ontario is estimated to exceed $100 million annually (Perchanok et al., 1991), representing 50% of its total highway maintenance budget (Buchanan and Gwartz, 2005). Costs for winter road maintenance are estimated at $1 billion in Canada, and over $2 billion in the United States (Transport Association of Canada, 2003). These estimates do not include significant indirect costs such as damage to the environment, road infrastructure, and vehicles due to salt use (Environment Canada, 2002; Perchanok et al., 1991). The environmental impact of salting has also been widely recognized. A recent study by Environment Canada concluded that road salts at high concentrations pose a risk to plants, animals and aquatic system (Transport Canada, 1999).

The substantial direct and indirect costs associated with winter road maintenance have stimulated significant interest in quantifying the safety and mobility benefits of winter road maintenance for systematic cost–benefit assessments. A number of studies have been initiated in the past decade to identify the links between winter road safety and factors related to weather, road, and maintenance operations. However, most of these studies have focused on the effects of adverse weather on road safety (Andreescu and Frost, 1998; Andrey et al., 2001; Handman, 2002; Knapp et al., 2002; Kumar and Wang, 2006). Limited efforts have been devoted to the problem of quantifying the safety benefits of winter road maintenance under specific weather conditions. Furthermore, most

* Corresponding author. Tel.: +1 519 888 4567x33872/635 9188.
E-mail addresses: tusman@engmail.uwaterloo.ca (T. Usman), lfu@uwaterloo.ca (L. Fu), luis.miranda-moreno@mcgill.ca (L.F. Miranda-Moreno).
[1] Tel.: +1 519 888 4567x33984; fax: +1 519 725 5441.
[2] Tel.: +1 514 398 6589; fax: +1 514 398 7361.

existing studies have taken an aggregate analysis approach, considering roads of all classes and locations together and assuming uniform road weather conditions over the entire day. This aggregate approach may average out some important environmental and operating factors that affect road safety at a microscopic level (e.g., a roadway section). Therefore, results may not be applicable for assessing decisions at an operational level with an analysis scope of a maintenance yard. Moreover, past studies usually do not control for the effects of traffic and road surface conditions simultaneously. The joint interactions between road driving conditions, traffic and maintenance and their impact on traffic safety have rarely been studied. In particular, few studies have investigated the link between road safety and road surface conditions resulting from the mixed effects of weather, traffic and road maintenance during snow storms.

This research aims to investigate the effect of road surface conditions on accident occurrence under adverse winter weather conditions. This paper is divided into six sections. The first section provides a brief introduction to the problems at hand and the research question. The second section provides a literature review of winter maintenance operations, weather and road safety. Section three presents a conceptual framework in which our modeling methodology is grounded. Section four presents modeling and exploratory data analysis. Section five presents a set of event-based models fitted to the empirical data. The developed models are then compared and the best model identified. Section six highlights the main conclusions and outlines directions for future research.

2. Safety effect of winter road maintenance — literature review

Significant past efforts have been directed towards road safety problems in general and winter road safety in particular. This section provides a review of studies that focused specifically on the effect of winter road maintenance on road safety. For other general winter road safety issues and research, readers are referred to, e.g., Andrey et al. (2001), Shankar et al. (1995), Hermans et al. (2006), Nixon and Qiu (2008), and Qin et al. (2007).

One of the first documented studies in the literature on the effect of winter road maintenance on road safety is provided by Hanbali and Kuemmel (1992). The authors conducted a simple before–after analysis to identify the effectiveness of salting on 570 miles of randomly selected divided and undivided roads from New York, Minnesota, and Wisconsin. Their analysis concluded that significant reductions in accidents were observed after salting operations. The average reduction in accident rates was 87% and 78% for two-lane undivided highways and freeways, respectively. It should be noted that, while easy to understand, the analysis approach is overly simplistic, accounting only for traffic volumes before and after salt spreading. The traffic volumes were estimated according to the historical temporal variation of traffic as opposed to observed counts during the events. Furthermore, the study did not consider several other most important factors, especially, the weather-related factors such as precipitation, temperature, and visibility.

Norrman et al. (2000) conducted a more elaborated study to quantify the relationship between road safety and road surface conditions. In their study, they classified road surface conditions into ten different types based on slipperiness, and then compared the crash rates associated with the different road surface types. They defined accident risk for a specific road surface condition type, as the ratio of the accident rate ($A_{t,m}/h_{t,m}$, where $A_{t,m}$ is the number of accidents that had occurred under road surface condition t in month m and $h_{t,m}$ the corresponding number of hours whereas) to

the expected number of accidents for each month (A_m/h_m, where A_m represents all accidents during a month and h_m the number of hours in that month). The accident risk computed was then compared to the percentage of maintenance activities performed. They concluded that in general, increased maintenance was associated with decreased number of accidents. However, their approach has several limitations. Firstly, it is an aggregate analysis in nature, considering roads of all classes and locations together. This approach may mask some important factors that affect road safety, such as road class and geometrical features, traffic, and local weather conditions. The resulting models may not be applicable for assessing safety effects of different maintenance policies and decisions at the level of maintenance yards. Secondly, the simple categorical method of determining crash rates may introduce significant biases if confounding factors exist, which is likely to be the case for a system as complex as highway traffic. Furthermore, the procedure cannot be used to compare the effect of different maintenance operations.

Recently, Fu et al. (2006) investigated the relationship between road safety and various weather and maintenance factors, including air temperature, total precipitation, and type and amount of maintenance operations. Two sections of Highway 401 were considered. They used the generalized linear regression model (Poisson distribution) to analyze the effects of different factors on safety. They concluded that anti-icing, pre-wet salting with plowing and sanding have statistically significant effects on reducing the number of accidents. Both temperature and precipitation were found to have a significant effect on the number of crashes. Their study also presents several limitations. First, the data used in their study were aggregated on a daily basis, assuming uniform road weather conditions over entire day for each day (record). Secondly, their study did not account for some important factors due to data problems, such as traffic exposure and road surface conditions. Furthermore, the data available for their analysis covered only nine winter months and thus the power of the resulting model needs to be further validated. One of the implications of these limitations is that their results may not be directly applicable for quantifying the safety benefit of winter road maintenance of other highways or maintenance routes.

Nordic countries have conducted extensive research on issues related to winter road safety and road maintenance. However, most of these studies were published in the form of project reports in the local language and few were published in academic journals. Wallman et al. (1997) provided a comprehensive review on this body of work. In terms of research methodology, most of these studies relied on simple comparative analyses instead of rigorous statistical modeling. Nevertheless, the findings were in general consistent, showing that winter weather increases the risk of accidents by virtue of poor road surface conditions and that maintenance lowers the crash risk by improving road surface conditions.

3. Proposed methodology

As discussed previously, there are a number of factors that could cause road accidents during a snow storm, including weather conditions, traffic characteristics, and maintenance operations, as schematically illustrated in Fig. 1. To quantify the relationship between road safety and these factors, an event-based modeling approach is proposed as an attempt to explain the variation of accident frequencies across individual snow storm events. The proposed methodology includes the following steps:

1. Selection of study sites and data sources (traffic, weather, maintenance and accident data).
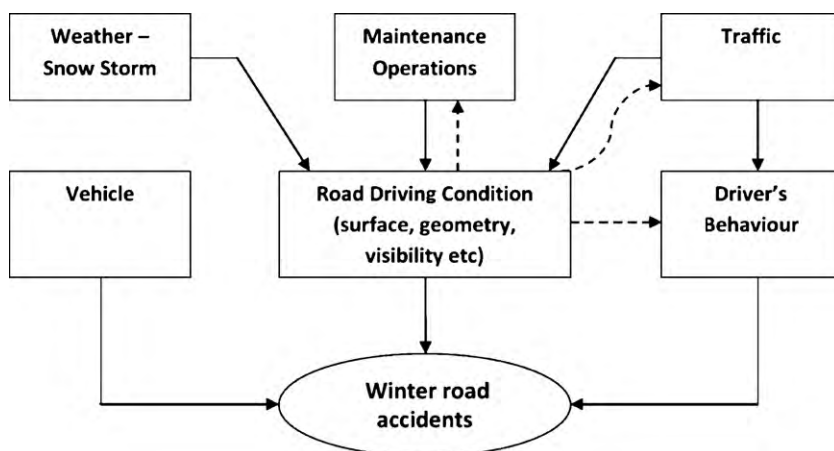2. Data processing (storm-event data).

Fig. 1. Relation between maintenance, weather, traffic and safety.

3. Modeling of road surface conditions.
4. Exploratory data analysis and development of statistical models.

3.1. Study sites

For the proposed analysis, historical information on both road accidents and possible influencing factors must first be obtained. This means that the sites that should be selected for this study must be well instrumented so that detailed data on all major factors of interest are available. In this research, after examining a number of candidate sites, four patrol routes were selected. Two are on Highway 401 and two are on the Queen Elizabeth Way (QEW) in the province of Ontario, Canada, as shown in Fig. 2:

- *401-R1*: Hwy 400 to Morningside Ave (28.0 km).
- *401-R2*: Trafalgar Road to Hwy 400 (31.1 km).
- *QEW-R1*: Burloak Drive to Erin mills parkway (17.4 km).
- *QEW-R2*: Erin mills parkway to Eastmall (13.1 km).

For the proposed investigation, five types of data sources are needed, including weather data, traffic data, accident data, road surface condition data and winter operations data. These data usually originate from different sources and are managed by different organizations, as described in the following section.

3.1.1. Traffic volume data

Traffic volume data was obtained from loop detectors collected by the Ontario Ministry of Transportation (MTO). The data were then processed in Excel and converted into hourly traffic volume data. Note that loop detectors are a commonly used technology for collecting traffic data such as volume, speed and density. This data was screened for any outliers caused by detector malfunction.

3.1.2. Traffic accident data

The Ontario Provincial Police (OPP) maintains a database of all of the collisions that have been reported on Ontario highways. A database including all of the collision records for the study routes was obtained from the MTO. The database includes detailed information on each collision, including accident time, accident location, accident type, impact type, severity level, vehicle information, driver information, etc. Note that the data on the accident occurrence time and location are needed for data aggregation over space (e.g., highway maintenance route) and time (e.g., by hour). The data item related to road surface conditions in the accident data represent the conditions at the time and location associated with the observed collisions only. Therefore, they do not necessar-

ily represent the condition of the whole maintenance route. As a result, we did not use this data field directly and instead used it to fill the missing RSC data from road condition weather information system (RCWIS) and road weather information system (RWIS). For this analysis only number of accidents and the time (hour) they occurred and road surface condition data, when no information was available from RCWIS/RWIS data, are considered from this database.

3.1.3. Road condition weather information system (RCWIS) data

This data contains information about road surface conditions, maintenance, precipitation type, accumulation, visibility and temperature. RCWIS data is collected by MTO maintenance personnel, who patrol the maintenance routes 3–4 times during a storm event on the average. One of the most important pieces of information in this data source is a description of road surface conditions based on the MTO's internal condition classification system. A detailed description of this data field and its processing for the subsequent modeling analysis is given later. This data is also used by MTO in their traveler's road information system; however, it is utilized here for the first time for research purposes.

3.1.4. Road weather information system (RWIS) data

This data source contains information about temperature, precipitation type, visibility, wind speed, road surface conditions, etc., recorded by the RWIS stations near the selected maintenance routes. All data except hourly precipitation were available on hourly basis. Hourly precipitations from RWIS sensors were either not available or not reliable. As a result, we derive this information from the daily precipitation reported by Environment Canada (EC). Temperature and RSC data from RWIS were used to fill in the missing data from RCWIS. For visibility and wind speed RWIS was used as the primary source.

3.1.5. Environment Canada (EC) data

Weather data from Environment Canada includes temperature, precipitation type and intensity, visibility and wind speed. Except precipitation related information, all data from this data source were used as a secondary source for filling in the missing data from RCWIS/RWIS. EC is the only reliable data source for precipitation type and intensity and it was therefore used as the primary source for these variables.

Fig. 2. Site map showing the location of the selected patrol routes.

3.2. Data processing

As described previously, there are five types of data available for each selected study site. Once these data are obtained, they need to be pre-processed for merging and integration.

Accident data are available as event records and therefore need to be aggregated into hourly records by totalling the accidents that have occurred within each hour of the day. Other attributes associated with accidents are averaged over each hour. The output of this process is hourly data records of the total number of accidents and the associated average conditions such as speed, light, road surface conditions, precipitation, road physical characteristics, etc.

The weather data is from three different sources, namely RCWIS, RWIS and EC. RCWIS data, organized by patrol yards, are descriptive (or categorical) and non continuous in nature, similar to event-based records with a time stamp. As a result, two immediate issues need to be addressed. First, the original RCWIS data classify road surface conditions (RSC) into a total of 160 classes, which make it difficult to be directly modeled as a categorical factor in a statistical analysis. The second issue is that a large number of hours do not have RSC observations. As a result, some assumptions have to be introduced to interpolate the RSC between observations, as detailed in the section. RWIS data are recorded every 20 min and can be easily aggregated into hourly records based on their time stamps.

Hourly data from Environment Canada was obtained for type of precipitation, visibility, wind speed, etc. However, precipitation intensity data are available only as a daily total, which is the water equivalent of the total precipitation amount over a day. Based on the data on "precipitation type" the hours with and without precipitation were decided. The total daily precipitation was then uniformly allocated to each hour of the hours with precipitation.

After all weather and road condition related data are converted into the hourly format, they were fused into a single data set on the basis of date and time. When multiple data were available for a given field, priority was given to RCWIS and RWIS data over EC data because these data sources are collected near the study sites and therefore considered to be more representative. Missing RSC data from RCWIS were retrieved from accident data or RWIS data. It was also assumed that the RSC at the hour right after a maintenance treatment was done could be considered as partially snow covered. This data field was then subsequently linearly interpolated for hourly conditions, as discussed in the following section. This gave us values of road surface conditions for all hours over individual storms. In case of any missing data for temperature, precipitation and wind in RCWIS, data from RWIS or EC data were used. This process resulted in a single hourly based weather data set.

Once the three sources of data were finalized, a single data set was formed by combining all the data sets on the basis of date and time.

3.3. Modeling of road surface conditions

As described previously, MTO reports RSC using qualitative descriptions, i.e., a categorical measure (e.g., 7 major categories and 160 subcategories). However, these categories have intrinsic ordering in terms of the severity, which means that a more sensible measure would be an ordinal one. While binary variables could be used to code ordinal data, it would mean loss of information on the ordering. We therefore decided to use an interval variable to map the RSC categories and at the same time make sure that the new variable has physical interpretations. Road surface condition index (RSI), a surrogate measure of the commonly used friction level, was therefore introduced to represent different RSC classes described in
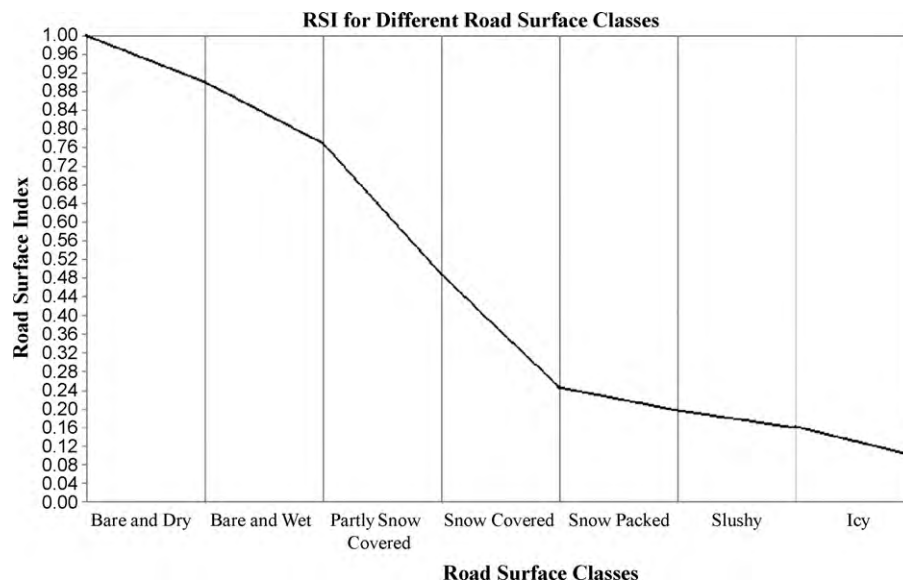
Fig. 3. RSI for different road surface classes.

RCWIS. The reason to use a friction surrogate is that there have been good amounts of field studies available on the relationship between descriptive road surface conditions and friction, which provided the basis for us to determine boundary friction values for each category. To map the categorical RSC into RSI, the following procedure was used:

(1) The major classes of road surface conditions, defined in RCWIS, were first arranged according to their severity in an ascending order as follows: *Bare and Dry < Bare and Wet < Partly Snow Covered < Snow Covered < Snow Packed < Slushy < Icy*.

This order was also followed when sorting individual sub categories in a major class.

(2) Road surface condition index (RSI) was defined for each major class of road surface state defined in the previous step as a range of values based on the literature in road surface condition discrimination using friction measurements (Wallman et al., 1997; Wallman and Astrom, 2001; Transportation Association of Canada, 2008; Feng et al., 2010). For convenience of interpretation, RSI is assumed to be similar to road surface friction values and thus varies from 0.1 (poorest, e.g., ice covered) to 1.0 (best, e.g., bare and dry).

(3) Each category in the major classes is assigned a specific RSI value. For this purpose, sub categories in each major category were sorted as per step 1 above. Linear interpolation was used to assign RSI values to the sub categories.

RSI values for major road surface classes are given below:

Road surface condition major classes	RSI range
Bare and dry	0.9–1.0
Bare and wet	0.8–0.89
Partly snow covered	0.5–0.79
Snow covered	0.25–0.49
Snow packed	0.2–0.24
Slushy	0.16–0.19
Icy	0.1–0.15

Fig. 3 shows variation of RSI for major categories.

3.4. Generation of event-based dataset

After the hourly data set as prepared, an event-based data set was generated by aggregating the hourly data over individual events. An important step in this data aggregation step was to identify the individual events with the available hourly data records. For the purpose of this research, events are defined on the basis of not only weather conditions but also of road surface conditions. This approach makes this research different from other event-based studies where events are mostly defined based on environmental data alone (e.g., Knapp et al., 2002).

Each event is defined with the following constraints:

- An event starts at the time when snow/freezing rain is observed.
- An event ends when snow/freezing rain stops and a certain predefined road surface condition is achieved after that time.
- Precipitation must be greater than zero (0).
- Air temperature must be less than 5 °C.
- Road surface conditions index value must not be equal to bare dry conditions.

This definition of storm events along with road weather, surface conditions and maintenance activities is schematically illustrated in Fig. 4. The numbers of snow storm events and accidents that were identified for individual routes are given in Table 1.

4. Model development

In road safety literature, the most commonly employed approach for modeling accident frequencies is the regression analysis for count data. In particular, the Negative Binomial (NB) model and its extensions have been found to be the most suitable distribution structures for road accident frequency (Shankar et al., 1995; Miaou and Lord, 2003; Miranda-Moreno, 2006; Sayed and El-Basyouny, 2006). Among the NB model extensions, we can mention the generalized NB, latent class NB and zero inflated NB

Table 1
Number of events for each maintenance route.

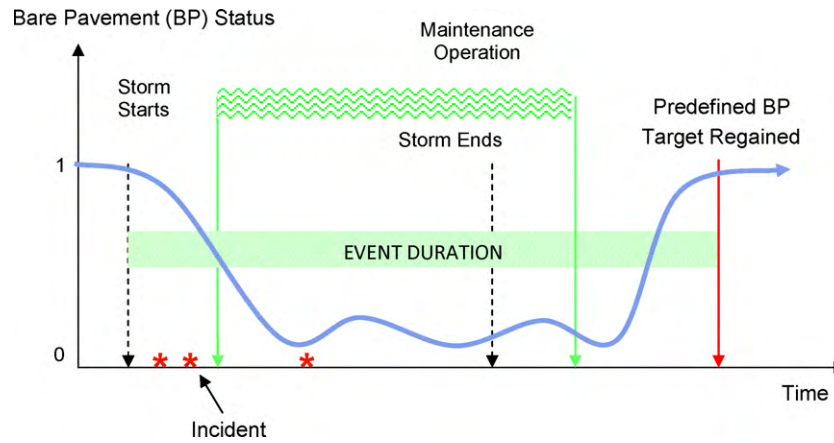| Patrol no.	Route length (km)	No. of events	No. of accidents	Accidents per event
401-R1	28	249	262	1.05
401-R2	31.1	259	158	0.61
QEW-R1	17.4	204	34	0.17
QEW-R2	13.1	200	32	0.16
Total	–	912	486	0.53

Fig. 4. Road surface condition index variation over a snow storm event.

models — (for instance see Miaou, 1994; Shankar et al., 1997; Miranda-Moreno, 2006). In this research, the NB model is therefore first evaluated for its performance in capturing observed and unobserved accident variations among individual snow storms. This model can be written as, $Y_i \sim NB(\mu_i, \alpha)$, where Y_i represents the number of accidents during an event i ($i = 1, \ldots, n$), μ_i stands for the mean accident frequency, and α is the over-dispersion parameter. Furthermore, it is assumed that the mean accident frequency (μ_i) is a function of a set of covariates through the log link function commonly used in the road safety literature, that is:

$$\mu_i = \exp(\beta_0 + \beta_1 \, \text{Ln(Exposure)} + \beta_2 x_{i1} + \beta_3 x_{i2} + \cdots + \beta_k x_{ik}) \quad (1)$$

$$\mu_i = (\text{Exposure})^{\beta_1} \exp(\beta_0 + \beta_2 x_{i1} + \beta_3 x_{i2} + \cdots + \beta_k x_{ik}) \quad (2)$$

where x_{ij} is the jth attribute associated with event i, Exposure is as defined in Section 4.1 and ($\beta_0, \beta_1, \ldots, \beta_k$) is a vector of regression parameters. One of the shortcomings of the NB model is the assumption of a constant over-dispersion parameter (α) for all observations. This assumption can be relaxed by assuming that the dispersion parameter is a function of a set of covariates, using an exponential link function as follows:

$$\alpha_i = \exp(\gamma_0 + \gamma_1 z_{i1} + \gamma_2 z_{i2} + \cdots + \gamma_k z_{im}) \quad (3)$$

where ($z_{i1}, \ldots, z_{im}$) is a vector of event-specific factors that may be different than those explaining μ_i and ($\gamma_0, \gamma_1, \ldots, \gamma_m$) is a vector of parameters. The resulting model is commonly referred to as generalized Negative Binomial (GNB) model (Miaou and Lord, 2003; Miranda-Moreno et al., 2005; Miranda-Moreno, 2006). This model may allow more flexibility than its alternatives to deal with the well known over-dispersion problem and unobserved heterogeneities among events.

The third model considered in this research is called zero-inflated NB or ZINB model, which is also an extension of the standard NB model. This assumes a dual state process in accident occurrence: one generating safe events with zero accidents ($Y_i \sim 0$ with probability p) and the other state following an NB distribution [$Y_i \sim NB(\mu_i, \alpha)$ with probability $1 - p$]. The ZINB model may be more flexible compared to the NB model since it can handle both over-dispersion due to unobserved heterogeneity and excess of zero counts. These three modeling options are all evaluated for their suitability to fit the given data.

4.1. Exploratory data analysis

After the events were extracted from the hourly data records, the following factors/variables were calculated for each storm event:

- Total number of accidents over the event.
- Event duration (h).
- Average visibility (km).
- Average air temperature (°C).
- Average wind speed (km/h).
- Precipitation type (1 = freezing rain, 2 = snow).
- Total precipitation[3] (mm).
- Hourly precipitation (mm).
- Hourly traffic volume (vehicles/h).
- Maintenance (1 = Sanding, 2 = Salting, 3 = Plowing, 4 = Sanding and Salting, 5 = Sanding and Plowing, 6 = Salting and Plowing, 7 = Sanding and Salting and Plowing).
- Exposure (product of total traffic volume[4] during the event and segment length, converted into million vehicle kilometers or MVkm).
- Average road surface conditions index (1 = bare pavement to 0.10 = icy pavement).

Five sets of data were formed: one for each patrol route and one including all patrol routes together. All the data sets were checked for any outliers using box plots of individual data fields in the database. A number of two way interactions were also considered for some of the variables. Note that these interaction terms were identified on the basis of some possible physical interpretation; terms that are more complicated are not recommended to include in regression models due to the added challenge of model interpretation.

A correlation analysis was also carefully conducted for each individual and combined dataset. As suspected, it was found that precipitation type and maintenance operations were consistently correlated with RSC across all datasets (with a correlation coefficient greater than 0.60) and were therefore excluded from further analysis. Descriptive statistics are presented for variables found significant in Table 2.

5. Model calibration

After the exploratory analysis, the next step is to specify the functional form and to fit the model to the data. Because of the relatively small number of events per patrol and the large number of independent variables that need to be evaluated for their effect on accident frequency, the sample size for each individual site may not be large enough for arriving at trustworthy results (Wright,

---

[3] Sum of precipitation for all the hours of an event.
[4] Sum of the hourly traffic volumes of an event.

Table 2
Descriptive statistics.

Variables	N	Minimum	Maximum	Mean	Std. deviation
Descriptive statistics – 401-R1					
Visibility (km)	249	0.8	24.1	13.093	6.689
Road surface condition index	249	0.225	0.986	0.764	0.196
Accidents	249	0	19	1.050	2.123
Total precipitation (mm)	249	0.12	141.53	12.684	21.212
Ln(Exposure)	249	−2.656	3.256	0.303	1.161
Hourly traffic	249	1254	60197	11907	8684
Descriptive statistics – 401-R2					
Visibility (km)	259	1.05	24.1	13.726	6.820
Road surface condition index	259	0.330	0.987	0.716	0.185
Accidents	259	0	8	0.610	1.167
Total precipitation (mm)	259	0.12	141.53	12.417	21.865
Ln(Exposure)	259	−2.812	3.581	0.354	1.292
Hourly traffic	259	966	60197	11101	8641
Descriptive statistics – QEW-R1					
Visibility (km)	204	2.67	24.1	13.116	6.711
Road surface condition index	204	0.13	0.983	0.675	0.238
Accidents	204	0	5	0.167	0.589
Total precipitation (mm)	204	0.2	246	14.403	27.264
Ln(Exposure)	204	−2.938	3.006	0.583	1.053
Hourly traffic	204	1485	77330	25670	14090
Descriptive statistics – QEW-R2					
Visibility (km)	200	1.8	24.1	12.923	6.319
Road surface condition index	200	0.13	0.98	0.660	0.244
Accidents	200	0	5	0.160	0.571
Total precipitation (mm)	200	0.2	246	15.101	27.763
Ln(Exposure)	200	−4.216	2.58	0.337	1.100
Hourly traffic	200	816	77330	25112	13970
Descriptive statistics – Patrol all					
Visibility (km)	912	0.8	24.1	13.241	6.649
Road surface condition index	912	0.13	0.987	0.708	0.217
Accidents	912	0	19	0.533	1.378
Total precipitation (mm)	912	0.12	246	13.523	24.355
Ln(Exposure)	912	−4.22	3.58	0.388	1.167
Hourly traffic	912	816	77330	17652	13281

1995). As a result, in this research we focused only on the combined dataset assuming that the direction and magnitude of the effect of individual factors on accident frequency are consistent across the four maintenance routes. A dummy variable characterizing the patrol routes is included to capture the remaining effect caused by other route specific factors such as location, driver population, and road geometry. A stepwise elimination process was followed to identify the significant factors. STATA[5] (Version 9) was used for this analysis.

Following the sequential regression process, we found that 4 of 11 factors were statistically significant at $p < 0.05$ for the three different model assumptions, including RSI, visibility, exposure and site effect. As shown in Table 3, the calibration results are quite consistent in terms of significant factors and coefficients across the models. After the three types of models were calibrated, the best fit model was identified using log likelihood value and Akaike Information Criterion (AIC) (Akaike, 1974). These test statistics are widely used to identify the best fit model from a set of models. The AIC statistic is defined as $-2LL + 2p$, where LL is the log likelihood of a fitted model and p is the number of parameters, which is included to penalize models with higher number of parameters: a model with smaller AIC value represents a better overall fit. Based on both the log likelihood value and AIC criterion, GNB was found to be the best fitted model for the data (see Table 3).

In addition to using log likelihood values and AIC for best fit model, we also compared the observed vs. the estimated relative frequencies of the number of accidents (Maher and Summersgill, 1996; Miranda-Moreno, 2006). Fig. 5 shows this plot and from this we can observe that all models fit well to the data and NB is com-

parable to GNB. Considering the small difference in LL and AIC between NB and GNB, NB might also be a model of choice based on the principle of parsimony.

5.1. Model interpretation and application

In general, results from the three models are consistent, as shown in Table 3. Most results obtained in our research are consistent with those reported in the literature, with some exceptions. The following specific observations could be made from the modeling outcome:

- The most interesting result is perhaps that the RSI was found to be a statistically significant factor influencing road safety across all sites. The negative sign associated to the factor sug-
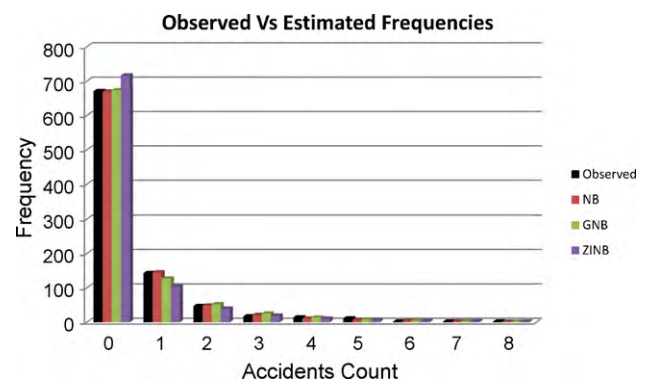

Fig. 5. Observed vs. estimated accident frequencies.

Table 3
Summary results of model calibration.

Variable	NB		GNB		ZINB	
	Coeff	Sig	Coeff	Sig	Coeff	Sig
Intercept	−0.400	0.174	−0.164	0.628	0.183	0.603
Visibility (km)	−0.041	0.000	−0.041	0.000	−0.026	0.017
Road surface condition index	−2.012	0.000	−2.276	0.000	−2.564	0.000
Ln(Exposure)	0.402	0.000	0.377	0.000	0.320	0.000
401-R1	1.986	0.000	2.034	0.000	1.879	0.000
401-R2	1.452	0.000	1.400	0.000	1.349	0.000
QEW-R1	−0.038	0.889	0.032	0.926	−0.031	0.910
QEW-R2	0.000		0.000		0.000	
Number of observations	912		912		912	
Log likelihood (constant only)	−784.57		−903.07		−911.62	
Log likelihood value at convergence	−728.56		−718.06		−719.09	
AIC	1473.12		1462.12		1462.18	
Ln(Alpha)	−0.014				−0.604	
Intercept			2.134	0.001	2.736	0.002
Ln(Exposure)			−0.329	0.017	−0.433	0.026
Hourly precipitation			−0.357	0.025	−1.014	0.005
Hourly traffic						
Road surface condition index					−3.966	0.002
401-R1			−1.545	0.008		
401-R2			−1.168	0.058		
QEW-R1			0.634	0.351		
QEW-R2			0.000			

gests that higher accident frequencies associated with poor road surface conditions. This result makes intuitive sense and has confirmed the findings of many past studies (Norrman et al., 2000; Wallman et al., 1997), mostly from Nordic countries. However, this research is the first showing the empirical relationship between safety and road surface conditions at a disaggregate level, making it feasible to quantify the safety benefit of alternative maintenance goals and methods.

- Visibility is also found to have a statistically significant effect on accident frequency during a snow storm. The negative model coefficient also makes intuitive sense, as it suggests that reduced visibility was associated with increased number of accidents. Note that this result is different from those from a past study by Hermans et al. (2006), which conducted a statistical study using data from 37 sites and found that visibility was significant only at two sites. Their study however considered collisions occurring at different roadways related to a single weather station. This approach may have masked the effect of visibility due to confounding of missing factors and large aggregation levels in both space (coastal areas vs. inter cities) and time (seasonal variation).

- As expected, exposure, defined as total vehicle kilometers traveled (product of the total traffic volume over a storm event and route length), was found to be significant, suggesting that an increase in traffic volume, storm duration, or route length would lead to increase in number of accidents. Inclusion of this term ensures that traffic exposure is accounted for when estimating the safety benefits of some specific policy alternatives. The coefficient associated with the exposure term has a value less than one (0.377), suggesting that the moderating effect of exposure is non-linear with a decreasing rate. This result is consistent with those from road safety literature (e.g., Andrew and Bared, 1998; Lord and Persaud, 2000; NCHRP Synthesis report No. 295, 2001; Roozenburg and Turner, 2005; Mustakim et al., 2006; Sayed and El-Basyouny, 2006; Sayed and Lovegrove, 2007; Jonsson et al., 2007; Lord et al., 2008; etc.).

- The model also suggests that, controlling for other factors (RSC, Visibility and Exposure), Hwy 401 (401-R1 and 401-R2) is more susceptible to crashes than the QEW (QEW-R1 and QEW-R2) while the difference in risk between the two maintenance routes on the same Highway is quite small. The exact reason for this discrepancy between 401 and QEW needs to be further investigated; however, one possible reason for this may be that Hwy 401 has more interchanges per kilometer than QEW, which is known to be an important factor influencing freeway safety in general.

- Both air temperature and precipitation had a statistically insignificant effect on accident frequency. It was found that they are both partially correlated with RSC (0.26 for air temperature and −0.46 for precipitation). These results confirm some of the earlier findings (Hermans et al., 2006) but are in contradiction to others (Andrey et al., 2001; Fu et al., 2006). Precipitation was also partially correlated with visibility (−0.38) (Andrey et al., 2001). Our result therefore makes intuitive sense as the effects of these two factors have already been accounted for through the RSC and visibility factors.

- In our analysis, we did not consider maintenance operations directly, which were assumed to be reflected in their end results, that is, road surface conditions. Our exploratory analysis also found that maintenance was correlated with road surface conditions and were not statistically significant after road surface conditions were accounted for.

Because of the exponential functional form, the exponent in the model is a measure of sensitivity of crash frequency to the corresponding variable. For example, the coefficient associated with RSI in the GNB is −2.276, which suggests that a 1% improvement in RSI would lead to approximately a 2.28% reduction in the expected number of accidents.

The calibrated model could be applied for assessing the safety benefit of alternative winter road maintenance Level of Service (LOS) goals for a specific maintenance route under a specific snow storm event. Fig. 6 shows the relationship between the number of accidents that are expected to occur on the 28 km stretch of Highway 401 (401-R1 in Fig. 2) and the average road surface condition over an assumed snow storm. Again, the GNB is chosen for this analysis. The assumed snow storm lasts for a total of 8 h with average visibility of 13 km. Three levels of traffic are considered, including 50,000, 100,000 and 150,000. The safety effect of maintaining the road section at a given level of road surface condition over the event, as represented by RSI, can be estimated as the mean number of accidents corresponding to the RSI. For example, an average reduction of about three accidents could be achieved by maintaining the route at an RSI of 0.8 as compared to the snow packed icy
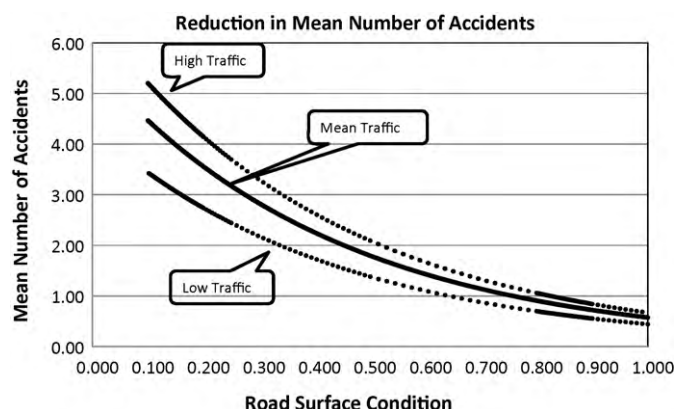
Fig. 6. Accident frequency as a function of RSI.

scenario (RSI = 0.2) for traffic level of 100,000 vehicles. It can be seen from Fig. 6 that effects of traffic volume are more pronounced under poor road surface conditions than good road surface conditions. For instance, the expected number of accidents for an RSI value of 0.1 and high traffic is 5.2 accidents, while this value is 3.44 for a low traffic scenario. This means that the difference between low and high traffic for RSI = 0.1 is 1.77 accidents. At the other extreme, for an RSI value of 1.0, the difference between the low and high traffic condition is 0.23 accidents.

To show further applications of the developed model, we apply the model for two hypothetical case studies. The first case study is to assess the safety implication of adopting different bare pavement (BP) recovery time, one of the critical policy variables in winter road maintenance operations. The same example described previously is used with a traffic flow rate of 12,500 vehicles per hour. The snow event is assumed to last 8 h: 4 h of precipitation and 4 h of BP recovery time.

In the first case, it is assumed that little maintenance work such as salting was done and the road surface was deteriorated from bare dry to bare wet passing through snow covered conditions. Two road surface conditions are considered, namely, snow covered and bare pavement. For snow covered conditions, the corresponding RSI is assumed to be 0.2 (average condition within the snow storm) while the bare pavement surface is assumed to have a RSI of 0.8. It is also assumed that before the start of the storm road surface is dry with a RSI of 1.0. For this scenario, the average RSI = $(1.0 + 7.0 \times 0.2 + 0.8)/9 = 0.356$. Mean number of accidents in this case are 2.50.

Now, we consider the alternative scenario of reducing the BP recovery time to 3 h. This means reducing the storm duration to 7 h (4 h of precipitation and 3 h of BP recovery time). Under the same conditions, the new average RSI = $(1.0 + 6 \times 0.2 + 0.8)/8 = 0.375$. Mean number of accidents in this case are 2.27, that is, a 9% reduction.

In second case it is assumed that some other maintenance work such as plowing has been done in the second hour raising RSI to 0.8 then dropping in a linear way to 0.4 at the end of fourth hour due to precipitation and remain so till the seventh hour after which it raises back to RSI = 0.8. For this case the new average RSI = $(1.0 + 2 \times 0.2 + 2 \times 0.8 + 0.6 + 4 \times 0.4)/10 = 0.520$. Mean number of accidents in this case are 1.72. The relative reduction in the mean number of accidents is therefore 31.2%.

6. Conclusions and future work

This paper has presented the results of a modeling approach aimed at explaining the variation of road accidents over different snow storm events. This approach allows for the relation of winter road safety to some direct road surface condition measures. As part of this work, three different regression models were developed and evaluated using data from four instrumented freeway sections in Ontario, Canada. From the calibrated models, interesting results were obtained. For instance, the study has shown the statistically significant link between road surface conditions represented by RSI and road safety. As illustrated in the paper, this result gives the opportunity to assess the safety benefit of alternative winter road maintenance goals under different weather and traffic conditions. This methodology can be applied to any other roadway sections with detailed high-quality data on road weather conditions, traffic, maintenance and accidents.

Our future efforts will concentrate on examining the validity of these findings across a wider spectrum of weather, highway, traffic and maintenance conditions, and exploring new model structures such as simultaneous equation models for addressing potential endogeneity problems between traffic and accidents. This can be caused by unobserved factors such as speed variations associated with both traffic volumes and accidents. A modeling approach accounting for potential correlation between events will be also attempted. These models can help to improve the explanatory power of the models and the accuracy of the model parameters.

References

Akaike, H., 1974. A new look at the statistical model of identification. IEEE Transaction on Automatic Control 19, 716–723.

Andreescu, M.P., Frost, D.B., 1998. Weather and traffic accidents in Montreal, Canada. Climate Research 9, 225–230.

Vogt, A., Bared, J., 1998. Accident models for two-lane rural segments and intersections. Transportation Research Record 1635. Paper No. 98-0294.

Andrey, J., Mills, B., Vandermolen, J., 2001. Weather Information and Road Safety. Institute for Catastrophic Loss Reduction, Toronto, Ontario, Canada.

Buchanan, F., Gwartz, S.E., 2005. Road weather information systems at the Ministry of Transportation, Ontario. In: Presented at Annual Conference of the Transportation Association of Canada Calgary, Alberta.

Eisenberg, D., Warner, K.E., 2005. Effects of snowfalls on motor vehicle collisions, injuries, and fatalities. American Journal of Public Health 95 (1), 120 (ABI/INFORM Global).

Environment Canada, 2002. Winter road maintenance activities and the use of road salts in Canada: a compendium of costs and benefits indicators.

Feng, F., Fu, L., Perchanok, M.S., 2010. Comparison of alternative models for road surface condition classification. TRB Annual Meeting 2010. Paper #10-2789.

Fu, L., Perchanok, M.S., Moreno, L.F.M., Shah, Q.A., 2006. Effects of winter weather and maintenance treatments on highway safety. TRB 2006 Annual Meeting. Paper No. 06 – 0728.

Hanbali, R.M., Kuemmel, D.A., 1992. Traffic accident analysis of ice control operation. http://www.trc.marquette.edu/publications/IceControl/ice-control-1992.pdf (accessed March 03, 2009).

Handman, A.L., 2002. Weather implications for urban and rural public transit. http://ams.confex.com/ams/pdfpapers/74636.pdf (accessed June 26, 2007).

Hermans, E., Brijis, T., Stiers, T., Offermans, C., 2006. The impact of weather conditions on road safety investigated on an hourly basis. TRB Annual Meeting 2006. Paper No. 06-1120.

Jonsson, T., Ivan, J.N., Zhang, C., 2007. Crash prediction models for intersections on rural multilane highways – differences by collision type. TRR Journal (2019), 91–98, doi:10.3141/2019-12.

Knapp, K.K., Smithson, D.L., Khattak, A.J., 2002. The mobility and safety impacts of winter storm events in a freeway environment. http://www.ctre.iastate.edu/pubs/midcon/knapp1.pdf (accessed December 01, 2007).

Kumar, M., Wang, S., 2006. Impacts of weather on rural highway Operations Showcase Evaluation # 2. http://www.coe.montana.edu/wti/wti/pdf/426243_Final_Report.pdf (accessed June 26, 2007).

Lord, D., Persaud, B.N., 2000. Accident prediction models with and without trend: application of the generalized estimating equations (GEE) procedure. TRB 79th Annual Meeting. Paper No. 00-0496.

Lord, D., Guikema, S.D., Geedipally, S., 2008. Application of the Conway–Maxwell–Poisson generalized linear model for analyzing motor vehicle crashes. Accident Analysis and Prevention 40 (3), 1123–1134.

Maher, M.J., Summersgill, I., 1996. A comprehensive methodology for the fitting predictive accident models. Accident Analysis and Prevention 28 (3), 281–296.

Miaou, S.P., 1994. The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. Accident Analysis and Prevention 26, 471–482.

Miaou, S.P., Lord, D., 2003. Modeling traffic crash–flow relationships for intersections: dispersion parameter, functional form, and Bayes versus empirical Bayes. TRR Journal 1840, 31–40.

Miranda-Moreno, L.F., Fu, L., Saccomano, F., Labbe, A., 2005. Alternative risk models for ranking locations for safety improvement. TRR Journal 1908, 1–8.

Miranda-Moreno, L.F., 2006. Statistical models and methods for identifying hazardous locations for safety improvements. PhD thesis report, University of Waterloo.

Mustakim, F.B., Daniel, B.D., Bin, K., 2006. Accident investigation, blackspot treatment and accident prediction model at federal route FT50 Batu Pahat-Ayer Hitam. Journal of Engineering e-Transaction vol. 1, no. 2, pp. 19–32, University of Malaya.

National Cooperative Highway Research Program, Synthesis 295, 2001. Statistical Methods in Highway Safety Analysis, A Synthesis of Highway Practice. TRB Executive Committee.

Nixon, W.A., Qiu, L., 2008. Effects of adverse weather on traffic crashes: systematic review and meta-analysis. TRB 87th Annual Meeting 2008. Paper No. 08 – 2320.

Norrman, J., Eriksson, M., Lindqvist, S., 2000. Relationships between road slipperiness, traffic accident risk and winter road maintenance activity. Climate Research 15, 185–193 (published September 05).

Perchanok, M.S., Manning, D.G., Armstrong, J.J., 1991. Highway Deicers: Standards, Practice, and Research in the Province of Ontario. Ministry of Transportation Ontario.

Qin, X., Noyce, D.A., Martin, Z., Khan, G., 2007. Road weather safety audit program development and initial implementation. TRB 2007 Annual Meeting. Paper No. 07 – 2684.

Roozenburg, A., Turner, S., 2005. Accident prediction models for signalised intersections. www.ipenz.org.nz/ipenztg/ipenztg_cd/cd/2005_pdf/03_Roozenburg.pdf (accessed June 2008).

Roskam, A.J., Brookhuis, K.A., Waard, D.d., Carsten, O.M.J., Read, L., Jamson, S., Ostlund, J., Bolling, A., Nilsson, L., Antilla, V., Hoedemaeker, M., Janssen, W.H., Harbluk, J., Johansson, E., Tevell, M., Santos, J., Fowkes, M., Engstrom, J., Victor, T., 2002. Human machine interface and safety of traffic in Europe. Project GRD1/2000/25361 S12.319626.

Sayed, T., El-Basyouny, K., 2006. Comparison of two negative binomial regression techniques in developing accident prediction models. TRR Journal (1950), 9–16.

Sayed, T., Lovegrove, G.R., 2007. Macrolevel collision prediction models to enhance traditional reactive road safety improvement programs. TRR Journal (2019), 65–73, doi:10.3141/2019-09.

Shankar, V., Mannering, F., Barfield, W., 1995. Effect of roadway geometrics and environmental factors on rural freeway accident frequencies. Accident Analysis and Prevention 27 (3), 371–389.

Shankar, V., Milton, J., Mannering, F., 1997. Modeling accident frequencies as zero-altered probability processes: an empirical inquiry. Accident Analysis and Prevention 29 (6), 829–837.

Transport Canada. Canadian Environmental Protection Act, 1999. Priority substances list assessment report, road salts.

Transport Association of Canada, 2003. Salt smart train, the trainer program. Salt smart learning guide.

Transportation Association of Canada, 2008. Winter maintenance performance measurement using friction testing. Final draft report, September 2008.

Velavan, K., 2006. Developing tools and data model for managing and analyzing traffic accident. MSc thesis report. University of Texas at Dallas, School of Economic, Political and Policy Sciences for Master of Science in Geographic Information Sciences.

Wallman, C.G., Wretling, P., Oberg, G., 1997. Effects of winter road maintenance. VTI rapport 423A.

Wallman, C.G., Astrom, H., 2001. Friction measurement methods and the correlation between road friction and traffic safety – a literature review. VTI report M911A.

Wright, R.E., 1995. Logistic regression. In: Grimm, L.G., Yarnold, P.R. (Eds.), Reading and Understanding Multivariate Statistics. American Psychological Association, Washington, DC, pp. 217–244.