Road crashes and field rainfall data: mathematical modeling for the Brazilian mountainous highway BR-376/PR

Acidentes rodoviários e dados pluviométricos de campo: modelagem matemática para a rodovia BR-376/PR em trecho montanhoso

Leandro Canezin Guideli[1], André Lucas dos Reis Cuenca[2], Milena Arruda Silva[3], Larissa de Brum Passini[4]

[1]Federal University of Paraná, Paraná – Brazil, leandro.guideli@gmail.com
[2]Federal University of Paraná, Paraná – Brazil, andrelrcuenca@gmail.com
[3]Federal University of Paraná, Paraná – Brazil, milena.arrudas@gmail.com
[3]Federal University of Paraná, Paraná – Brazil, larissapassini@hotmail.com

ABSTRACT
Recent studies analyze the influence of rainfall on traffic crashes, indicating that precipitation intensity is an important factor, for modeling crashes occurrence. This research presents a relationship between daily-basis traffic crashes and precipitation, from 2014 to 2018, in a rural mountainous Brazilian Highway (BR-376/PR), where field rain gauges were used to obtain precipitation data. Data modeling considered a Negative Binomial regression for precipitation influence in crash frequency. Separate regression models were estimated to account for the rainfall effect in different seasons, and for different vehicle types. All models analyzed presented a positive relationship between daily rainfall intensity and daily crashes number. This can indicate that generally rainfall presence is a hazardous factor. Different critical seasons for rainfall influence were also highlighted, alerting for the possible necessity of distinct road safety policies concerning seasonality. Finally, for the vehicle type analysis, typically, rainfall seemed to have a greater effect in lighter vehicles. Moreover, results are useful for traffic control, in order to increase safety conditions.

RESUMO
Pesquisas recentes estudaram a influência de precipitação em acidentes de trânsito, indicando que a intensidade da precipitação é um fator importante na ocorrência de acidentes. Esta pesquisa apresenta a relação entre números de acidentes diários e precipitação, de 2014 a 2018, numa Rodovia brasileira montanhosa e rural (BR-376/PR), onde pluviógrafos locais forneceram dados de precipitação. Uma Regressão Binomial Negativa possibilitou obter a influência da precipitação na frequência de acidentes. Modelos distintos consideraram o efeito da chuva em diferentes estações do ano, e para diferentes tipos de veículos. Houve uma relação positiva consistente entre a intensidade de chuva e o número de acidentes diários, indicando a presença de chuva como fator de risco. Efeitos de sazonalidade apontaram diferentes estações críticas para a influência das chuvas, alertando para a possível necessidade de políticas de segurança viária distintas. Finalmente, para o tipo de veículo, normalmente, a chuva pareceu ter um efeito maior em veículos mais leves. Os resultados são úteis para o controle de tráfego, a fim de aumentar as condições de segurança.

1. INTRODUCTION

In road safety study, there is a common perception that adverse weather conditions, such as

rain and snowfall, favors the occurrence of traffic accidents. In fact, several studies performed in expressways and/or urban freeways have found that these conditions play an important role in higher crash counts (Levine *et al.*, 1995; Eisenberg, 2004; Keay and Simmonds, 2006; Bergel-Hayt *et al.*, 2013). Precipitation as a variable, included in mathematical models, works as a proxy of other aspects that might be hard to quantify, identify or collect (Winston *et al.*, 2006), such as visibility and road slipperiness. Hence, it is important that precipitation data is correlated spatially to accident information (Mannering and Bhat, 2014).

The use of local precipitation data (instead of data from pluviometric stations) is, therefore, more likely to provide a better correspondence of the effect that is actually verified in a highway. In fact, in association with traffic detection systems, infield rain gauges can be used to provide adequate estimates of a crash risk in real time. Other than that, adverse weather conditions can also create a situation in which there is an increased effect in traffic disturbance, which is extremely important for traffic monitoring and predicting crash behavior (Ahmed and Abdel-Aty, 2012).

The present work sought to develop a statistical model using *in situ* rainfall data (from rain gauges) and traffic/crash frequency for a segment of the BR-376/PR Highway, in order to correlate accident frequency and weather conditions in a key Brazilian highway segment. In addition, two other models were developed to address the seasonality effect of rainfall and vehicle type response in crash occurrence.

2. BACKGROUND

Adverse weather conditions have been suggested in literature to prompt traffic crashes. This correlation has been analyzed considering various aspects, on different countries.

For the primary Dutch road system, a mathematical model reported that 5% of the variation in crashes resulting in injuries could be explained by weather conditions (Hermans *et al.*, 2006). Similarly, a research on crash occurrence in Connecticut (USA), found that the relative risk of an accident was roughly 40% lower in dry conditions, when compared to wet situations (Zhang *et al.*, 2005). Rainfall is consistently associated with crashes by an increase in the number of accident records (Theofilatos and Yannis, 2014). The volume of vulnerable drivers in the road is also affected by the weather effect, due to the change in number of trips, as well as in means of transportation choices (Martensen *et al.*, 2016). In this context, a model developed for the contiguous USA, based on monthly and daily analyses (Eisenberg, 2004), found a pronounced positive rainfall relationship in the proposed daily models, while also identifying a negative effect in the monthly analysis, which, according to the author, may have occurred due to a lagged rainfall effect in crash numbers. Also, a study on accident frequency on a major transportation Highway in Taiwan (Chang and Cheng, 2005), stressed that traffic volume and rainfall were indispensable factors determining accident frequency.

The seasonality effect has also been examined in studies evaluating precipitation and risk of crashes. Crash behavior in a mountainous Freeway (Zhang *et al.*, 2005) had distinct effects for the combination of weather and road geometrics on dry and wet seasons, and the risk of accident in the wet period was 82% higher when compared to the dry season. Moreover, different effects may be observed when modeling rainfall and traffic crashes: the rain effect may be less pronounced in the wettest season of the year, as reported for an urban Freeway (Keay and Simmonds, 2006), and a different effect of rainfall intervals could also be found in the four seasons. In contrast, a study on single vehicle crashes in an US Highway (Dabbour and Haider,

2019) indicated that in the wettest season crash risk was at its highest, and weather effect may not be a significant factor depending on the year of analysis.

Another factor considered by traffic research is the vehicle type. According to police reported crashes, incidents with light trucks were found to be more severe, in comparison to passenger cars (Grabler and Hollowell, 1998). However, few researches, if any, have examined rainfall effect in crash frequency with respect to vehicle type.

Furthermore, it is noticeable that traffic crash-precipitation studies usually rely on crash data for large road segments, sometimes even an entire State area (see Eisenberg, 2004), and combine this to regional weather stations data. Even though the contributions from research using this methodology are important, they could possibly lack spatial correlation, since regional climate effects may be present and significant variation in some areas can be observed. In addition, the majority of studies that correlate crash frequency with rainfall target urban expressways or freeways, whereas rural areas have been sparsely examined (see Theofilatos and Yannis, 2014 for a thorough review on the subject).

3. MATERIAL AND METHODS
3.1. Study Area

The study area (Figure 1a) is located in the BR-376 Highway, Paraná State (PR), Brazil. It covers km 644+000 m to 680+000 m. This specific segment is at the "Serra do Mar", which is a mountain range with strongly undulated relief. The "Serra do Mar" is exposed to high annual average rainfall with prolonged periods of rain, with one of the largest volumes of rainfall in the country (Hasui *et al.*, 2012; Vieira and Gramani, 2015; de Almeida and Carneiro, 2017).
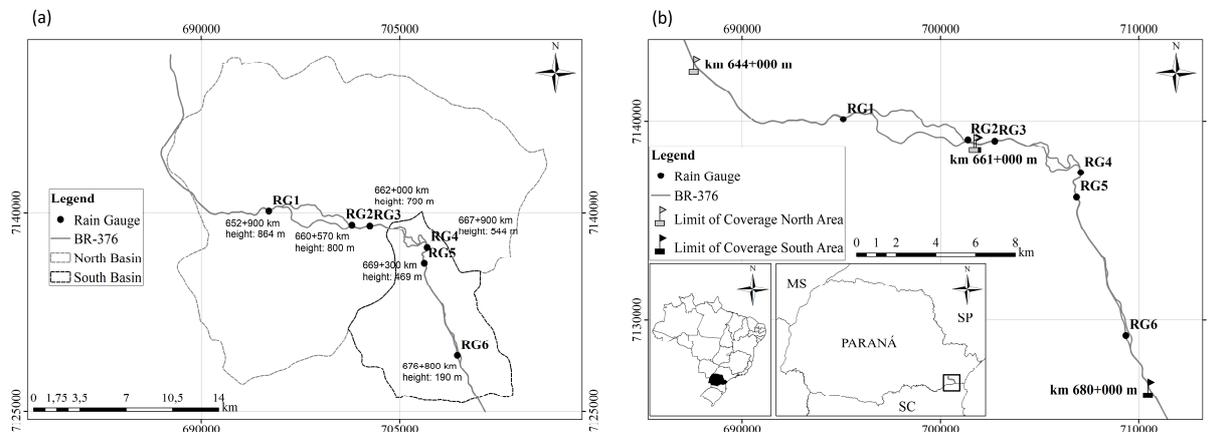

Figure 1. (a) Study area and (b) distribution of rain gauges along BR-376 highway

BR-376 Highway is one of the most important traffic routes for Brazilian products and passenger transports. The area has strategic economic importance, since it connects the exporting and importing port area and the most industrialized and populated regions of the country (Hasui and Carneiro, 2012). Between 2009 and 2012 the BR-376 presented the highest numbers of accidents from all the Paraná highways (Sehaber, 2013). In its total extension, 17,429 accidents were registered involving 31,677 vehicles. In spite of this scenario, only one study (Sehaber, 2013) has described and analyzed accident behavior along the BR-376.

During the years of analysis used in the present study (i.e., 2014 to 2018), in the entire extension of the BR-376, 15,314 total crashes have been registered (PRF, 2018).

3.2. Precipitation Input Data

This study considered rain data from six rain gauges (RG1 to RG6) distributed along the roadway in two hydrographic basins, North and South (Figure 1b), from March, 2014 to September, 2018. These two areas are inserted in the "São João" River Basin. The precipitation data used herein is part of a research project related to monitoring of landslide risks in the area (Victorino, 2015; González Acevedo, 2017; Ferreira *et al.*, 2018; Trevizzoli *et al.*, 2018, Silva *et al.*, 2019).

It would be possible to employ the rain data for each of the rain gauges influence areas. However, this would result in too many regression models, thus making it very difficult to analyze the outcomes. This issue was overcome by grouping areas of influence that correspond to rain gauges that present similar recordings. An Analysis of Variance was performed comparing daily precipitation from each of the six rain gauges. The analysis indicated that data from RG1 and RG2 were similar in each other, while the same happened for RG3, RG4, RG5 and RG6 data. The formers' recordings also presented significant difference from the latter's; therefore, the two groups were determined: RG1 + RG2, (henceforth North Area); and RG3 + RG4 + RG5 + RG6 (South Area). Despite RG3 being located on the North hydrographic basin, the analysis indicated that data from this rain gauge is statistically more similar to the data from the South basin rain gauges (RG4, RG5 and RG6), thus the combination groups were not based on the hydrographic basins.

Daily rainfall data was categorized into six classes, according to the rain gauge register limit value and intervals based on quantiles (Xavier and Xavier, 1987; de Lima Leite *et al.*, 2011; Souza *et al.*, 2012; González Acevedo *et al.*, 2017, Silva *et al.*, 2020). In general, there were lesser days of moderate to extreme precipitation. The South Area registered more days, 98 and 23 days respectively, of these intense events. In the North Area, the days without registered rainfall (480 days) is also evidently greater than that registered in the South (344 days), while in the latter, in most of days (518 days) there was occurrence of drizzle (0.1 to 2.5 mm).

3.3. Crash Input Data

The crash data applied in this study was obtained in the public database from the Federal Highway Police (PRF, 2018), which provided information related to a total of 2,314 crash occurrences in a period of 1,490 days from March, 2014 to September, 2018. The number of crashes analyzed were registered in the segment of interest, which covers km 644+000 m to 680+000 m. The database has a complete description of each accident including location, date and time, road direction, road and weather conditions, day of the week, and accident severity. The main input from this database was the daily frequency of crashes.

The location was used to correspond the accidents count to either North (33.72 km extension) or South Area (38.96 km). In addition, vehicle type and date information were accounted for the type and seasons modeling. Other qualitative variables were not included in the analysis as they did not necessarily relate to the research's purpose. However, some of the unused qualitative factors were analyzed in previous work (Silva *et al.*, 2020). It was demonstrated that in this area, adverse precipitation tends not to shape crashes characteristics,

such as location and severity. The absence/presence of rain, though, was found to be defining to the crashes peculiarities.

A summary of the data used in the analysis is presented in Table 1, separated in (i) areas (North or South), (ii) seasons (December, January and February – DJF; March, April and May – MAM; June, July and August – JJA; September, October and November – SON) and (iii) vehicle type (Light or Heavy).

Table 1 – Summary of the data applied

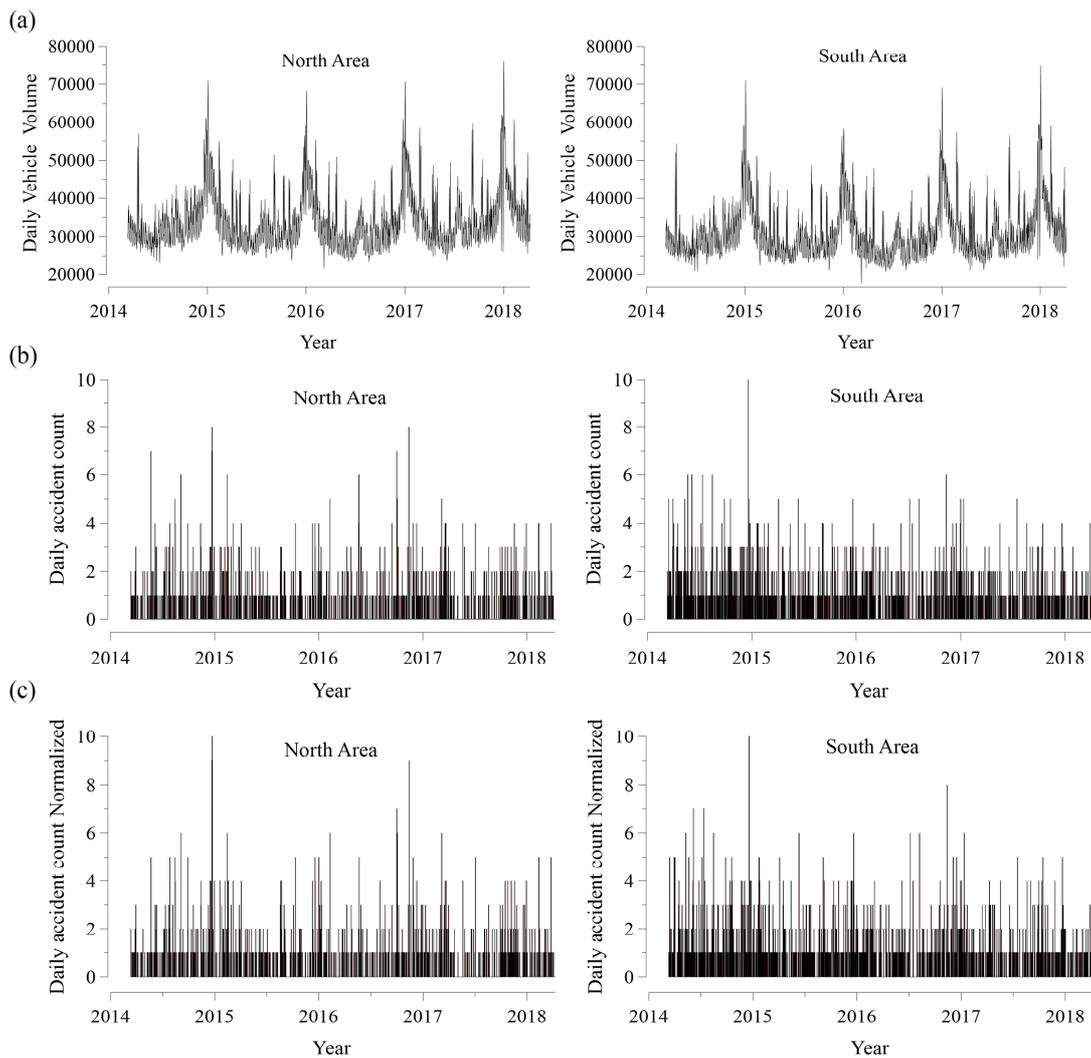|	North Area	South Area
Number of observations (days)	1490	
Total Crashes	2314	
Crashes per area	980	1334
Crashes per Season		
DJF	291	362
MAM	234	335
JJA	203	310
SON	252	327
Crashes per vehicle type		
Light Vehicles	829	836
Heavy Vehicles	212	576

Figure 2. (a) Daily vehicle volume (V), (b) Daily accident count (C), (c) Daily accident count normalized (VNC)

Data from traffic volume (Figure 2a) comes from toll stations registers of "São José dos Pinhais" and "Garuva" municipalities, for North and South Areas, respectively. The register ranges from 2014 to 2018. Throughout the five years, both areas showed a similar distribution of traffic volume, with peaks above the average mostly between the end of December and January, which is mostly associated with Christmas and Summer holidays, as the highway connects the state capital and other cities to the Santa Catarina state coast.

The daily accident count is illustrated in Figure 2b, for North and South Areas. The South Area presents a higher average daily accident count compared to the North Area, with values around 0.92 and 0.65, respectively. Traffic volume data of the highway (Figure 2a) was used to normalize the road accident count (Figure 2b).

The normalization approach was used to account for traffic volume exposure. The detailed procedure adopted to carry out the normalization is described in section 3.4. As shown in Figure 2b, it is evident that in both areas there are expressive fluctuations in traffic volume during the years, with much greater traffic during Christmas season.

3.4. Negative Binomial Model

In order to account for traffic exposure, the model considered the application of a volume-normalized traffic crash count. This consisted of applying Equation 1 to the available data.

$$VNC = \frac{C}{V} \tag{1}$$

And:

$$V = \frac{V_d}{AADT} \tag{2}$$

Where VNC (Equation 1) is the normalized crash count, C stands for the daily traffic crash count, and V (Equations 1 and 2) is the volume normalization factor, calculated by dividing the daily traffic volume (V_d) by the Annual Average Daily Traffic (AADT). This same approach has been used by Keay and Simmonds (2006) and, as stated by the authors, it allows variations in daily volume relative to days of the week, as well as holidays, to be incorporated to the model.

The model was developed in the R environment. Initial trials of modelling took into consideration the application of a Poisson distribution (Ma and Kockelman, 2006; Li *et al.*, 2013; Ye *et al.*, 2013). However, the data did not meet the assumptions required, being over-dispersed. Thus, a Negative Binomial (NB) model (Eisenberg, 2004; Keay and Simmonds, 2006; Poch and Mannering, 2013) was chosen, since it is able to account for data dispersion. Its application in the present study can be described as shown in Equation 3.

$$ln(VNC) = X_i + \varepsilon \tag{3}$$

Where X_i consists of the vector representing the explanatory variables on the highway segment and time period, and ε is the model error. It should be noted that ε has a gamma distribution (Hauer *et al.*, 1988), which the mean value is 1 and has a variance of $\theta = 1/\alpha^2$, with θ being defined as the dispersion parameter. If θ can be assumed as 1, it shows evidence that NB model is preferable than the Poisson model. That was the case for every model presented herein, justifying the choice for NB methodology.

Concerning the explanatory variables, the decision criteria of whether a subset of independent variables should or not be added to the crashes estimation model consisted of an application of the Akaike's Information Criterion (AIC). The procedure started considering the model dummy variables for days of the week (DOW), Christmas period, Brazilian Calendar Holidays and Rain classes. If the Holiday dummy is removed, there is an improvement in the

model fitting, however the same did not happen when Christmas variable was left out, thus Christmas was maintained in the model. That is the same case for DOW when removed, hence they were also not removed. This procedure was not applied to the Rain classes' dummies, since these variables constitute the main interest of this work. The final *Xi* is presented in Equation 4.

$$X_i = \alpha_j + \beta_1 Mo + \beta_2 Tu + \beta_3 We + \beta_4 Th + \beta_5 Fr + \beta_6 Sa +$$
$$+ \beta_7 X + \beta_8 Drizzle + \beta_9 Light + \beta_{10} Moderate + \beta_{11} Heavy$$
$$+ \beta_{12} Extreme$$

(4)

Where α_j is the model intercept, *Mo* (Monday) through *Sa* (Saturday) are dummy variables representing DOW (*Mo* equals 1 if it is a Monday, and 0 if another day), *Tu* through *Sa* follow the same definition. The model does not include "Sunday" variable due to the linear restriction imposed on the fact that there are 7 days of the week, thus if there was a *Su* variable, it would always equal zero. Thus, the model coefficients β_1 through β_6 represent the day effect in the accident occurrence in comparison to Sunday. The same applies to the other variables representing Christmas, "*X*", and "*Drizzle*" through "*Extreme*", which are dummies that insert weather variables into the model, i.e., rainfall classes. In this case the coefficients β_8 through β_{12} will dictate the difference in means from a rainy day (e.g., if it rained 5 mm, it will fit in the Light dummy).

Finally, for overall goodness-of-fit assessment, the approach selected was the evaluation of the deviance value, *2(LL(β)−LL(0))*, which follows a squared-chi distribution (Agresti, 2003). The criteria consists of determining the model deviance, and compare that to the chi-square value corresponding to the degree of freedom of the model. A deviance value greater than the chi-square value should indicate that the evaluated model presents an overall adequate goodness-of-fit in representing the data.

4. RESULTS AND DISCUSSION
4.1. Accident Frequency Modeling

The relative coefficients for North and South Areas presented in Table 2 can be interpreted as changes in incidence rates for unit change in crash occurrence (Eisenberg, 2004). Hence, a coefficient of 0.01, for example, shows a 1% increase in accidents. Similarly, negative coefficients would mean a reduction in the crash count. The comparison for DOW dummies should be made with respect to Sunday, while for Precipitation, with respect to no precipitation occurrence (0.0 to 0.1 mm of daily rainfall). As mentioned, the deviance values were used in the determination of the model overall goodness-of-fit. For both models, the chi-squared test of deviance values, North (deviance of 198) and South (deviance of 125), rejects the null hypothesis, thus confirming the data distribution has an adequate fitting with respect to NB regression.

For the North Area model, all but Drizzle (0.1–2.5 mm) rain classes are significant to the model. Additionally, North and South Area coefficients show that with the increase of daily precipitation, a greater number of accidents followed, hence confirming this positive relationship on a daily analysis. For this area, rainfall dummies coefficients are positive, showing that in any precipitation condition (rainfall greater than 0.1 mm) accident frequency is increased. For daily rainfall greater than 50 mm, an increase of 6% would be expected, when compared to a dry day. Very similar result checks for the South Area model, where all precipitation coefficients were significant. Moreover, all rainfall coefficients present a consistent growth with the increase in the rain intensity. As in the North Area model, no negative

rainfall-crash frequency relationship was found. In general, the rain effect observed in the South Area is smaller than that observed for the North Area.

Table 2 – Accident frequency results in North and South Areas of study

Variables	North Area Coefficient	North Area z-statistics	South Area Coefficient	South Area z-statistics
Monday	0.009	-0.96	0.010	0.20
Tuesday	0.007	-2.16	0.008	-1.70
Wednesday	0.008	-1.76	0.008	-1.67
Thursday	0.007	-2.69	0.009	-0.73
Friday	0.010	-0.14	0.011	1.12
Saturday	0.011	1.06	0.011	0.52
Christmas Period	0.016	2.42	0.011	0.55
Drizzle [0.1 mm, 2.5 mm]	0.011	0.69	0.014	3.66
Light precipitation [2.5 mm, 10 mm]	0.025	6.63	0.021	8.16
Moderate precipitation [10 mm, 25 mm]	0.034	8.34	0.022	7.73
Heavy precipitation [25 mm, 50 mm]	0.039	7.58	0.025	5.81
Extreme precipitation [50 mm, ∞ mm]	0.058	6.83	0.037	5.16
Summary of statistics				
Days analyzed	1490		1490	
Log-lik at zero	-1657		-1985	
Log-lik at convergence	-1558		-1922	
$2(LL(\beta)-LL(0))$	198		125	
$\chi 2$, df = 12	21.03		21.03	

Despite the lack of significance of some DOW dummies for the model, an important insight is that in all days that were presented in Table 2 (i.e., from Monday to Saturday), there is a tendency of greater crashes number than that would be verified in a Sunday. Friday and Saturday would be the most critical days in terms of increased accident risk. As abovementioned, the positive association between precipitation and crash frequency is consistent with previous reports. A model developed for crashes occurrence in a mountainous freeway in the Colorado state of the USA (Yu and Abdel-Aty, 2013) found that risk of crash is greater on weekends (Saturday and Sunday), and weekend crashes are more likely to happen under conditions of smooth traffic flow, with no speed variations.

4.2. Seasonal Effect in Accident Frequency Modeling

Eight models of daily accident frequency were estimated in order to quantify the effect of rain in each season, with four models for each segment. In these models, all variables from Equation 4 were included, with the exception of the Christmas variable, since it would appear only in the DJF period (Southern hemisphere summer). Although a few insights would follow from including the DOW dummies in the Seasonal model, only rainfall coefficients were added to this analysis, focusing on the effects of weather conditions throughout the year. The same procedure as for the DOW analysis was performed for goodness-of-fit on Seasonal effect model. Since all models presented overall adequate goodness-of-fit, this information was omitted from Table 3, to avoid data overcrowding.

Table 3 – Seasonal effect in accident frequency results

Method: negative binomial regression	Model - Seasons			
	DJF	MAM	JJA	SON
North Area				
Number of observations	362	398	369	365
Drizzle [0.1 mm, 2.5 mm]	0.012 (0.50)	0.012 (0.77)	0.0136 (1.24)	0.007 (-1.24)
Light precipitation [2.5 mm, 10 mm]	0.027 (2.94)	0.024 (3.19)	0.0309 (4.50)	0.020 (2.58)
Moderate precipitation [10 mm, 25 mm]	0.044 (4.38)	0.028 (3.61)	0.0367 (4.35)	0.030 (3.79)
Heavy precipitation [25 mm, 50 mm]	0.046 (4.03)	0.042 (4.08)	0.0400 (3.89)	0.034 (3.33)
Extreme precipitation [50 mm, ∞ mm]	0.096 (4.83)	0.044 (3.16)	0.0376 (1.72)	0.043 (1.88)
South Area				
Number of observations	362	398	369	365
Drizzle [0.1 mm, 2.5 mm]	0.013 (1.47)	0.011 (0.65)	0.016 (2.62)	0.016 (2.40)
Light precipitation [2.5 mm, 10 mm]	0.020 (3.87)	0.020 (3.70)	0.020 (3.32)	0.019 (3.18)
Moderate precipitation [10 mm, 25 mm]	0.019 (3.15)	0.018 (2.87)	0.033 (5.49)	0.024 (4.01)
Heavy precipitation [25 mm, 50 mm]	0.020 (2.33)	0.017 (1.48)	0.043 (4.50)	0.024 (3.03)
Extreme precipitation [50 mm, ∞ mm]	0.021 (1.79)	0.043 (3.46)	0.067 (3.04)	0.034 (1.72)

Note: 95% confidence interval. Coefficients presented are relative coefficients, and in parenthesis are the respective z-statistics values.

Similarly to what was found in the entire year analysis, there was a distinct seasonality effect in the North and South segments. First, for the North Area, Drizzle intensity was reported as non-significant, which is coherent with the entire period model. Hence, no matter the time of the year, this range seems not to be of considerable effect in the risk of crash for the North Area. In contrast, the other classes are significant. Coefficients show that DJF (i.e., summer) season is the most critical in terms of rainfall influence in crashes, with an increased risk of up to 10% (for Extreme precipitation).

In fact, for all rainfall classes, summer is the season where rainfall substantially influences crash frequency. This is generally the wettest season, and it could have been expected that drivers would be more careful at this time of the year (Keay and Simmonds, 2006; Mannering and Bhat, 2014). Nonetheless, that was not observed for this particular area. It is deduced that the reason for this is because the segment consists mostly of streamline extensions, or light curves, without accentuated curves or high road grade, which possibly gives the drivers a false impression of safety. However, for other seasons it seems that there is no clear tendency of a more critical period or intensity that outstands in affecting crash frequency.

In parallel, the South Area presented generally lighter rain effect than the North Area. The greatest contrast here is that JJA (i.e. winter) season was the most critical, with up to 7% more crashes at Extreme precipitation. Summer, however, was the season with less expressive effects. This highlights the importance on segmenting the region for the analysis. The South Area has greater grades and an increased incidence of curves with large degrees (Figure 1b), and the segment also contrasts with North in terms of radars and maximum speed, as it has a higher number of speed radars (three), and higher maximum speed allowed (80 km/h). These are some of the underlying factors that might be shaping the difference between the two study areas.

Another possible explanation for that is the evidence that the degree of curvature may have a negative relationship with accidents (Anastasopoulos *et al.*, 2008). This is an indication that

segments with sharper curves could be in safer conditions than those where curves are flatter. In concern to the South Area seasonal model, the increased presence of sharp curves may have resulted in the reduced rainfall effect. It's likely that the sharper curves make drivers more aware and/or that they compensate for the increased risk by driving more slowly, resulting in lower accident rates (Anastasopoulos *et al.*, 2008).

The effect of seasonality on crash risk under rainfall conditions for an urban freeway in Melbourne, Australia (Keay and Simmonds, 2006) indicated that for the wettest season, spring, the precipitation influence on accident risk was reduced when compared to the other seasons. This observation is coherent to the results observed in the South Area of the present study, in which for the wettest season (Summer), rainfall effect on crash frequency was reduced in comparison with other seasons.

A previous study model on a mountainous area (Ahmed *et al.*, 2012) suggested that high precipitation and speed variation were factors that caused increased crash risk; and unexpectedly, in the dry season, low speeds together with low visibility increased the odds of accident occurrence. Hence, the interaction of weather and traffic variables was different in the two seasons, indicating distinctive combined factors.

4.3. Vehicle category effect in accident frequency modeling

Four models of daily accident frequency were estimated in order to investigate the rain effect in two categories of vehicles: light and heavy vehicles. This division was based in the Brazilian traffic code (Brasil, 2008). In summary, motorcycles, automobiles and others, such as carriages, are considered "light" vehicles. Trucks (articulated and non-articulated) and buses are categorized as "heavy".

As presented for the seasonal modeling, DOW coefficient and goodness-of-fit results are also omitted from the table, since overall values were adequate. It is possible to identify that in the North Area rain is a more significant factor for light vehicles, but also has an expressive influence in heavy vehicles (see Table 4). As for the other models, Drizzle was not significant for light vehicles, however the same does not happen for heavy vehicles, where there was a 2% increase in crashes. For the South Area, light vehicles consistently showed an increase in crashes followed by higher rainfall classes, with plus 7% accidents for Extreme precipitation. The results for heavy vehicles in this area, though, are remarkable. All precipitation classes' coefficients were encountered as non-significant. Hence, they seem to have minimal impact in explaining crash frequency, suggesting that factors other than rainfall are shaping accident frequency in the extension of South Area for this vehicle category.

In order to address the reason behind the non-significance of rain regarding the South Area, an accident type analysis with respect to heavy vehicles was key. We bring the analysis of the two accident categories, involving heavy vehicles, which are the most frequent in each of the two areas. The first category is "runway exit", which in the North Area corresponds to 49% (104) of total heavy vehicles accidents. Meanwhile, in the South Area this category represents a lesser proportion, of 33% (192) of the total heavy vehicles' crashes.

The second category is overturning. While in the North Area, 25% (53 in absolute numbers) of total accidents involving this vehicle category is "overturning", in the South Area, 43% (212) accidents fit into this crash type. In addition, the total crash number for heavy vehicles verified in the South Area are 171% greater than the North Area total numbers (576 for South Area, compared to 212 for North Area).

This information can be analyzed in relative and absolute terms. Although the herein developed model did not included geometric characteristics, they can be qualitatively included in the discussion. The South Area has not only a greater number of horizontal curves, but also sharper ones. These are likely inducing the accidents in the overturning category (which, as mentioned, is the area's most common accident type for heavy vehicles), whether or not rain is present. In turn, in the North Area, there are fewer heavy vehicle accidents in absolute numbers, along with runway exits holding the greatest number by accident types. Since this kind of accident is directly related to road slick, and rain presence and/or intensity is crucial in determining slipperiness, the model is capturing rain as un underlying factor (a model significant factor), causing the runway exit accidents in the North Area, while in the South Area it could not be captured (i.e., not significant), likely due to overturning preponderancy in the crashes.

Table 4 – Vehicle category effect results in accident frequency

	Model–Vehicle weight			
Method: negative binomial regressions				
	Light		Heavy	
North Area				
Number of observations	1490		1490	
Drizzle [0.1 mm, 2.5 mm]	0.010	(-0.12)	0.016	(1.79)
Light precipitation [2.5 mm, 10 mm]	0.026	(6.25)	0.017	(2.14)
Moderate precipitation [10 mm, 25 mm]	0.035	(7.80)	0.026	(3.54)
Heavy precipitation [25 mm, 50 mm]	0.040	(7.13)	0.032	(3.89)
Extreme precipitation [50 mm, ∞ mm]	0.056	(6.38)	0.044	(3.38)
South Area	Light		Heavy	
Number of observations	1490		1490	
Drizzle [0.1 mm, 2.5 mm]	0.021	(5.75)	0.010	(-0.05)
Light precipitation [2.5 mm, 10 mm]	0.036	(9.40)	0.012	(0.20)
Moderate precipitation [10 mm, 25 mm]	0.041	(9.50)	0.010	(-0.01)
Heavy precipitation [25 mm, 50 mm]	0.055	(8.48)	0.010	(-0.03)
Extreme precipitation [50 mm, ∞ mm]	0.067	(5.71)	0.018	(0.58)

Note: 95% confidence interval. Coefficients presented are relative coefficients, and in parenthesis are the respective z-statistics values.

Analysis of previous literature showed that vehicles most affected by rainfall are from non-work-related travels (i.e. leisure) in automobiles (Bijleveld and Churchill, 2009). In addition, single vehicles accidents seem to be the ones most affected by weather (Satterthwaite, 1967).

With respect to the behavior of accidents with heavy vehicles, a significant factor that can imply greater crash numbers is wind. A previous study indicated that in wind-induced accidents that resulted in injuries and/or fatalities, only 27% of vehicles involved were light vehicles, with the majority of accidents involving taller, heavier vehicles (Baker and Reynolds, 1992). Moreover, the most common accident for this type of vehicle is overturning, induced by their broader area of contact with the wind. Additionally, for a specific highway segment, other variables affecting heavy vehicles accidents frequency are the average daily traffic, the percent of heavy vehicles in traffic, and the difference in speed between heavy and light vehicles (Joshua and Garber, 1990).

Logistic regression models were previously used to explore the probability of involvement in a truck crash on road segments with negative grades, i.e., downgrade (Momen *et al.*, 2019). In clear weather this probability was reduced when compared to the adverse condition, confirming the positive effect that unfavorable weather has in increasing crash occurrence.

4.4. Remarks from Accident Frequency Modeling

The models developed in this work aimed at interpreting and correlating time-course of rainfall variations on BR-376, therefore the intrinsic characteristics of the highway (such as geometrics) were not considered for this specific analysis as it does not correspond to the research's aim. The literature indicates that conclusive research on rain effect on crashes can be built, in spite of the absence of space related characteristics, which include geometrics, pavement surface, and others (Eisenberg, 2004; Keay and Simmonds, 2006; Mannering, 2014). Also, the implementation of rainfall data obtained *in situ* for each of the regions of interest in this study is of great relevance to address particularities of the segments of the Highway.

In addition to the models presented, there was an attempt to develop a model to evaluate the effect of rainfall in crashes with distinct injury severity using separate samples, just like for vehicle classes. However, this analysis did not fit in a NB regression model. Opportunities for further research include exploring the relationship of severity/injury with rainfall variables, while also considering new methodological approaches. Other factors could be investigated, such as driver's gender, age, type of accident and downgrade/upgrade distinction. Policy lessons can be drawn from the results here presented, which can be useful not only for governmental traffic management agencies, but also to highway concessionaires.

5. CONCLUSIONS

This study presents a daily accident frequency modeling based on a NB regression, focusing on local rainfall effect, with precipitation data provided by in-field rain gauges. In this specific area of the BR-376, no past research has explored the effects of precipitation in traffic crashes numbers.

The method considered two separate geographical areas (North and South), where three models were built. Results from all models consistently showed that rain seems to be a significant factor in accident frequency, and that under rainfall conditions crash frequency tends to be higher. Moreover, when the road is subjected to greater rain intensity, there is a tendency of steady increase on the number of crashes. In the seasonal analysis, rain presence and intensity affected positively crash numbers for the two geographical areas modeled. Nevertheless, the two areas indicated different critical seasons, in terms of rain influence. For the North Area, summer was the period in which rain mostly affected the crash frequency, whilst for the South, winter seemed to be the season in which that occurs. For vehicle type analysis, strong rainfall influences in crashes numbers for both light and heavy vehicles in the North Area. For the South Area, the same effect was found only for light vehicles, whereas rain presence/intensity seemed to be a non-significant factor for heavy vehicles, which we suspect to be linked with the crash type (majority overturning in the South Area).

Rainfall should raise an alert to safety agencies, as well as to road management, since greater accident number is expected. Management should consider implementing traffic signs indicating increased accident risk in a rainy situation, and specific actions can be taken for each of the regions analyzed when considering the most critical seasons. These actions could be

applied for the entire year, but more locally specified, since in the South Area the biggest concern seems to be during winter, while for the North Area, summer appeared to be the most critical period of the year.

REFERENCES

Abdel-Aty, M.A. and A.E., Radwan (2000) Modeling traffic accident occurrence and involvement. *Accident Analysis & Prevention*, v. 32, n. 5, p. 633-642. DOI: 10.1016/S0001-4575(99)00094-9.

Ahmed, M.M.; M. Abdel-Aty and R. Yu (2012) Assessment of interaction of crash occurrence, mountainous freeway geometry, real-time weather, and traffic data. *Transportation research record*, v. 2280 n. 1, p.51-59. DOI: 10.3141/2280-06.

de Almeida, F.F.M. and C. D. R. Carneiro (2017) Origem e evolução da Serra do Mar. *Revista Brasileira de Geociências*, v. 28, n. 2, p. 135-150. Available at <http://bjg.siteoficial.ws/1998/n.2/3.pdf> (Access in 26/05/2021).

Anastasopoulos, P.C.; A. P. Tarko. and F. L. Mannering (2008). Tobit analysis of vehicle accident rates on interstate highways. *Accident Analysis & Prevention*, v. 40, n. 2, p. 768-775. DOI: /10.1016/j.aap.2007.09.006.

Baker, C.J. and S. Reynolds (1992) Wind-induced accidents of road vehicles. *Accident Analysis & Prevention*, v. 24, n. 6, p. 559-575. DOI: 10.1016/0001-4575(92)90009-8. DOI: 10.1016/0001-4575(92)90009-8.

Bergel-Hayat; R.; M. Debbarh; C. Antoniou and G. Yannis (2013) Explaining the road accident risk: weather effects. *Accident Analysis & Prevention*, v. 60, p. 456-465. DOI: 0.1016/j.aap.2013.03.006.

Bijleveld, F. and T. Churchill (2009) *The influence of weather conditions on road safety*. Leidschendam (The Netherlands): SWOV Institute for Road Safety Research.

Brasil. *Código de Trânsito Brasileiro* (CTB). Lei Nº 9.503, de 23 de setembro de 1997 que institui o Código de Trânsito Brasileiro. Presidência da República. Casa Civil. Subchefia para Assuntos Jurídicos. Brasília: 2007.

Chang, L.Y. and W. C. Chen (2005) Data mining of tree-based models to analyze freeway accident frequency. *Journal of safety research*, v. 36, n.4, p. 365-375. DOI: 10.1016/j.jsr.2005.06.013.

Dabbour, E.; M. Haider and E. Dias (2019) Using random-parameter and fixed-parameter ordered models to explore temporal stability in factors affecting drivers' injury severity in single-vehicle collisions. *Journal of traffic and transportation engineering (English edition)*, v. 6, n. 2, p. 132-146. DOI: 10.1016/j.jtte.2018.04.002.

Eisenberg, D. (2004) The mixed effects of precipitation on traffic crashes. *Accident analysis & prevention*, v. 36, n. 4, p. 637-647. DOI: 10.1016/S0001-4575(03)00085-X.

Ferreira, K. S. M.; L. P. Sestrem; A. M. G. Acevedo; A. C. M. Kormann and V. P. Faro (2018) Instrumentação geotécnica como ferramenta de gestão de riscos rodoviários: um estudo de caso no trecho de serra do mar da BR 376/PR. *Revista Técnico-Científica*, v. 1, n. 11. Available at <http://creaprw16.creapr.org.br/revista/sistema/index.php/revista/article/view/400> (Access in 26/05/2021).

Gabler, H.C. and W. T. Hollowell (1998) The aggressivity of light trucks and vans in traffic crashes. *SAE transactions*, v. 107, p. 1444-1452. Available at <www.jstor.org/stable/44741082> (Access in 26/05/2021)

González Acevedo, A. M. G.; L. B. Passini and A. C. M. Kormann (2017) Rainfall Effects on Pore Pressure Changes in a Coastal Slope of the Serra do Mar in Santa Catarina. *Soils and Rocks*, v. 40, n. 3, p. 263-278. DOI: 10.28927/SR.403263.

González Acevedo, A.M. (2017) *Simulação geologico-geotécnica para avalização de estabilidade de taludes a partir de técnicas de geoprocessamento*. Thesis (Doctorate). Programa de Pós-graduação em Engenharia de Construção Civil, Federal University of Paraná. Curitiba. Available at <https://www.acervodigital.ufpr.br/handle/1884/52842> (Access in 26/05/2021).

Hauer, E.; J. C. Ng and J. Lovell (1988) Estimation of safety at signalized intersections. *Transportation Research Record*, v. 1185, p. 48-61. Available at <http://onlinepubs.trb.org/Onlinepubs/trr/1988/1185/1185-005.pdf> (Access in 26/05/2021).

Hasui, Y., C. D. R. Carneiro; F. F. M. de Almeida and A. Bartorelli (2012) *Geologia do Brasil* (p. 900). São Paulo: Beca.

Hermans, E.; T. Brijs; T. Stiers and C. Offermans (2006) The impact of weather conditions on road safety investigated on an hourly basis. In *Washington, DC: Proceedings of the 85thAnnual meeting of the Transportation Research Board*.

Joshua, S.C. and N. J. Garber (1990) Estimating truck accident rate and involvements using linear and Poisson regression models. *Transportation planning and Technology*, v. 15, n.1, p. 41-58. DOI: 10.1080/03081069008717439.

Keay, K. and I. Simmonds (2006) Road accidents and rainfall in a large Australian city. *Accident Analysis & Prevention*, v. 38, n. 3, p. 445-454. DOI: 10.1016/j.aap.2005.06.025.

Levine, N.; K. E. Kim and L. H. Nitz (1995) Daily fluctuations in Honolulu motor vehicle accidents. *Accident Analysis & Prevention*, v. 27, n. 6, p. 785-796. DOI: 10.1016/0001-4575(95)00038-0.

de Lima Leite, M.; P. A. Adacheski and J. S. das Virgens Filho (2011) Análise da frequência e da intensidade das chuvas em Ponta Grossa, Estado do Paraná, no período entre 1954 e 2001. *Acta Scientiarum Technology*, v. 33, n. 1, p. 57-54. DOI: 10.4025/actascitechnol.v33i1.6957.

Li, Z.; W. Wang; P. Liu; J. M. Bigham and D. R. Ragland (2013) Using geographically weighted Poisson regression for county-level crash modeling in California. *Safety science*, v. 58, p. 89-97. DOI: 10.1016/j.ssci.2013.04.005.

Ma, J. and K. M. Kockelman (2006) Bayesian multivariate Poisson regression for models of injury count, by severity. *Transportation Research Record*, v. 1950, n.1, p. 24-34. DOI: 10.1177/0361198106195000104.

Mannering, F. L. and C. R. Bhat (2014) Analytic methods in accident research: Methodological frontier and future directions. *Analytic methods in accident research*, v.1, p. 1-22. DOI: 10.1016/j.amar.2013.09.001.

Martensen, H.; N. Focant and K. Diependaele (2016) Let's talk about the weather–Interpretation of short term changes in road accident outcomes. *Transportation research procedia*, v. 14, p. 96-104. DOI: 10.1016/j.trpro.2016.05.045.

Moomen, M.; M. Rezapour and K. Ksaibati (2019) An investigation of influential factors of downgrade truck crashes: A logistic regression approach. *Journal of traffic and transportation engineering (English edition)*, v. 6, n.2, p. 185-195. DOI: 10.1016/j.jtte.2018.03.005.

Poch, M. and F. Mannering (1996) Negative binomial analysis of intersection-accident frequencies. *Journal of transportation engineering*, v. 122, n. 2, p. 105-113. DOI: 10.1061/(ASCE)0733-947X(1996)122:2(105).

Satterthwaite, S. P. (1976) An assessment of seasonal and weather effects on the frequency of road accidents in California. *Accident Analysis & Prevention*, v. 8, n. 2, p. 87-96. DOI: 10.1016/0001-4575(76)90002-6.

Sehaber V. F. (2013) *Análise estatística multivariada dos acidentes de trânsito da BR-116 no período entre os anos de 2009 e 2012*. Dissertation (Masters). Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Federal University of Paraná. Curitiba. Available at <http://hdl.handle.net/1884/32018> (Access in 26/05/2021).

Silva M. A.; I. C. Salah; A. M. A. González; L. P. Sestrem; V. P. Faro; A. C. M. Kormann and L. B. Passini (2019) Pluviometric thresholds application to predict slope failures along BR-376 Highway, in the state of Paraná – Brazil. In *Cancun – México: Proceedings of the XVI Congreso Panamericano de Mecánica de Suelos e Ingeniería Geotécnica.* Available at <https://ebooks.iospress.nl/pdf/doi/10.3233/STAL190227> (Access in 26/05/2021)

Silva, M.A.; L. C. Guideli; A. C. Neto; L. B. Passini and A. C. M. Kormann (2020) Multiple correspondence analysis applied to the study of the relationship between traffic crashes and precipitation on a highway in Brazil. *Transportes*, v. 28, n. 3, p. 196-211. DOI: 10.14295/transportes.v28i3.2044.

Souza, W.M.; P. V. de Azevedo and L. E. de Araújo (2012) Classificação da precipitação diária e impactos decorrentes dos desastres associados às chuvas na cidade do Recife-PE. *Revista Brasileira de Geografia Física*, v. 5, n. 2, p. 250-268. DOI: 10.26848/rbgf.v5i2.232788.

Theofilatos, A. and G. Yannis (2014) A review of the effect of traffic and weather characteristics on road safety. *Accident Analysis & Prevention*, v. 72, p. 244-256. DOI: 10.1016/j.aap.2014.06.017.

Trevizolli, M. N. B.; A. C. M. Kormann; V. P. Faro; L. P. Sestrem and L. B. Passini (2018) Análise de registros históricos de precipitações e rupturas de encostas para composição de limiares pluviométricos visando monitoramento de taludes rodoviários. *Revista Técnico-Científica*, v. 1 n. 11. Available at < http://creaprw16.crea-pr.org.br/revista/sistema/index.php/revista/article/view/397> (Access in 26/05/2021).

Victorino M. M. (2015) *Influência da chuva nas poropressões e estabilidade dos taludes rodoviários de um trecho da BR 376 na Serra do Mar paranaense*. Dissertation (Masters). Programa de Pós-graduação em Engenharia de Construção Civil, Federal University of Paraná. Curitiba. Available at <https://www.acervodigital.ufpr.br/handle/1884/40608> (Access in 26/05/2021).

Vieira, B.C. and M. F. Gramani (2015) Serra do Mar: the most "tormented" relief in Brazil. In *Landscapes and landforms of Brazil* (p. 285-297). Dordrecht (The Netherlands): Springer.

Winston, C.; V. Maheshri and F. Mannering (2006) An exploration of the offset hypothesis using disaggregate data: The case of airbags and antilock brakes. *Journal of Risk and Uncertainty*, v. 32, n.2, p. 83-99. DOI: 10.1007/s11166-006-8288-7.

Xavier, T. M. B. S. and A. F. S. Xavier (1987) Classificação e monitoração de períodos secos ou chuvosos e cálculo de índices pluviométricos para a região Nordeste do Brasil. *Revista Brasileira de Engenharia-Caderno de Recursos Hídricos*, v. 5, n.2, p. 7-31.

Ye, X.; R. M. Pendyala; V. Shankar and K. C. Konduri (2013) A simultaneous equations model of crash frequency by severity level for freeway sections. *Accident Analysis & Prevention*, v. 57, p. 140-149. DOI: 10.1016/j.aap.2013.03.025.

Yu, R. and M. Abdel-Aty (2013) Investigating the different characteristics of weekday and weekend crashes. *Journal of safety research*, v. 46, p. 91-97. DOI: 10.1016/j.jsr.2013.05.002.

Zhang, C.; J. N. Ivan; W. M. El Dessouki and E. N. Anagnostou (2005) Relative risk analysis for studying the impact of adverse weather conditions and congestion on traffic accidents. In *Transportation Research Board – 84th Annual Meeting*, Washington, USA.