G Model
AAP-1952; No. of Pages 12
ARTICLE IN PRESS

Accident Analysis and Prevention xxx (2009) xxx–xxx

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Rainfall effect on single-vehicle crash severities using polychotomous response models

Soyoung Jung[a,*], Xiao Qin[b,1], David A. Noyce[c,2]

[a] University of Wisconsin-Madison, Department of Civil and Environmental Engineering, 1241 Engineering Hall, 1415 Engineering Drive, Madison, WI 53706-1691, United States
[b] University of Wisconsin-Madison, Department of Civil and Environmental Engineering, B239 Engineering Hall, 1415 Engineering Drive, Madison, WI 53706-1691, United States
[c] University of Wisconsin-Madison, Department of Civil and Environmental Engineering, 1210 Engineering Hall, 1415 Engineering Drive, Madison, WI 53706-1691, United States

ARTICLE INFO

Article history:
Received 6 March 2009
Received in revised form 25 July 2009
Accepted 29 July 2009

Keywords:
Crash severity
Ordinal logistic regression
Sequential logistic regression
Rainfall effect
Road weather safety

ABSTRACT

As part of the Wisconsin road weather safety initiative, the objective of this study is to assess the effects of rainfall on the severity of single-vehicle crashes on Wisconsin interstate highways utilizing polychotomous response models.

Weather-related factors considered in this study include estimated rainfall intensity for 15 min prior to a crash occurrence, water film depth, temperature, wind speed/direction, stopping sight distance and deficiency of car-following distance at the crash moment. For locations with unknown weather information, data were interpolated using the inverse squared distance method. Non-weather factors such as road geometrics, traffic conditions, collision types, vehicle types, and driver and temporal attributes were also considered. Two types of polychotomous response models were compared: ordinal logistic and sequential logistic regressions. The sequential logistic regression was tested with forward and backward formats. Comparative models were also developed for single vehicle crash severity during clear weather.

In conclusion, the backward sequential logistic regression model produced the best results for predicting crash severities in rainy weather where rainfall intensity, wind speed, roadway terrain, driver's gender, and safety belt were found to be statistically significant. Our study also found that the seasonal factor was significant in clear weather. The seasonal factor is a predictor suggesting that inclement weather may affect crash severity. These findings can be used to determine the probabilities of single vehicle crash severity in rainy weather and provide quantitative support on improving road weather safety via weather warning systems, highway facility improvements, and speed limit management.

1. Introduction

Driving in rainy conditions is more challenging than in clear weather due to low visibility and slippery road conditions. These challenges contribute to a sizable portion of severe crashes including fatalities and injuries in Wisconsin. According to Wisconsin Traffic Crash Facts from the Wisconsin Department of Transportation (WisDOT), there are 3047 injury and fatal crashes in rainy weather, the greatest number of all kinds of inclement weather conditions from 1999 to 2005. In addition, the proportion of injury and fatal crashes to total crashes in rainy weather is 0.37, the second highest during the same period. Despite the fact that fog-related crashes have the highest injury and fatal crash proportion, the total number of injury and fatal crashes in foggy weather is merely 249 from 1999 to 2005, less than 10% of injury and fatal crashes that occurred in rainy weather.

In Wisconsin, rainy weather-related crash severity is distributed differently between single and multi-vehicle crashes. From 1999 to 2006, 659 single-vehicle crashes and 899 multi-vehicle crashes occurred on Wisconsin interstate highways in rainy weather. The proportion of serious crashes including fatalities and incapacitating injuries to total crashes is 10% for single-vehicle crashes while it is only 4% for multi-vehicle crashes. Since the consequence of a single-vehicle crash is more serious, there is a need to investigate the contributing factors to single-vehicle crash severity, especially fatal or incapacitating injury severity.

Crash consequences can be minimized through the improvements in roadway and roadside design, appropriate use of safety devices, and changes in driver behavior. Numerous studies have been conducted by safety researchers and practitioners in hopes of identifying the contributing factors to crash severities (Abdel-Aty and Pemmanaboina, 2006a; Dissanayake and Lu, 2002; Donnell and Mason, 2004; Golob and Recker, 2003; Khan et al., 2008; Qin

* Corresponding author. Tel.: + 1 608 262 2524.
E-mail addresses: jung2@wisc.edu (S. Jung), xqin@engr.wisc.edu (X. Qin), noyce@engr.wisc.edu (D.A. Noyce).
[1] Tel.: +1 608 262 3649; fax: +1 608 262 5199.
[2] Tel.: +1 608 265 1882; fax: +1 608 262 5199.

0001-4575/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aap.2009.07.020

et al., 2006; Savolainen and Tarko, 2005; Shankar et al., 1995; Yau, 2004; Yau et al., 2006). Overall, research demonstrates that such factors as weather, roadway geometries, traffic conditions, driver, and temporal-related predictors as well as their interaction lead to different injury severities. In many crash severity-based studies, weather is just one of the contributing factors, not necessarily the focus. Using only rainy weather related crashes, this study identified a variety of significant predictors that contribute more serious crash consequences. Furthermore, a comparison between rainy weather and clear weather crash prediction models revealed important factors that may potentially lead to appropriate countermeasures against severe crash occurrences in rainy weather.

2. Literature review

The topic of crash severity has been of interest to traffic safety community because of the direct impact on occupants involved. Weather is frequently cited and found as one of the factors contributing to either a more or less severe crash. The approaches used to model injury severities vary from one to another, depending on the purpose of the study and data availability.

Based on the purpose of this research, the focus of this literature review will be specifically on factors caused by rain precipitation at the time of the crash.

For crash severity models, the crash outcome of a single vehicle or multi-vehicle crash is usually modeled separately. Malyshkina and Mannering (2009) explained unobserved heterogeneity related to variant weather conditions over time for single- and two-vehicle crash severity potentials using a Markov switching multinomial logit model. In their study, daily averaged or maximal weather data over one week were used as follows: rain precipitation, temperature, snowfall, visibility, gust wind, and fog/frost. Weather variables such as rain precipitation, low visibility, gust wind were key factors generating time-related two-state nature of severities in single-vehicle accidents on high-speed roads, but not in two-vehicle accidents on high-speed roads. Savolainen and Mannering (2007) predicted motorcyclists' injury severities in single and multiple crashes using nested logit and multinomial logit models, respectively. Wet pavement was significant to increase no injury severity only in single-vehicle motorcycle crashes while none of the weather related factors were found to be significant to motorcyclists' injury severities in multi-vehicle crashes involving motorcyclists.

Rainfall-related effect on crash severity outcomes has been also identified along with roadway characteristics. Ordered probit models were used in a Abdel-Aty's study (2003) to predict driver injury severity in Central Florida, with crashes occurring in specific roadway sections, signalized intersections and toll plazas in expressway systems. It was found that crashes happening in signalized intersections with bad weather and dark street lighting had a significantly higher probability of severe injury. Abdel-Aty stated that an angle and turning collision in the adverse weather and dark street light conditions was a possible reason to contribute higher probability of injuries in signalized intersections. Donnell and Mason (2004) employed ordinal and nominal logistic regression models to predict interstate highway crash severity in cross-median and median-barrier collisions, respectively. Researchers found wet or icy pavement surface to be significant factor in decreasing crash severity. In contrast, in a study by Lee and Mannering (2002), the nested logit analysis showed that wet roadway surfaces increased the likelihood of evident and disabling injury/fatality in run-off-roadway accidents. The conflicting results from these two studies suggest the need for a detailed analysis for weather-related crash severity by collision type.

Special attention was given to crash severities by vehicle types in several studies. Shankar and Mannering (1996) used a multinomial logit model to predict five-level rider injury severities from statewide single-vehicle motorcycle crashes in Washington. In their study, pavement surface and weather interaction were significant factors to increase the likelihood of being injured. Interestingly, the crash severity on wet pavement without rainfall was limited only to property damage and possible injury crashes. Similarly, wet pavement for single-vehicle motorcycle crashes was more likely to result in no injury according to a study by Savolainen and Mannering (2007). The authors argued that it could be a result of lower speeds or longer headway maintained by riders in these conditions as they adjust for the perceived higher risk. Khorashadi et al. (2005) also utilized the multinomial logit model to analyze driver injury severities involving large-truck accidents. Results in their study showed that rainy weather was significant to increase injuries in urban area accidents. Ulfarsson and Mannering (2004) showed that wet roads led to increasing relatively higher injury severities involved in single sport utility vehicle (SUV)/minivan accidents using multinomial logit model. In a study by Kim et al. (2007), bicyclist injury severities in bicycle-motor vehicle accidents were predicted by the same method. Inclement weather including rain, snow, and fog was found to increase the probability of fatal injury approximately by 129%. The authors stated that the inclement weather effect was largely due to increased slipperiness which reduced both the vehicle's and bicycle's maneuverability and visibility. Pai and Saleh (2008) used ordered logit model to assess motorcyclist injury severities. In their study, on the contrary, fine weather was a statistically significant factor to increase the most severe category of motorcyclist injury.

Driver characteristics such as age or gender also play important roles in the likelihood of injury severity associated with weather conditions. Hill and Boyle (2006) utilized a logistic regression model in the fatality and incapacitating injury prediction. In their study, females in the older age groups (age of 54 or older) were more likely to suffer severe injuries in poor weather. Ulfarsson and Mannering (2004) found that rainy weather significantly affected the increase of property damage only level in female single SUV/minivan accidents but not in male driver single SUV/minivan accidents.

Weather impact was also evaluated in crash count-based models with the emphasis on severity counts. Abdel-Aty et al. (2006b) used a seemingly unrelated Negative Binomial regression model to estimate the number of property damage only and injury crashes, respectively. The result showed that crash severity in adverse weather conditions causing wet pavement surface was more likely to increase at curves or ramps. Caliendo et al. (2007) grouped crashes by total, fatal and injury crashes on curves and tangent roadways sections, and compare them using Poisson, Negative Binomial and Negative Multinomial regression models. In their study, rain was found to be a highly significant variable increasing the expected number of severe crashes for curves by a factor of 3.26 and for tangents by a factor of 2.81. Their study suggests wet-skidding for the higher number of severe crashes on curves.

Compared with previous studies, our study applied a rain-related crash dataset and included microscopic data at the crash moment to predict crash severity outcomes. To be specific, variables used in this study were real-time information at the crash moment, such as momentary weather and traffic data, and other non-weather data such as driver characteristics and roadway geometries. Additionally, rain-related single-vehicle crash severity models were compared to clear weather models to identify the common factors that contributed to crash severity regardless of weather. Research findings from this study will provide guidance on countermeasures to prevent severe crashes related to rainy weather and improve overall safety.

Fig. 1. Study area.

3. Data collection and processing

The study area consisted of 74.99 miles of southeastern Wisconsin highway segments including I-43, I-94, I-43/94 and I-43/894, where rainy weather crash frequency normalized by average annual daily traffic and vehicle miles traveled was higher than any other interstate highway segments between 2004 and 2006. The study area is shown in Fig. 1.

Single-vehicle crash and traffic data were collected from police accident reports (MV 4000) and traffic detector database (V-SPOC) from the WisTransPortal, respectively. The WisTransPortal project serves the data archiving and management needs of WisDOT.

In the study area, 255 single-vehicle crashes were found occurring in rainy weather between 2004 and 2006. The crash dataset included information for crash severity, roadway geometries, driver demographics, collision types, vehicle types, pavement conditions, and temporal and weather conditions. In this study, incapacitating injury (type A) and fatal injury (type K) crashes were combined as the highest level of crash severity to obtain a meaningful sample size (Agresti, 1996). Possible injury (type C) and non-incapacitating injury (type B) crashes were combined as the second highest level of crash severity because they were not clearly distinguishable. Property damage only (PDO) crashes made up the lowest level of crash severity. Crash frequency by the severity and the coding are provided in Table 1.

Next, average vehicle volume, speed, and occupancy data for 5-min intervals were obtained for 1 h prior to each crash moment. The associated standard deviations providing 1 h temporal buffers prior to the crash moment were also computed based on the archived traffic detector data due to the difference in density between crash

Table 1
Frequency distribution of crashes occurred in rainy weather.

Injury severity	Ordinal logistic	Sequential logistic			
		Forward format		Backward format	
		Stage 1	Stage 2	Stage 1	Stage 2
Fatal and incapacitating injury	10 (3)	10 (1)	10 (1)	10 (1)	–
Non-incapacitating and possible injury	57 (2)	57 (1)	57 (0)	57 (0)	57 (1)
PDO	188 (1)	188 (0)	–	188 (0)	188 (0)
Total	255	255	67	255	245

Note. SAS coding of crash severity level are in parentheses.

Please cite this article in press as: Jung, S., et al., Rainfall effect on single-vehicle crash severities using polychotomous response models. Accid. Anal. Prev. (2009), doi:10.1016/j.aap.2009.07.020

locations and detector locations in study area. The traffic detectors were installed at approximately 0.7 mile interval in the study area. Real-time average vehicle volume, speed and occupancy were collected at short-time (5-min) intervals.

State Trunk Network (STN) highway log from WisDOT contains roadway geometric attributes, including the number and width for travel lane and shoulder as well as pavement surface material. Using the STN highway log, the geometric attributes were linked to the crash dataset.

One of the keys to the success of this study was to obtain accurate weather data at a crash moment. However, there were few weather data sources that can provide minute-based measurement interval. After a thorough investigation, Weather Underground Inc. was found to deliver the most reliable and real-time weather data for Wisconsin. Considering crash distribution in study area and weather station proximity to each crash location, six airport weather stations and ten private weather stations were selected to obtain microscopic weather data measured by the minute based interval in this study. For each crash observation, three weather stations out of the 16 weather stations were used by their proximity to the crash location to estimate the microscopic weather data.

From the four databases, the following explanatory variables provided in Table 2 were utilized in the crash severity prediction models. The category coding in each explanatory variable was based on sample size and characteristic for each category, which is also shown in Table 2.

3.1. Weather parameter estimation

Weather data directly collected from a weather station include temperature, wind speed/direction, rainfall precipitation and rainfall duration. To reflect real-time weather conditions at the crash moment, some weather data can be estimated by interpolating between weather stations because weather data such as rainfall intensity or wind speed show geometrical and temporal variety. In other words, rainfall precipitation for 15 min (rainfall intensity) can be calculated using rainfall precipitation and the duration at each of the three weather stations near the crash location and the estimated 15-min rainfall intensities are interpolated for each crash. Water film depth, stopping sight distance (SSD), and deficiency of car-following distance (DCD) are estimated by hourly rainfall precipitation, traffic, and road geometry data.

3.1.1. Rainfall intensity

Rainfall intensity is defined as the rainfall precipitation divided by measurement interval. The rainfall intensity reflects visibility on highway in rainy weather conditions. Using three weather station data for each crash location, the average measurement interval of rainfall precipitation was 15 min. Therefore, rainfall precipitation for 15 min prior to a crash was adopted as the real-time rainfall intensity at the crash moment. Compared to the weather data measurement intervals mentioned in the previous studies, 15-min measurement interval used in this study was a more microscopic reflection of the real-time rainfall intensity at a crash moment.

Table 2
Explanatory variables used in prediction models.

Variable	Minimum	Maximum	Mean	Category coding
Driver's sex	–	–	–	Female = 1, male = 2
Alcohol or drug	–	–	–	Sobriety = 1, under alcohol/drug effect = 2
Safety belt	–	–	–	Use of safety belt = 1, non-used = 2
Driver action	–	–	–	Going straight = 1, others = 2, negotiating curve = 3
Curve direction	–	–	–	Curve to the right = 1, curve to the left = 2
Injury transport	–	–	–	Injured people transported to hospital = 1, others = 2
Terrain	–	–	–	Horizontal curve = 1, vertical curve = 2, horizontal/vertical curve = 3, tangent/flat = 4
First harmful spot	–	–	–	Ramp/gore = 1, Shoulder/outside shoulder = 2, median = 3, On roadway = 4
Pavement surface	–	–	–	Asphaltic cement plant mix/rigid base = 1, others = 2
Lighting condition	–	–	–	Daylight = 1, dusk/dawn/dark = 2, night but street light = 3
Collision type	–	–	–	Median related = 1, Non-collision = 2, Fixed object = 3
Vehicle type	–	–	–	Car = 1, truck (straight)/truck-tractor = 2, motor cycle = 3
Time of day	–	–	–	Peak-hour (6–8 a.m. and 3–5 p.m.) = 1, off-peak = 2
Day of week	–	–	–	Tuesday to Thursday = 1, Monday/Friday = 2, Saturday/Sunday = 3
Quarter of year	–	–	–	December to February = 1, March to May = 2, June to August = 3, September to November = 4
Wind direction	–	–	–	No wind = 1, North = 2, East = 3, South = 4, West = 5
Driver's age	16	80	34	–
Number of lanes	1	4	3	–
Lane width (ft)	12	17	12	–
Shoulder width (ft)	0/0[a]	16/20	7/11	–
Speed limit (min/h)	45	65	56	–
Average 5-min V[b]	2	129	51	–
Average 5-min SPD[c]	16	99	53	–
Average 5-min O[d] (%)	0.08	41.02	5.77	–
S.D.[e] of V	0.69	61.23	9.26	–
S.D. of SPD	0.27	29.33	6.25	–
S.D. of O	0.13	19.96	1.86	–
SSD[f] (ft)	71	1305	447	<330 = 1, [330,546] = 2, >546 = 3
DCD[g] (ft)	0	363	68	<16 = 1, [16,130] = 2, >130 = 3
Wind speed (km/h)	0	48	12.8	<6 = 1, [6,18] = 2, >18 = 3
Temperature (°C)	0.9	24.1	12.1	<7 = 1, [7,18] = 2, >18 = 3
Water film (mm/h)	0	0.74	0.15	<0.04 = 1, [0.04, 0.24] = 2, >0.24 = 3
RI[h] (mm/15 min)	0	9.53	0.79	<0.06 = 1, [0.06, 0.83] = 2, >0.83 = 3

[a] Left shoulder width/right shoulder width.
[b] V: volume.
[c] SPD: vehicle speed.
[d] O: occupancy.
[e] S.D.: standard deviation.
[f] SSD: stopping sight distance.
[g] DCD: deficiency of car-following distance.
[h] RI: rainfall intensity.

3.1.2. Water film depth

Water film is created by rainfall between the tire and pavement surface, causing a decrease in skid resistance. Therefore, water film depth in this study was used as an explanatory variable for measuring the slippery pavement condition. Russam and Ross (1968) gave the following empirical method to estimate the water film depth:

$$D = \frac{0.046(W \cdot S/Sc \cdot I)^{1/2}}{S^{1/5}} \tag{1}$$

$$S = (S_l^2 + S_c^2)^{1/2} \tag{2}$$

where D = water film depth (mm/h); I = rainfall intensity (mm/h); S = flow path slope (%); S_l = longitudinal slope (%); S_c = slope of pavement cross section (%); W = width of pavement (m).

3.1.3. SSD and DCD

In this study, there were not direct visibility data for highways. Therefore, SSD and DCD were considered as the surrogate measures for highway visibility at the time of the crash. First of all, SSD formula is as follows.

$$\mathrm{SSD} = \frac{1.47V \cdot t + 1.075V^2}{a} \tag{3}$$

where V = vehicle speed (min/h); t = brake reaction time (2.5 s); a = deceleration rate (ft/s^2).

According to a detailed study about pavement conditions (Kokkalis and Panagouli, 1998), the coefficient of wet pavement friction is associated with water film depth and vehicle speed shown in Fig. 2. In Fig. 2, vertical axis is Skid Number (or Friction Number) representing longitudinal friction force and horizontal axis is vehicle slip speed given a specific water film depth. That is, Fig. 2 combines relations among friction force, vehicle speed and water film depth.

Using Fig. 2 and pavement surface material information from Wisconsin STN highway log, deceleration rate in SSD formula can be obtained by multiplying wet pavement friction coefficient by gravity acceleration. Consequently, SSD was calculated by the obtained deceleration rate, vehicle speed from traffic detector data, and brake reaction time.

DCD represents the risk of losing control caused by driver over correction for avoiding any potential conflict. DCD is calculated by the following formula.

$$\mathrm{DCD} = \mathrm{SSD} - \mathrm{AVG} \tag{4}$$

where SSD = stopping sight distance; AVG = average vehicle gap.

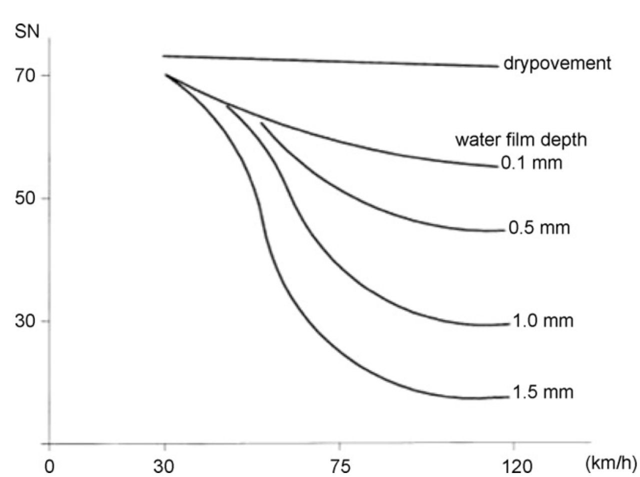

Fig. 2. Influences of water film depth and vehicle speed on skid resistance.

In Eq. (4), AVG is the average vehicle spacing obtained by subtracting average vehicle length from vehicle density data calculated by traffic detector data (Roess et al., 2004).

Strictly speaking, vehicle speed in SSD formula should be individual vehicular speeds, so is the gap between every pair of cars. In this study, the average of 5-min traffic detector data containing the crash occurrence time was used to surrogate the real-time prevailing traffic conditions at the crash moment.

3.2. Weather data interpolation

To estimate weather data at a crash location, a study (Patrick and Stephenson, 1990) regarding the comparison of interpolation methods concludes that the inverse squared distance method is stable and appropriate for the localized field with short spatial correlation length scale and large variability, and the minimum number of weather stations to apply the inverse squared distance interpolation is three (Press et al., 2007). Therefore, the inverse squared distance interpolation was utilized to estimate localized weather data at the crash moment:

$$Z_k = \Sigma w_i Z_i \tag{5}$$

$$w_i = \frac{\left(1/dik\right)^2}{\Sigma(1/d_{ik})^2} \tag{6}$$

where Z_k = weather data estimated at a crash location, k; Z_i = weather data measured by each weather station, I; w_i = weighting for each weather station, I; d_{ik} = distance between each weather station i and a crash location k; i = one of the three nearest weather stations, i = 1–3.

In Eqs. (5) and (6), the larger weight is applied to a closer weather station data. According to the weather database used in this study, rainfall precipitation and wind speed from weather stations showed spatial and temporal variety in the values. Hence, 15-min rainfall intensity, hourly rainfall precipitation for water film depth, and wind speed were interpolated between three weather stations near to each crash location by the inverse squared distance interpolation method. However, temperature data in the weather station nearest to each crash location was used without data interpolation because temperature was spatially and temporally consistent at the crash time.

4. Methodology

To model discrete outcome data, several alternative modeling techniques such as ordered probability, multinomial and nested logit models can be considered but the application to the dataset varies from one to another due to their limitations. For example, the multinomial and nested logit models do not account for the ordering of crash severities, which was stated in several previous studies (Abdel-Aty, 2003; Milton et al., 2008; Wang and Abdel-Aty, 2008). Note that the crash severities are not only multiple discrete outcomes but also inherently ordered. Therefore, neither multinomial nor nested logit analysis would be able to account for the ordinal nature of the crash severities. Moreover, traditional ordered probability approaches impose a critical restriction that regression parameters have to be the same for different crash severity levels, so called proportional odds. Since it is not clear whether the distance between adjacent crash severity levels is equal, it is rather arbitrary to assume that all coefficients of ordered probability models are the same. This restriction was also mentioned in other studies (Milton et al., 2008; Wang and Abdel-Aty, 2008).

Alternatively, a generalized version of the standard ordered logit model was introduced to relax the restrictions of the same parameters for explanatory variables across the response levels imposed by the standard ordered logit model (Eluru et al., 2008). The generalized standard ordered logit model allows the treatment of the utility thresholds across the ordered response levels by separate parameter coefficients for explanatory variables and heterogeneity in the effects of injury severity determinants. However, the generalized ordered response model is recommended only to conclude that the proportional odds assumption is valid because the model is very anti-conservative (Peterson and Harrell, 1990). Based on the purpose of this study, the difference in the set of predictors across various severity levels is one of the most important issues. Even though the generalized ordered logit model allows a separate parameter for each explanatory variable across crash severity levels, the set of significant explanatory variables is invariant over all the crash severity comparisons.

Furthermore, crash data used in this study were filtered through several criteria: rainy weather, wet pavement, single vehicle included in a crash, interstate highway divided by barrier, no construction zone, no hit and run and no pedestrian involved in a crash. Additionally, microscopic weather and traffic attributes at the crash time as well as roadway geometries, vehicle, and driver attributes were used in this study. Therefore, the conditions restricted by several criteria at the crash moment and detailed explanatory variables can reduce moderate influence of unobserved factors that may cause heterogeneity in the effects of crash severity determinants. Specifically, eight crashes out of 255 total crashes were related to intoxicated drivers and seven crashes out of the eight crashes by the intoxicated drivers were property damage only levels in this study, which implies a certain level of homogeneity in drivers' alertness.

Consequently, given the certain level of data homogeneity, sequential logistic regression approach was selected to predict rainy weather crash severities in this study because this method not only accounts for the inherent ordering of crash severities but also allows different regression parameters for the severity levels. Nevertheless, as a comparison, ordinal logistic regression was used to predict rainy weather crash severities with the inclusion of only the independent variables satisfying the proportional odds assumption. Since sequential logistic regression model does not require the proportional odds, all the independent variables can be included in the model.

4.1. Model structure

The functional form of a logistic is S-shaped, making it possible to deal with the multi-level response outcomes and the probability of a certain outcome (Keinbaum and Klein, 2002).

An interpretation of the logistic regression model is based on the odds and the odds ratio of an event. The odds of an event are expressed as a ratio of the probability that the event will occur divided by the probability that it will not. The odds ratio is a ratio of the predicted odds for a one-unit change in X_i with other variables in the model held constant. Logarithm of the odds ratio is called as logit and it is shown as follows.

$$\ln\left[\frac{P(X)}{1-P(X)}\right] = \operatorname{logit} P(X) = \alpha + \beta X \tag{7}$$

where $P(X) = P(Y = y|X)$ = probability of a response outcome, Y = response variable, $y = 0$ or 1, X = set of variables, α = intercept parameter, β = set of parameter estimates for X.

Crash severity can be fitted to a proportional odds model meaning that the odds ratio assessing the effect of a predictor for any ordinal response categories is equal regardless of where cut-point to classify the response categories is made. The logistic model assuming the proportional odds is called an ordinal logistic regression model. In this study, the ordinal logistic regression fits the following equations:

$$\operatorname{logit} P_1 = \ln\left[\frac{P_1}{1-P_1}\right] = \alpha_1 + \sum \beta_i X_i \tag{8}$$

$$\operatorname{logit}(P_1 + P_2) = \ln\left[\frac{P_1 + P_2}{1 - P_1 - P_2}\right] = \alpha_2 + \sum \beta_i X_i \tag{9}$$

where P_1 = probability of PDO severity; P_2 = probability of non-incapacitating/possible injury severity; P_3 = probability of fatality/incapacitating injury severity.

The LOGISTIC procedure was used to fit logistic regression models. To reverse the default ordering of the response variable in a statistical program used for this study, accordingly, the ordinal logistic regression can be refitted in the following way:

$$\operatorname{logit} P_3 = \ln\left[\frac{P_3}{1-P_3}\right] = \ln\left[\frac{P_3}{P_1+P_2}\right] = \alpha_1 + \sum \beta_i X_i = h1 \tag{10}$$

$$\operatorname{logit}(P_2 + P_3) = \ln\left[\frac{P_2 + P_3}{1 - P_2 - P_3}\right] = \ln\left[\frac{P_2+P_3}{P_1}\right] = \alpha_2 + \sum \beta_i X_i = h2 \tag{11}$$

Combining Eqs. (10) and (11), the probability of each crash severity is written as follows:

$$P_3 = \frac{\exp(h1)}{[1+\exp(h1)]} \tag{12}$$

$$P_2 = \frac{[\exp(h1) - \exp(h2)]}{[1+\exp(h1)][1+\exp(h2)]} \tag{13}$$

$$P_1 = 1 - P_2 - P_3 \tag{14}$$

If the proportional odds assumption is not satisfied in the crash dataset, the probability of a crash severity can be estimated by the sequential logistic regression model. The primary difference between ordinal logistic regression and sequential logistic regression is that the sequential logistic regression handles a set of predictors at each stage independent of the set used at the previous stage.

In this study, the standard logistic regression concept is applied at both stages to fit the sequential logistic regression model. At the second stage, a sub-sample is used after removing observations of a certain crash severity used in the previous stages (Maddala, 1983). In order to explore whether there is an impact in the development of the sequential structure, forward and backward formats were conducted in this study as follows:

Forward format:

Stage 1: Crash types K, A, B, and C vs. PDO.
Stage 2: Crash types K and A vs. Crash types B and C.

Backward format:

Stage 1: Crash types K and A vs. Crash types B, C, and PDO.
Stage 2: Crash types B and C vs. PDO.

Using the standard logistic regression concept at each stage of both formats, the probabilities of crash severity levels can be written as follows:

Forward format:

$$\text{Stage 1}: \quad \frac{(P_2+P_3)}{P_1} = \exp\left(\alpha_1 + \sum \beta_i X_i\right) = g1 \tag{15}$$

$$\text{Stage 2}: \quad \frac{P_3}{P_2} = \exp\left(\alpha_2 + \sum \beta_j X_j\right) = g2 \tag{16}$$

$$P_1 = \frac{1}{1+g1} \tag{17}$$

$$P_2 = \frac{g1}{(1+g1)(1+g2)} \quad (18)$$

$$P_3 = \frac{g1 \cdot g2}{(1+g1)(1+g2)} \quad (19)$$

Backward format:

$$\text{Stage 1}: \quad \frac{P_3}{(P_1 + P_2)} = \exp\left(\alpha_1 + \sum \beta_i X_i\right) = g1 \quad (20)$$

$$\text{Stage 2}: \quad \frac{P_2}{P_1} = \exp\left(\alpha_2 + \sum \beta_j X_j\right) = g2 \quad (21)$$

$$P_1 = \frac{1}{(1+g1)(1+g2)} \quad (22)$$

$$P_2 = \frac{g2}{(1+g1)(1+g2)} \quad (23)$$

$$P_3 = \frac{g1}{1+g1} \quad (24)$$

4.2. Measures of model performance

Typical measures of model performance for goodness of fit and prediction accuracy are likelihood ratio test and classification table, respectively. These measures of model performance are synthetically considered to assess crash severity prediction models in rainy weather condition.

The likelihood ratio (LR) test reveals whether or not global null hypothesis for a specific model is rejected. In other words, an estimated model containing at least one non-zero parameter coefficient is better fit than constant only model when *P*-value of LR test is less than a conventional criterion.

Standard logistic regression model classifies an observation as an event if the estimated probability of this observation is greater than or equal to a given cut-point. Otherwise, it is classified as a non-event. In the statistical term, the rate of actual events that are also predicted to be events is called sensitivity. Similarly, the rate of actual non-events that are also predicted to be non-events is called specificity. The overall predictive power of a model depends on the proportion of correctly predicted observations (i.e., the sum of sensitivity and specificity). The classification table for the prediction accuracy is as follows.

	Predicted event	Predicted non-event
Actual event	Sensitivity	False negative
Actual non-event	False positive	Specificity

Even though the predictive power of a model can be measured for all severity levels, sensitivity and false negative rate are often emphasized for the highest crash severities (fatal and incapacitating injuries) because of their resulting economic loss. Hence, a model that produces high sensitivity and low false negative rate at the classification stage for fatal and incapacitating injuries is considered as a good one.

Note that the classification accuracy is dependent on the probability cut-point since the model classifies an observation based on the given probability cut-point. The probability cut-point of 0.5 can be considered first assuming an equal opportunity for observing event or non-event. However, the value of 0.5 may yield low model accuracy. In reality, more severe crashes (event) do not frequently occur compared with less severe crashes (non-event), which implies the probability cut-point of 0.5 is too high to be used in the prediction model. From this perspective, the probability cut-point may be determined by practical consideration. Since desirable prediction models should fit the field data well, the probability cut-point used in this study is the proportion of actual events in the field data.

5. Results and discussion

In this study, PROC LOGISTIC statement in SAS 9.1 was used to estimate ordinal logistic regression and sequential logistic regression models on the basis of rain-related single-vehicle crashes with a significance level of 0.10 for retaining explanatory variables in the models. The modeling process is as follows.

First of all, bivariate logistic regression of each explanatory variable was performed to choose an individual predictor one by one which correlated to crash severity. Especially for ordinal logistic regression model, the individual predictor that was also satisfied with proportional odds assumption was selected in the step of bivariate logistic regression. In this study, six continuous weather-related variables were considered: 15-min rainfall intensity, wind speed, temperature, water film depth, SSD, and DCD. Since weather effect on crash severity is the primary interest of this study, the continuous weather data were specifically transformed to a categorical variable by quantile if any continuous weather variable was not selected by the bivariate logistic regression.

Next, correlation between predictors selected by the bivariate logistic regression was identified by Pearson's correlation coefficient or likelihood ratio chi-squared test in order not to omit significant predictors in multiple logistic regression models. After the correlation test, several combinations that contain the maximum number of uncorrelated predictors were constructed for the next step.

Finally, stepwise variable selection was conducted to select the best multiple logistic regression model. Using the best-fitted model, a classification table by the model was produced at a given probability cut-point to check prediction power of the fitted model. Given an appropriate probability cut-point, the fitted model produces reasonable prediction accuracies for both actual events and actual non-events. In this study, the probability cut-point is determined by the proportion of actual events to total observations at each modeling stage because desirable prediction models should fit the field data well and the probability cut-point should reflect the field data condition.

5.1. Ordinal logistic regression model

The proportional odds assumption in the best model provided in Table 3 was not rejected, which means that ordinal logistic regression is appropriate to the response outcomes. *P*-value of LR test revealed quite small value in Table 3, indicating the global null hypothesis is rejected.

The odds of response categories for types K and A vs. types B, C and PDO or for injury vs. PDO were 0.041 times higher for wearing safety belt than for not wearing safety belt, implying crash severity reduction by the safety belt. The odds ratio for posted speed limit was positive, but close to 1, indicating slight effect on increasing crash severity at higher posted speed limit. All the significant parameter effects were consistent to general expectation.

The probability cut-point to classify the predicted event was based on the proportion of actual event observations to total observations. That is, the probability cut-point was 0.26 to predict injury crash severity (event) while the probability cut-point was 0.04 to predict the fatal and incapacitating injury crash severity (event). For two event comparisons, reasonable overall prediction accuracies for total observations were produced. However, sensitivity was 50% particularly for fatal and incapacitating injury crashes vs. others at the reasonable overall prediction accuracy of 65%, which means that the prediction power of ordinal logistic regression was not sufficiently strong for the highest level of crash severity.

Table 3
Multiple ordinal logistic regression model.

	ChiSq	D.F.	Pr > ChiSq
Score test	4.5793	2	0.1130
LR test	48.3070	2	<0.0001

	Parameter	Estimate	S.E.	Odds ratio	Pr > ChiSq
MLE	Intercept 3	−4.2021	1.4022	–	0.0027
	Intercept 2	−1.4351	1.0581	–	0.3037
	Safety belt	−3.1990	0.5345	0.041	<0.0001
	Speed limit	0.0580	0.0233	1.060	0.0130

Classification	Event		Non-event		Overall accuracy
Others vs. PDO (P_{cutoff} = 0.26)	Correct = 41	Incorrect = 26	Correct = 135	Incorrect = 53	69%
	Sensitivity	Specificity	False POS	False NEG	
	61%	72%	56%	16%	
K + A vs. Others (P_{cutoff} = 0.04)	Correct = 5	Incorrect = 5	Correct = 161	Incorrect = 84	65%
	Sensitivity	Specificity	False POS	False NEG	
	50%	66%	94%	3%	

Note. MLE: maximum likelihood estimate; false POS: false positive rate; false NEG: false negative rate.

5.2. *Sequential logistic regression model*

The modeling process of sequential logistic regression model was identical to that of ordinal logistic regression. The differences between the sequential logistic regression and the ordinal logistic regression were the use of sub-sample and variant sets of predictors across stages for the sequential logistic regression. In other words, PDO crashes were removed at the second stage in the forward format while fatal and injury crashes were removed in the backward format.

5.2.1. *Forward format*

Based on low *P*-values of LR test in Table 4, the global null hypothesis was rejected at both stages, indicating the estimated models were better to predict crash severities than constant only model.

At the first stage, female drivers were much more likely to be involved in more severe crashes than PDO crashes based on the high odds ratio. An interesting effect of left shoulder width was identified. The odds ratio for left shoulder width was approximately 1.3, implying wide left lane was more likely to increase crash sever-

Table 4
Multiple model for forward format of sequential logistic regression.

	Stage 1		
	ChiSq	D.F.	Pr > ChiSq
LR test	22.4048	3	<0.0001

	Parameter	Estimate	S.E.	Odds ratio	Pr > ChiSq
MLE	Intercept 1	−3.5445	0.7718	–	<0.0001
	FD	2.6658	0.9072	14.3794	0.0016
	LSW	0.2436	0.0769	1.2758	0.0015
	FD × LSW	−0.2247	0.0950	0.7987	0.0180

	Event		Non-event		Overall accuracy
Classification (P_{cutoff} = 0.26)	Correct = 48	Incorrect = 19	Correct = 105	Incorrect = 83	60%
	Sensitivity	Specificity	False POS	False NEG	
	72%	56%	63%	15%	

	Stage 2		
	ChiSq	D.F.	Pr > ChiSq
LR test	12.8808	3	0.0049

	Parameter	Estimate	S.E.	Odds ratio	Pr > ChiSq
MLE	Intercept	−2.3430	1.5934	–	0.1414
	SSD	0.0050	0.0027	1.005	0.0896
	Wind speed 2	−1.4784	0.8929	0.228	0.0978
	Vehicle type 1	−1.5171	0.8156	1.085	0.0629

	Event		Non-event		Overall accuracy
Classification (P_{cutoff} = 0.17)	Correct = 8	Incorrect = 2	Correct = 44	Incorrect = 13	78%
	Sensitivity	Specificity	False POS	False NEG	
	80%	78%	62%	4%	

Note. FD: female driver; LSW: left shoulder width; FD × LSW: interaction of female driver with left shoulder width; SSD: stopping sight distance; Wind speed 2: wind speed from 6 km/h to 18 km/h; Vehicle type 1: passenger car.

Please cite this article in press as: Jung, S., et al., Rainfall effect on single-vehicle crash severities using polychotomous response models. Accid. Anal. Prev. (2009), doi:10.1016/j.aap.2009.07.020

Table 5
Multiple model for backward format of sequential logistic regression.

		Stage 1					
		ChiSq		D.F.	Pr > ChiSq		
LR test		34.5496		4	<0.0001		
	Parameter		Estimate		S.E.	Odds ratio	Pr > ChiSq
MLE	Intercept		−0.4724		0.4839	–	0.4866
	RI		0.5806		0.1655	1.787	0.0005
	Wind speed 2		−2.5170		1.0440	0.081	0.0159
	Terrain 3		2.7768		1.1082	16.067	0.0094
	Safety belt		−4.2795		1.0090	0.014	<0.0001
		Event			Non-event		Overall accuracy
Classification (P_{cutoff} = 0.04)		Correct = 9	Incorrect = 1		Correct = 214	Incorrect = 31	88%
		Sensitivity	Specificity		False POS	False NEG	
		90%	87%		78%	0.5%	
		Stage 2					
		ChiSq		D.F.	Pr > ChiSq		
LR test		41.8949		2	<0.0001		
	Parameter		Estimate		S.E.	Odds ratio	Pr > ChiSq
MLE	Intercept		1.6464		0.7680	–	0.0321
	Female driver		1.1858		0.3419	3.273	0.0005
	Safety belt		−3.6222		0.7986	0.027	<0.0001
		Event			Non-event		Overall accuracy
Classification (P_{cutoff} = 0.23)		Correct = 39	Incorrect = 18		Correct = 128	Incorrect = 60	68%
		Sensitivity	Specificity		False POS	False NEG	
		68%	68%		61%	12%	

Note. RI: 15-min rainfall intensity; Wind speed 2; wind speed from 6 km/h to 18 km/h; Terrain 3: horizontal/vertical curve.

ity from PDOs to injuries, which is opposite to general expectation. However, the odds ratio for the interaction with female drivers was less than 1, which implies that wide left shoulder can be helpful to decrease injury crashes especially for female drivers.

Weather factors were more explicitly identified at stage two of forward sequential logistic regression than ordinal logistic regression. At the second stage, severe crashes including types K and A increased slightly as stopping sight distance increased. On the other hand, the second class of wind speed from 6 km/h to 18 km/h was likely to decrease the most severe crashes. The effect of wind speed could be caused by drivers' cautions due to low visibility derived by moderately strong wind in the rainfall. Passenger cars were also less likely to be involved in the most serious crashes than other types of vehicles.

The probability cut-point was 0.26 to predict injury crash severity (event) at the first stage while the probability cut-point was 0.17 to predict the fatal and incapacitating injury crash severity (event). Overall prediction accuracies were reasonable at both stages. Especially, sensitivity at the second stage was 80%, which is much higher than the sensitivity for the fatal and incapacitating injury crash severity in ordinal logistic regression.

5.2.2. *Backward format*

Table 5 shows the best model selected by backward sequential logistic regression. Small *P*-values in LR test at both stages reveals that the selected model provided in Table 5 is better fit than the global null model.

Weather-related factors were identified to predict fatal and incapacitating injury crash severity at stage one. The most severe crashes were 1.787 times more likely to occur as rainfall intensity for 15 min was getting stronger. This result implies that drivers tend to less sensitively perceive the risk of driving by the rainfall intensity. However, wind speed effect on the highest crash severity in backward format was consistent to the effect in forward format. That is, the odds ratio for the second class of wind speed from 6 km/h to 18 km/h was less than 1, indicating moderately strong wind speed is likely to decrease the fatal and incapacitating injury crashes. Also note that the odds ratio for horizontal/vertical curves was extremely higher than 1, which indicates that the horizontal/vertical curves increases the likelihood of the most severe crashes in rainy weather. At stage two, female driver was more likely to be involved in types B and C crashes in rainy weather than male driver based on the positive parameter estimate and the corresponding odds ratio. At both stages, wearing safety belt decreased more severe crashes.

At stage one, 0.04 was used as the probability cut-point to predict fatal and incapacitating injury crash severity (event) while 0.23 was used to predict non-incapacitating and possible injury crash severity (event) at stage two. According to Table 5, overall prediction accuracies were 88% at the first stage and 68% at the second stage, which are reasonable. In particular, the first stage shows the highest sensitivity (90%) and overall prediction accuracy of all the rates provided by ordinal and sequential logistic regression models. Additionally, false negative rate at stage one was found to be the lowest of all false negative rates from ordinal and sequential logistic regression models. These prediction accuracy rates at the first stage imply that backward format is the most desirable to predict crash severity levels especially for fatal and incapacitating injury crash severity.

Comparing multiple regression results based on goodness of fit, parameter significance and prediction accuracies, on the whole, the backward sequential logistic regression model outperforms ordinal and forward sequential logistic regression models in predicting crash severity levels in rainy weather. Specifically, the backward

Table 6
Crash frequencies and coding in study area in clear weather.

Crash severity	Frequency	Category coding	
		Stage 1	Stage 2
Fatal and incapacitating injury	30	1	–
Non-incapacitating and possible injury	175	0	1
PDO	350	0	0
Total	555	555	525

sequential logistic regression model was found to be comparatively effective to predict the highest level of crash severity based on the overall prediction accuracy, sensitivity and false negative rate. Moreover, 15-min rainfall intensity impact on crash severity was significantly identified in the backward format and wind speed was more significant in the backward format than in the forward format. In addition, the estimated parameter coefficients in the backward sequential model consistently agree with the common expectations. Therefore, the backward format of sequential logistic regression model is recommended as the final model for predicting all levels of single vehicle crashes that occurred on high-speed highways in rainy weather.

5.3. Comparison of clear weather crash severity

The goal of comparing rainy and clear weather conditions using the best crash severity prediction model is to identify factors that are independent of weather conditions. In this study, the clear weather condition indicates dry pavement under clear or cloudy weather, when rainfall and other weather factors do not exist. Crash data in clear weather condition only were separately collected. In other words, crash data used in rainy weather model were not included in modeling clear weather crash severity. Study area, data sources, data collection duration, variable coding, and modeling process were the same as those of the backward sequential logistic regression model that was selected as the best model format to account for rainy weather effect on crash severity levels. Crash frequency for the clear weather is presented in Table 6.

The best multiple regression model for clear weather condition is provided in Table 7. At the first stage, the odds ratio for posted speed limit was greater than 1, indicating that higher posted speed limit increased the likelihood of fatal and incapacitating injury crashes. In particular, the odds ratio for street lighting at night was approximately 10, implying street light condition at night has high potential to increase fatal and incapacitating injury crashes in clear weather. At night, glaring caused by street lighting might affect the strong street light effect on increasing the most severe crashes. However, interaction of the street lighting at night with safety belt was significantly identified to decrease the most severe crashes, which implies that safety belt can be effective to improve highway safety under the street lighting at night. The effect of wearing safety belt on decreasing crash severity was consistent regardless of weather conditions.

At the second stage, the odds ratio of possible/non-incapacitating injury crashes to being PDO crashes was 0.567 for spring season. Intuitively, this seasonal factor is expected to be an excellent indicator representing weather conditions. It is plausible that crash severity is affected by weather conditions. Additionally, median related collisions increased the likelihood of possible/non-incapacitating injury crashes.

The overall prediction accuracies were reasonable at both stages: 64% with event classification probability cut-point of 0.05 at stage one and 60% with event classification probability cut-point of

Table 7
Multiple model for backward sequential logistic regression (clear weather).

	Stage 1				
	ChiSq	D.F.	Pr > ChiSq		
LR test	38.7058	4	<0.0001		
	Parameter	Estimate	S.E.	Odds ratio	Pr > ChiSq
MLE	Intercept	−10.1080	2.6479	–	0.0001
	Speed limit	0.1316	0.0421	1.141	0.0017
	LGT 3	2.3159	0.7182	10.134	0.0013
	Safety belt	−1.1575	0.5665	0.314	0.0410
	SFTB × LGT 3	−1.7601	0.8779	0.172	0.0450
	Event		Non-event		Overall accuracy
Classification ($P_{cutoff} = 0.05$)	Correct = 23	Incorrect = 7	Correct = 332	Incorrect = 193	64%
	Sensitivity	Specificity	False POS	False NEG	
	77%	63%	89%	2%	
	Stage 2				
	ChiSq	D.F.	Pr > ChiSq		
LR test	15.2776	2	0.0005		
	Parameter	Estimate	S.E.	Odds ratio	Pr > ChiSq
MLE	Intercept	−0.9102	0.1496	–	<0.0001
	Spring season	−0.5682	0.2460	0.567	0.0209
	Collision type 1	0.5936	0.1902	1.811	0.0018
	Event		Non-event		Overall accuracy
Classification ($P_{cutoff} = 0.33$)	Correct = 89	Incorrect = 86	Correct = 222	Incorrect = 128	60%
	Sensitivity	Specificity	False POS	False NEG	
	51%	63%	59%	28%	

Note. LGT 3: night but street light; SFTB: safety belt; Collision type 1: median related collisions.

0.33 at stage two. In addition to the overall accuracies, sensitivities at both stages were also reasonable, implying backward sequential model are effective to predict more severe crashes (event) even in clear weather.

6. Conclusions

Even though rainfall effect on crash severity has been investigated in previous studies, the rainy weather-related factors in the studies lack the accuracy and sophistication to reflect real-time pavement surface conditions and visibility practically during the rainfall. For instance, wet or dry pavement surface, average annual rainfall precipitation, and even hourly rainfall are not sufficient to capture the real-time rainy weather conditions prior to or during the crash occurrence. Using more microscopic weather data, this study assessed rainfall effect on the severities of single vehicle crashes on the selected Wisconsin interstate highways. To comprehensively characterize weather conditions and their effects on crash occurrences, this study used several novel variables at the crash moment, in particular, 15-min rainfall intensity, water film depth, stopping sight distance, and deficiency of car-following distance that have not been frequently considered in the previous studies. In addition, estimated or measured weather factors were interpolated between three weather stations by inverse squared interpolation method for each crash location.

In this study, both ordinal and sequential logistic regression models were applied to predict crash severity that is a polychotomous response. The sequential logistic regression models were further divided into the forward format from the lowest injury severity to the highest one and the backward format reversing the sequence. Additionally, the severity of crashes occurring on dry pavement in cloudy or clear weather was estimated to compare with its counterpart in rainy condition. As a result, the backward format of sequential logistic regression model outperformed others in predicting crash severity levels, especially fatal and incapacitating injuries, and detected rainfall effect on the severities. In the backward sequential logistic regression model, following variables were significantly identified: 15-min rainfall intensity, wind speed, horizontal/vertical curve, female driver, and safety belt usage. The backward format produced the highest prediction accuracy and clearly significant weather effect especially on fatal and incapacitating injury prediction.

Note that ordinal logistic regression imposes restrictions that all parameter estimates have to be the same and only the intercept is allowed to be different while sequential logistic regressions allows variant parameter estimates for different response outcomes. In other words, the sequential logistic regression is more flexible to reflect variant predictor effects on response categories and to predict the response accurately for this study. Thus, the backward sequential logistic regression model, specifically, is considered to be the most appropriate for determining the probability of crash severity in rainy weather. In addition, a seasonal indicator was significant in the clear weather backward sequential logistic regression model. This implies that weather condition may affect crash severity outcomes because the weather condition is associated with the seasonal factor. The resultant findings in this study can be used to provide quantitative support on improving road weather safety via weather warning systems, highway facility improvements, and speed limit management.

Acknowledgement

The authors thank Andrea Bill at the University of Wisconsin-Madison for her contribution to this paper.

References

Abdel-Aty, M., 2003. Analysis of driver injury severity levels at multiple locations using ordered Probit models. Journal of Safety Research 34 (5), 597–603.
Abdel-Aty, M., Pemmanaboina, R., 2006a. Calibrating a real-time traffic crash prediction model using archived weather and ITS traffic data. IEEE Transactions on Intelligent Transportation Systems 7 (2), 167–174.
Abdel-Aty, M., Pemmanaboina, R., Hsia, L., 2006b. Assessing crash occurrence on urban freeways by applying a system of interrelated equations. Transportation Research Record 1953, 1–9.
Agresti, A., 1996. An Introduction to Categorical Data Analysis. John Wiley & Sons, Inc., New York.
Caliendo, C., Guida, M., Parisi, A., 2007. A crash-prediction model for multilane roads. Accident Analysis and Prevention 39 (4), 657–670.
Dissanayake, S., Lu, J., 2002. Analysis of severity of young driver crashes: sequential binary logistic regression modeling. Transportation Research Record 1784, 108–114.
Donnell, E.T., Mason Jr., J.M., 2004. Predicting the severity of median-related crashes in Pennsylvania by using logistic regression. Transportation Research Record 1897, 55–63.
Eluru, N., Bhat, C.R., Hensher, D.A., 2008. A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. Accident Analysis and Prevention 40 (3), 1033–1054.
Friction coefficient: http://www.engineershandbook.com/Tables/frictioncoefficients.htm.
Golob, T.F., Recker, W.W., 2003. Relationships among urban freeway accidents, traffic flow, weather, and lighting conditions. Journal of Transportation Engineering 129 (4), 342–353.
Hill, J.D., Boyle, L.N., 2006. Assessing the relative risk of severe injury in automotive crashes for older female occupants. Accident Analysis and Prevention 38 (1), 48–154.
Keinbaum, D.G., Klein, M., 2002. Logistic Regression: A Self-Learning Text, 2nd edition. Springer, New York.
Khan, G., Qin, X., Noyce, D.A., 2008. Spatial analysis of weather crash patterns, ASCE. Journal of Transportation Engineering 134 (5), 191–202.
Kim, J., Kim, S., Ulfarssonb, G.F., Porrello, L.A., 2007. Bicyclist injury severities in bicycle–motor vehicle accidents. Accident Analysis and Prevention 39 (2), 238–251.
Khorashadi, A., Niemerier, D., Shankar, V., Mannering, F.L., 2005. Differences in rural and urban driver-injury severities in accidents involving large-trucks: an exploratory analysis. Accident Analysis and Prevention 37 (5), 910–921.
Kokkalis, A.G., Panagouli, O.K., 1998. Factual evaluation of pavement skid resistance variation: surface wetting from chaos. Chaos, Solitons and Fractals 9 (11), 1875–1890.
Lee, J., Mannering, F.L., 2002. Impact of roadside features on the frequency and severity of run-off-roadside accidents: an empirical analysis. Accident Analysis and Prevention 34 (2), 149–161.
Maddala, G.S., 1983. Limited-Dependent and Qualitative Variables in Econometrics. Cambridge University Press, New York.
Malyshkina, N.V., Mannering, F.L., 2009. Markov switching multinomial logit model: an application to accident-injury severities. Accident Analysis and Prevention 41 (4), 829–838.
Milton, J.C., Shankar, V.N., Mannering, F.L., 2008. Highway accident severities and the mixed logit model: an exploratory empirical analysis. Accident Analysis and Prevention 40 (1), 260–266.
Pai, C., Saleh, W., 2008. Modeling motorcyclist injury severity by various crash types at T-junctions in the UK. Safety Sciences 46, 1234–1247.
Patrick, N.A., Stephenson, D., 1990. Spatial variation of rainfall intensities for short duration storms. Hydrological Science 35, 667–680.
Peterson, B., Harrell Jr., F.E., 1990. Partial proportional odds model for ordinal response variables. Applied Statistics 39 (2), 205–217.
Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2007. Numerical Recipes: The Art of Scientific Computing, 3rd edition. Cambridge University Press, New York.
Qin, X., Noyce, D.A., Lee, C.Y., 2006. Snowstorm event-based crash analysis. Transportation Research Record 1948, 135–141.
Roess, R.P., Prassas, E.S., Mcshane, W.R., 2004. Traffic Engineering, 3rd edition. Pearson Education, Inc., New Jersey.
Russam, K., Ross, N.F., 1968, The Depth of Rain Water on Road Surfaces, Ministry of Transport Report No. LR 236. Road Research Laboratory.
SAS Institute Inc., 1998. Logistic Regression Examples Using the SAS System. Publications Division of SAS Institute Inc., SAS Campus Drive, Cary, NC 27513.
Savolainen, P.T., Tarko, A.P., 2005. Safety impacts at intersections on curved segments. Transportation Research Record 1908, 130–140.
Savolainen, P., Mannering, F.L., 2007. Probabilistic models of motorcyclists' injury severities in single- and multi-vehicle crashes. Accident Analysis and Prevention 39 (5), 955–963.
Shankar, V., Mannering, F.L., 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., Mannering, F.L., 1996. An explanatory multinomial logit analysis of single-vehicle motorcycle accident severity. Journal of Safety Research 27 (3), 183–194.
Ulfarsson, G.F., Mannering, F.L., 2004. Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents. Accident Analysis and Prevention 36 (2), 135–147.

Please cite this article in press as: Jung, S., et al., Rainfall effect on single-vehicle crash severities using polychotomous response models. Accid. Anal. Prev. (2009), doi:10.1016/j.aap.2009.07.020

Wang, X., Abdel-Aty, M., 2008. Analysis of left-turn crash injury severity by conflicting pattern using partial proportional odds models. Accident Analysis and Prevention 40 (5), 1674–1682.
Weather Underground Inc., 2009. http://www.wunderground.com/US/WI/.
Wisconsin Traffic Crash Facts, 1999–2005, Wisconsin Department of Transportation.
Yau, K.K.W., 2004. Risk factors affecting the severity of single vehicle traffic accidents in Hong Kong. Accident Analysis and Prevention 36 (3), 333–340.
Yau, K.K.W., Lo, H.P., Fung, S.H.H., 2006. Multiple-vehicle traffic accidents in Hong Kong. Accident Analysis and Prevention 38 (6), 1157–1161.