Modeling Highway Safety and Simulation in Rainy Weather

Soyoung Jung, Xiao Qin, and David A. Noyce

Research was done to examine comprehensively the safety impact of rainy weather conditions on multivehicle crash frequency and severity and to validate the impact on traffic operations through microsimulation modeling. Three primary tasks were performed to meet these objectives. For weather data processing, available data were used to estimate the following factors: rainfall intensity, water film depth, and deficiency of car-following distance. For statistical modeling, negative binomial regression was used for crash frequency, and sequential logistic regression was tested with forward and backward formats for crash severity. A better format for the crash severity estimation was determined by combining all model performance measures. VISSIM was used to design traffic simulation models to reflect the effect of weather on traffic operation with five scenarios of the following weather-sensitive parameter adjustments: desired deceleration rate function, desired speed distribution, and headway time. As weather-related determinants, daily rainfall and wind speed were found to be statistically significant to crash frequency and severity estimations, respectively. VISSIM provided the most similar traffic data to the observed data when both desired speed distribution and deceleration rate function were adjusted. Statistical modeling in this research can be used to examine highway safety in rainy weather and to provide quantitative support on implementing road weather safety management strategies. Correspondingly, the adjustments of weather-sensitive traffic parameters will be the preliminary step to measure the strategy efficiencies through safety surrogate indexes in traffic simulation.

Severe crashes with injuries or fatalities have occurred during rainfall on wet pavement surfaces on Wisconsin highways. Many studies have focused on the relationship between snowy weather conditions and highway safety, and Wisconsin as one of the Snow Belt states. In Wisconsin, however, rain is the most frequent of all inclement weather events with average annual rainy days ranging from 100 to 150 (*1*). This implies that the rain event is more likely to be a potential risk to Wisconsin highway safety, prompting a need for comprehensive analysis of rainfall-related effects on Wisconsin highway safety.

Generally, highway safety is measured by crash frequency and severity. Crash frequency indicates the sum of crash occurrences at a specific location during a specific period. The severity of a crash occurrence is classified into five levels (*2*):

- Fatality (Type K);
- Incapacitating injury (Type A)—any visible injuries to a person who had to be carried from the scene;
- Nonincapacitating injury (Type B)—any visible injuries, such as bruises or abrasions;
- Possible injury (Type C)—no visible signs of injuries but complaint of pain or momentary unconsciousness; and
- Property damage only (PDO).

In risk analysis of highway safety, severe crashes with injuries and fatalities are more emphasized because of their significant human and economic loss potential. In Wisconsin from 1999 to 2006, 4,919 crashes with injuries and fatalities occurred on highways in all adverse weather conditions: rain, fog, snow, sleet, and wind (*3*). Of these crashes, 1,805, approximately 37%, were crashes with injuries and fatalities occurring in rainy weather. This proportion is the highest for all adverse weather conditions. During the same period, 899 multivehicle crashes occurred specifically on Wisconsin Interstate highways in rainy weather; the multivehicle crash frequency is approximately 1.5 times that of the single-vehicle crash frequency. Moreover, the proportion of severe multivehicle crash occurrences is 1.4 times more than the proportion of the severe single-vehicle crash occurrences.

Regardless of the number of crashes in rainy weather, comparatively few studies have been conducted in a disaggregated fashion to assess rainfall-derived factor effects on highway safety, particularly crash severity. Therefore, the objectives for this research were to develop a novel methodology for microscopic weather data estimation and to apply this methodology to statistical modeling of highway safety and weather-sensitive parameter adjustments during traffic simulation.

LITERATURE REVIEW

Rain-derived factors were found to affect crash frequencies through manner of collision in the crash, crash types, and road geometries. In a crash frequency study by negative binomial model, maximum daily rainfall per month, average daily rainfall per month, and number of rainy days per month were identified as significant for sideswipe, parked-vehicle, fixed-object, overturn, and rear-end crashes on Interstate highways (*4*). Shankar et al. estimated roadside crashes for the Washington State highway network by using a zero-inflated negative binomial model (*5*). In this study, rain precipitation during September and October increased the probability of a roadside crash, whereas less rain precipitation during April and June reduced the probability of a roadside crash.

S. Jung, Institute of Urban Sciences, Integrated Urban Research Center, 13 Siripdae-Gil, University of Seoul, Dongdaemun-Gu, Seoul 130-743, South Korea. X. Qin, Department of Civil and Environmental Engineering, South Dakota State University, 148 Crothers Engineering Hall, Brookings, SD 57007. D. A. Noyce, Department of Civil and Environmental Engineering, University of Wisconsin–Madison, 1415 Engineering Drive, Madison, WI 53706-1691. Current affiliation for S. Jung: Korea Advanced Institute of Science and Technology, Department of Civil and Environmental Engineering, 291 Daehak-ro, Yuseong-gu Daejeon 305-701, South Korea. Corresponding author: S. Jung, jung2@kaist.ac.kr.

Transportation Research Record: Journal of the Transportation Research Board, No. 2237, Transportation Research Board of the National Academies, Washington, D.C., 2011, pp. 134–143.
DOI: 10.3141/2237-15

Abdel-Aty et al. segregated Florida freeway crashes by apparently unrelated binary categories, including crash type, hour of the day, lighting condition, crash severity, and pavement condition (*6*). They provided insight into specific contributing factors that affected crash frequencies of each category by using a negative binomial regression model. In estimating crash frequencies by pavement condition, roadway curvature, on- and off-ramps, and annual average daily traffic (AADT) were found to be significant for increasing crashes on wet pavement. Caliendo et al. estimated multilane road crash occurrences classified by road geometries by comparing Poisson, negative binomial, and negative multinomial regression models (*7*). In their study, rain was found to be highly significant for increasing the number of severe crashes, especially on horizontal curves.

The impact of rain-related factors on crash severity has been identified in previous studies by manner of collision in the crash, vehicle types, road characteristics, and driver attributes. Shankar et al. estimated a nested logit model of crash severity that occurred on a Washington State rural Interstate (*8*). Wet-pavement rear-end collision indicators were found to increase the likelihood of possible injuries, capturing the effect of rear-end collisions occurring in rainy weather.

Duncan et al. used an ordered probit model to identify specific variables significantly influencing levels of injury in truck–passenger car rear-end involvements on divided roadways (*9*). In that study, interaction of wet and grade was found to significantly increase all injury propensities.

Khorashadi et al. explored the differences between urban and rural driver injuries in crashes involving large trucks by using multinomial logit analysis (*10*). The authors found that rain increased the likelihood of complaint of pain crashes only in urban areas.

Hill and Boyle investigated fatality and incapacitating injuries to occupants of passenger vehicles by using a logistic regression model (*11*). Their study showed that crashes in adverse weather conditions with rain, snow, or fog increased the risk of severe injuries to females that were 55 and older.

Several previous studies tried to integrate weather-related safety issues with simulating traffic operations via safety surrogate measures by identifying weather-sensitive traffic parameters. Zhang et al. evaluated the impact of weather-sensitive traffic parameters on freeway traffic operations by using microsimulation (*12*). As a result of sensitivity analysis, in a medium or high level of congestion, the mean free-flow speed and the car-following multiplier were found to most affect measures of effectiveness of freeways.

Lieu and Lin also examined the issues regarding the development of weather-related signal timing plans at an arterial of intersections by using CORSIM (*13*). The authors selected key weather-sensitive traffic parameters among available CORSIM parameters as follows: maximum speed, start-up lost time, queue discharge headway, and additional gaps between vehicles for safety and maximum deceleration rates for collision avoidance. The simulation results of this study demonstrated that potential benefits could be realized by retiming signals in inclement weather.

Tantillo and Demetsky also examined the impact of wet weather on traffic flow at signalized intersections by using VISSIM (*14*). The authors identified that the acceleration and deceleration rates were the most sensitive parameters to weather conditions. However, no data on the acceleration or deceleration rates were available in this study. Only free-flow speed and saturation flow rate data were used as weather-sensitive traffic parameters in this study. Correspondingly, this study showed statistically significant differences between dry and wet weather conditions, resulting in decreasing free-flow speeds and saturation flow rates on wet conditions.

STATISTICAL MODELS

In this study, negative binomial (NB) regression model was used to estimate crash frequencies caused by overdispersed response. NB regression assumes the count response Y follow a gamma distribution. The formula for NB regression model is specified so that

$$\ln\big(Y(r,t)\big) = \sum X(r,t)\beta + \epsilon \tag{1}$$

where

$Y(r, t)$ = random variable in area r during fixed period of time t,
$X(r, t)$ = explanatory variables,
β = coefficients of explanatory variables, and
$\exp(\epsilon)$ follows gamma distribution.

As the model performance, deviance is defined as two times the difference of log likelihood (LL) for the maximum achievable model and LL for the fitted model. As another performance measure, the Pearson chi-square statistic is defined as the squared difference between the observed and predicted values divided by the predicted value summed over all observations in the model. The deviance and Pearson chi-square statistic value divided by degree of freedom should be approximately one, indicating a good model fit for the data set.

For estimating crash severities, several modeling techniques for the discrete ordered outcome are available (*15–19*). A sequential logistic regression was selected in this study for the crash severity estimation not only to account for the inherent order of crash severities, but also to allow for different sets of predictors in the model specification by severity (*20*). The cumulative density function for the logistic regression is used to express the probability of a certain outcome in the standard logistic regression as follows (*21*):

$$\frac{P(Y)}{1-P(Y)} = \exp(\alpha + \beta X) \tag{2}$$

where

$P(Y)$ = probability of response outcome,
Y = binary response variable,
α = intercept parameter,
β = vector of parameter estimate, and
X = vector of explanatory variable.

A series of standard logistic regression was applied over two stages of two following formats to explore whether there was an impact in the ascending and descending development of crash severity levels, respectively: forward format and backward format. In this study, Crash Types K, A, and B were combined as the highest level of crash severity because of sample size issues. Crash Type C and PDO crashes were considered as the second highest and lowest level of crash severity, respectively. In the forward format,

Stage 1. PDO versus Types K, A, B, and C and
Stage 2. Type C versus Types K, A, and B

and in the backward format,

Stage 1. Types K, A, and B versus Type C and PDO and
Stage 2. Type C versus PDO.

A standard logistic regression model classifies a crash observation as an event, if the estimated probability of the crash observation severity is greater than or equal to a given cut point. The event means a crash observation classified as a more severe crash among binary crash severity levels on each stage. Otherwise, it is classified as a nonevent.

In statistical terms, sensitivity measures the proportion of actual events that are also predicted to be such. Similarly, specificity measures the proportion of actual nonevents that are also predicted to be such. The false-positive rate is the ratio of the number of nonevents incorrectly classified as events to the total nonevents, and the false negative rate is a ratio of the number of events incorrectly classified as nonevents to the sum of total events. A model that produced a high sensitivity and low false-negative rate at the stages of the highest crash severity classification was considered good because of the enormous economic loss.

As the final step of model evaluation, the estimated crash severity model was validated statistically by a bootstrap sampling method called jackknife. In each step, one observation was withheld from the data set used for model building. This restricted model was then compared with the model with use of the full data set. The process was repeated until all observations were tested. High and similar prediction accuracies between an estimated model and the model in the validation process suggest the estimated model is fairly robust.

DATA PROCESSING

In rainy weather, crash occurrences normalized by segment length, AADT, and vehicle miles traveled were the most frequent on southeastern Wisconsin Interstate highways, including I-43, I-94, I-43/94, and I-43/894 (*3*). According to Wisconsin weather station data, annual rainfall in the concurrent segment of I-43 and I-94 was approximately 14% higher than average annual rainfall on any other segment of southeastern Wisconsin Interstate highways. Moreover, this concurrent segment was the most congested of all Wisconsin Interstate highways according to AADT data. Therefore, the study areas for modeling highway safety and traffic simulation consisted of approximately 75 mi of southeastern Wisconsin highways and a 2.7-mi segment of I-43/94 between Howard Avenue and Mitchell Street, respectively. These study areas are shown in Figure 1.

In the study areas, traffic flow rate and the variation were found to be much greater on weekdays than on weekends. The weekday traffic pattern implies more likelihood of traffic conflicts, especially in adverse weather conditions. Consequently, only weekday traffic data during morning and evening peak hours in rainy weather conditions were used for northbound and southbound traffic simulation because of data homogeneity.

Databases

The crash data set for multivehicle crashes occurring in rainy weather was obtained from the Wisconsin Department of Transportation (DOT) crash database (MV4000 database). The crash data used in this study were filtered through several criteria to form data homogeneity: wet pavement during rainfall, multivehicle included in a crash, Interstate highway divided by barrier, no construction zone, and no hit and run. Consequently, 634 crashes were collected in the study area from 2004 to 2009. The crash occurrences by 1-mi segment of study area were counted for crash frequency estimation. For crash severity estimation, the crash severity was classified into three levels as mentioned in the section of statistical model to obtain a

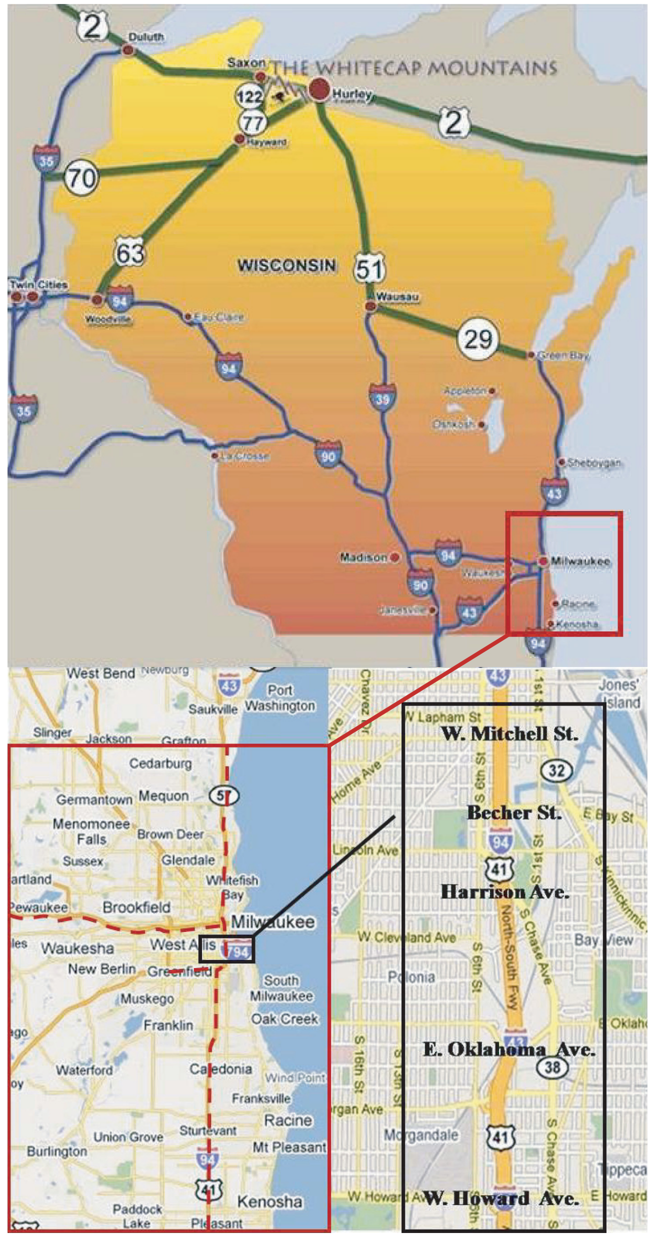

FIGURE 1 Study areas.

meaningful sample size (*22*). Crash frequencies by the severities and SAS program coding are provided in Tables 1, 2, and 3.

On each stage, crash severity was coded as binary values: 0 for lower severity and 1 for higher severity. A subsample was used on Stage 2 of each format after observations of a certain crash severity used in the previous stages were removed.

In addition to the crash database, V-SPOC traffic detectors installed at approximately 0.7-mi intervals collect and archive traffic data in the study area every 30 s. Considering the difference in density between crash and detector locations, average vehicle volume, speed, and occupancy data measured by 5-min intervals were obtained for 1 h before each crash.

The State Trunk Network (STN) highway log from the Wisconsin DOT contains the roadway geometric attributes of number and width for travel lane and shoulder and pavement surface materials. The

TABLE 1 Variables Used in Statistical Models: Crash Distribution

	Forward		Backward	
Injury Severity	Stage 1	Stage 2	Stage 1	Stage 2
Types K, A, and B	53 (1)[a]	53 (1)	53 (1)	—[b]
Type C	170 (1)	170 (0)	170 (0)	170 (1)
PDO	411 (0)	—[b]	411 (0)	411 (0)
Total	634	223	634	581

[a]SAS coding of crash severity level.
[b]Not applicable.

STN highway log was used to link the geometric attributes to the crash data set.

As one of the most important tasks in this research, real-time surface weather data at the time of crash were collected from Wisconsin weather station data run by Weather Underground Inc., which are the most reliable for obtaining the minute base measurements (*23*).

Weather Data Estimation

Weather data used in this study were temperature, wind speed and direction, and rainfall intensity. Additionally, existing rainfall intensity, traffic, and road geometry data were used to estimate water film depth and deficiency of car-following distance (DCD) as the secondary weather data. For crash frequency estimation, average daily rainfall per season was used, because the rain precipitation pattern was found to vary by season. For crash severity estimation, rain precipitation measured for 15 min before a crash was adopted because crash severity estimation is based on disaggregate analysis of a crash. The average measurement interval of rain precipitation was 15 min over weather station data in study areas.

Water Film Depth

Water film created by rainfall leads to a decrease in skid resistance between the tire and the pavement surface. To estimate water film depth, the empirical formula was produced by a study as follows (*24*):

$$D = 0.046\frac{(WS'I)^{1/2}}{S^{1/5}} \tag{3}$$

where

D = water film depth (mm/h),
I = rainfall intensity (mm/h),
$S' = S/S_c$,
$S = (S_l^2 + S_c^2)^{1/2}$,
S_l = longitudinal slope (%),
S_c = slope of pavement cross section (%), and
W = width of pavement (m).

Car-Following Distance Factor

In this research, there were no sight distance observations. Therefore, DCD was considered as a surrogate measure for the sight distance observed at the time of the crash. DCD represents the risk of losing control caused by driver overcorrection for avoiding any potential conflict, which is calculated with the following formula:

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

where SSD is stopping sight distance and AVG is average vehicle gap.

AVG is obtained by subtracting average vehicle length from inverse of vehicle density (*25*). Specifically, SSD is computed as follows (*26*):

$$\text{SSD} = 1.47Vt + 1.075\frac{V^2}{a} \tag{5}$$

where

V = vehicle speed (mph),
t = brake reaction time (s), and
a = deceleration rate (ft/s^2).

A detailed study about pavement conditions shows the relation between friction force and vehicle speed by levels of water film depth (*27*). Combining the relation in the study with pavement surface material information, deceleration rates to apply to the SSD equation were obtained by correlating to the pavement friction coefficient (*28*). In this research, 2.5 s was used as brake reaction time to encompass the capabilities of most drivers (*26*). Average 5-min traffic detector data containing the crash occurrence time was used to surrogate the real-time individual vehicular speeds at the crash moment because of data deficiency.

As a result of data processing, explanatory variables used in this research are shown in Tables 1, 2, and 3.

TABLE 2 Explanatory Variables Used in Crash Frequency Estimation

Roadway	Traffic	Weather
On- and off-ramp indicator	AADT/1,000	Avg. daily temperature/season
Speed limit change indicator	AADT/lane/1,000	Avg. hourly wind speed/day
Lane width and number change indicator	Avg. 5-min traffic volume/day	Avg. daily rainfall/season
Right and left shoulder width change indicator	Avg. 5-min traffic speed/day	Avg. hourly water film depth/season
Posted speed limit (mph)	Avg. 5-min occupancy/day	Avg. number of rainy days/month
Asphaltic cement pavement ratio	Avg. SD of 5-min volume/day	
Number of horizontal curves	Avg. SD of 5-min speed/day	
Lane width (ft) and number	Avg. SD of 5-min occupancy/day	
Right and left shoulder width		

NOTE: SD = standard deviation; avg. = average.

TABLE 3 Explanatory Variables Used in Crash Severity Estimation

Roadway at Crash Spot	Traffic[a]	Weather	Time and Crash Type	At-Fault Driver
Posted speed limit	Volume	Wind speed at the time of crash (mph)	Peak-hour (6–8 a.m./3–5 p.m.) = 1, off-peak = 2	Car = 1, truck–truck-tractor = 2, Motorcycle = 3
Asphaltic cement pavement indicator	Speed	15-min rainfall intensity (mm/15 min)	Tuesday to Thursday = 1, Monday and Friday = 2, Saturday and Sunday = 3	Driver's sex: Female = 1, male = 2
Right curve = 1, left curve = 2	Occupancy	Water film depth for an 1 h before crash moment (mm/h)	Season: Dec.~Feb. = 1, March~May = 2 June~Aug. = 3, Sept.~Nov. = 4	Sobriety = 1, under alcohol or drug effect = 2
Lane width	Standard deviation of volume	Temperature at crash moment (°F)	Median related crash = 1, fixed object outside roadway crash = 2, crash on travel lane = 3	Use of safety belt = 1, Not used = 2
Lane number	Standard deviation of speed		Sideswipe collision = 1, rear-end collision = 2, others = 3	Driver's age
Left shoulder width	Standard deviation of occupancy			Driver's action[e]
Right shoulder width				
First harmful location[b]				
Terrain[c]				
Light condition[d]				

[a]Average 5-min traffic data for 1 h before crash moment.
[b]Ramp = 1, shoulder–outside shoulder = 2, median = 3, on roadway = 4.
[c]Curve = 1, grade = 2, curve–grade = 3, tangent = 4.
[d]Daylight = 1, dusk–dawn–dark = 2, night but street light = 3.
[e]Action at crash moment: straight = 1, lane change–merge = 2, negotiating curve = 3, slowing–stopped = 4.

STATISTICAL MODEL RESULTS AND DISCUSSION

Because of the geographical and temporal variety of weather station data, rainfall intensity, water film depth, and wind speed were interpolated between three weather stations nearest to the crash spot. The interpolation was conducted by the inverse squared distance method that is appropriate for the field with short spatial correlation length scale and large variability (*29, 30*). However, temperature data from one weather station nearest to each crash was used because of its proximity.

For statistical modeling, negative binomial and sequential logistic regression models were estimated with PROC GENMOD and LOGISTIC statements in SAS 9.1. In both estimation models, the relation between a response and each of the explanatory variables was tested so a single significant predictor to the response could be individually selected.

Next, backward elimination was conducted to select the best multiple regression model with conventional *p*-value of parameter estimate significance for crash frequency and severity estimations. The backward elimination performs better to remove multicollinearity than forward and stepwise selection (*31*).

To measure the prediction accuracy of the crash severity estimation model, the event probability cut-points on each stage of each format were referred to the average proportions of actual crashes with higher severity on the stage occurred from 2000 to 2009.

Crash Frequency Estimation

The best multiple regression models for crash frequency estimation are provided in Tables 4 and 5.

The dispersion parameter greater than zero implies that NB regression was more appropriate than Poisson regression. Moreover, the fitted NB regression model performed well to predict multivehicle

TABLE 4 Multiple NB Regression, Goodness of Fit

Criteria	Values
Deviance/DF	1.1980
Pearson chi-square/DF	0.9508
LL_0[a]	1,016.7376
LL (NB)[b]	1,052.0650
LR statistic	77.932
Dispersion	0.3740

NOTE: DF = degrees of freedom; LR = likelihood ratio.
[a]LL_0 = log likelihood for intercept-only model.
[b]LL (NB) = log likelihood for the best fitted NB regression model.

TABLE 5 Multiple NB Regression, Maximum Likelihood Estimates

Parameter	Estimate	Standard Error	*p*-Value
Intercept	−3.4622	0.6263	<.0001
Off-ramp indicator	0.7403	0.2228	.0009
LSWI[a]	0.7223	0.2029	.0004
AADT/1,000	0.0312	0.0087	.0004
SDV[b]	0.0392	0.0119	.0010
Average daily rainfall in winter	1.0270	0.2455	<.0001
SLCI*LSW[c]	0.0608	0.0244	.0128

[a]Indicator of left shoulder width change.
[b]Average standard deviation of 5-min traffic volume in a day.
[c]Interaction between indicator of speed limit change and left shoulder width.

crash frequency based on deviance, Pearson chi-squared values, and likelihood ratio statistic.

Average daily rainfall in winter, indicator of off-ramp existence, and change of left shoulder width were found to be statistically significant to highly increased multivehicle crash occurrences in rainy weather. These findings imply that changes in roadway geometry and weather condition are the most significant factors in multivehicle crash occurrences in rainy weather.

Daily rainfall in winter was most likely to increase multivehicle crash frequency among all significant predictors. Less sunlight in winter and high rainfall intensity can obscure drivers' visibility. This can cause an increase in multivehicle crash occurrences.

A change in the left lane width was also likely to significantly increase multivehicle crash frequency in rainy weather. Generally, vehicles easily go into skid on wet pavement because of reduced pavement friction force. Changing left-lane width can increase the difficulty of maneuvering vehicles on wet travel lanes.

Standard deviation of 5-min traffic volume, AADT, and interaction between speed limit change and left shoulder width were also likely to increase the number of multivehicle crashes, but these variable effects were relatively lower.

Adjusted *R*-squared value for this multiple regression model was found to be .32, which is a reasonable global goodness of fit based on a comprehensive study (*32*).

Crash Severity Estimation

The best multiple model for each format of sequential logistic regression is shown in Table 6.

For global model fit, parameter estimates, prediction accuracies, and cross-validation, the backward format was found to perform better than forward format, especially for predicting the most severe crashes with nonincapacitating, incapacitating, and fatal injuries.

On the first stage of the backward format, wind speed as a weather-related factor was likely to decrease the most severe crashes in rainy weather. Strong wind during rainfall can obstruct a driver's vision, increasing driver attention. This is a possible explanation to a decrease in the likelihood of occurrence of the most severe crash. At-fault driver and crash type related factors were found to affect the most severe multivehicle crashes more than were weather-related factors. Wearing a safety belt was found to decrease the likelihood of the most severe level of crash. Negotiating a curve by an at-fault driver and collision with fixed objects outside the roadway were found to be likely to increase the most severe crashes. Because of reduced pavement friction during rainfall, vehicles are more likely to go into skid, especially on curves, which may cause collisions with fixed objects outside the roadway. This process could cause an increase in the most severe crashes.

In the second stage, passenger car, lane change or merge by the at-fault driver, standard deviation of 5-min traffic occupancy, and Monday or Friday decreased the likelihood of possible injury crashes. However, female driver and median-related crashes were likely to increase possible injury crashes. The effects for lane change or merge, traffic occupancy, and day of the week may be opposite to expectations. Traffic congestion with long traffic occupancy usually occurs during commuting peak hours on Monday or Friday. Correspondingly, this traffic condition can make drivers more cautious during lane changing or merging, especially in rainy weather, because drivers recognize the risk of driving in rainfall from experience; this could be why crash severity decreased in Stage 2.

Thus, it is straightforward to explain the effects of weather-related factors on highway safety in rainy weather by integrating crash frequency and severity estimation models. Applying the methodologies used in the models, microsimulation will be conducted to improve traffic safety in rainy weather as follows.

TRAFFIC SIMULATION MODEL

To complete a comprehensive analysis of highway safety related to rainy weather, safety improvement strategies based on the results in statistical models should be evaluated through traffic simulation before the strategies are implemented. In current traffic simulation programs, safety surrogate indices are used to indirectly measure traffic safety during traffic simulation (*33*). However, the preliminary step for the simulation is to replicate traffic observed on wet pavement under rainfall. That is, weather-sensitive traffic parameters should be adjusted preferentially to reflect rainy weather effects on traffic operations before the simulated safety surrogate indices are used. The significance of traffic simulation in this research is emphasized in conducting the preliminary step. Correspondingly, traffic simulation in this research was designed to validate previous findings of weather-sensitive traffic parameters and to develop a novel method for rainy weather microsimulation models. The novel methodology of weather data estimation was used in both statistical highway safety models and weather-sensitive parameter adjustments during microsimulation.

From the findings in the previous studies, operating speed in free-flow conditions, deceleration rate, and headway time were found to be key weather-sensitive traffic parameters (*11–13*). In this research, therefore, the following parameters in VISSIM, a frequently used traffic simulation software program, were selected as weather-sensitive traffic parameters: desired vehicle deceleration function, desired speed distribution, and headway time.

There were three primary tasks for simulating rainy weather traffic in this research. First, traffic simulation was conducted with loading traffic volume observed in rainy weather under the default VISSIM parameter setting, which is Scenario 5 in the following. Second, weather-sensitive parameters were adjusted by the following scenarios:

Scenario 1. Change desired speed distribution only;

Scenario 2. Change both desired speed distribution and vehicle decelerate rate function;

Scenario 3. Change both desired speed distribution and headway time value;

Scenario 4. Change desired speed distribution, vehicle deceleration rate function, and headway time simultaneously; and

Scenario 5. Change nothing.

To reflect rainy weather effects on traffic operations, the following criteria were applied to propose the first four simulation scenarios. First, default values of any other VISSIM parameters are not changed in five scenarios, except for three weather-sensitive traffic parameter values. Second, desired speed distribution is adjusted in every scenario because traffic speed in rainy weather was observed to validate the simulated traffic speed. That is, scenarios for individually adjusting deceleration rate function or headway time were not considered because only traffic speed data were observed. Finally, the simulation system performance under each scenario was measured to select the most realistic scenario that shows the highest similarity between observed and simulated data.

TABLE 6 Multiple Sequential Logistic Regression

Fit Measure	Parameter	Forward Format		Backward Format	
		Stage 1	Stage 2	Stage 1	Stage 2
Global null	Chi-square (Chisq)	48.5202	32.3974	49.6830	33.5044
	Degrees of freedom	6	4	5	6
	P > Chisq	<.0001	<.0001	<.0001	<.0001
Maximum likelihood estimate	Intercept	2.7206/0.7208/.0002[a]	0.9417/0.5551/0.0898[a]	0.7059/0.5968/0.3979[a]	–0.3628/0.2290/0.1132[a]
	Passenger car	–0.5041/0.1798/.604/0.0050[b]			–0.5198/0.1956/0.595/0.0079[b]
	Female driver				0.3845/0.1916/1.469/0.0448[b]
	DR 2[c]	–0.9126/0.2913/0.401/0.0017[b]			–0.8060/0.3187/0.447/0.0114[b]
	DR 3[d]			1.1823/0.5222/3.262/0.0236[b]	
	DR 4[e]	–0.3980/0.1918/0.672/0.0380[b]			
	Safety belt	–1.6092/0.6817/0.200/0.0183[b]		–1.3376/0.6304/0.262/0.0399[b]	
	Standard deviation of traffic occupancy	–0.0684/0.0303/0.934/0.0238[b]			–0.0679/0.0333/0.934/0.0411[b]
	Traffic volume			–0.0180/0.0043/0.982/<.0001[b]	
	Deficiency of car-following distance		–0.0060/0.0020/0.994/0.0009[b]		
	Wind speed			–0.0502/0.0231/0.951/0.0299[b]	
	SIDE[f]		–1.5869/0.8715/0.205/0.0686[b]		
	SIDE * CT 3		2.0054/0.9833/7.429/0.0023[b]		
	Median-related crash				1.2690/0.4232/3.557/0.0027[b]
	CT 2[g]			1.7650/0.6897/5.841/0.0105[b]	
	CT 3[h]	–1.0364/0.3558/0.3550/0.0036[b]	–1.8194/0.5687/0.162/0.0014[b]		
	Monday or Friday				–0.4985/0.2047/0.607/0.0149[b]
Prediction accuracy	Total	56.3	65.5	66.9	53.9
	Sensitivity	66.8	62.3	69.8	50.6
	Specificity	50.4	66.5	66.6	55.2
	False positive	56.8	63.3	84.0	68.1
	False negative	27.1	15.0	4.0	27.0
Cross validation	Total	56.5	65.9	66.9	54.0

[a]Estimate, standard error, *p*-value for intercept.
[b]Estimate, standard error, odds-ratio, *p*-value.
[c]Lane changing or merging by at fault driver before crash occurrence.
[d]Negotiating curve by at-fault driver before crash occurrence.
[e]Slowing or stopping by at-fault driver before crash occurrence.
[f]Sideswipe collisions.
[g]Crash type related to fixed object outside roadway.
[h]Crash occurred on travel lane.

Base Simulation Data

Weather station data, V-SPOC traffic detector data, and the Wisconsin STN log in the traffic simulation area were combined and a sample of rainy days was selected for traffic simulation with the following criteria: continuous rain during morning or evening peak period with rainfall precipitation greater than 0.01 in per hour and temperature greater than 32°F to exclude the effect of icy pavement surface. Traffic data on dry weather weekdays near the rainy weather weekdays were also collected to be compared with the rainy weather traffic pattern. For Scenario 5, the base parameter setting is summarized as follows:

- Vehicle type, class, category: car and truck with 20% truck in traffic (*34*);
- Link type and length: 2.7-mi two-way freeway;
- Travel lane number and width: three 12-ft lanes; and
- Speed limit: 55 mph.

The VISSIM parameter setting was commonly applied to five simulation scenarios.

Desired Speed Distribution

Real-time traffic detector speed data were used to obtain the mean and standard deviation of hourly traffic speed on rainy weekdays. Assuming the normality of continuous speed data and a 95% confidence interval, approximately 66% of the traffic speed observed in rainy weather would fall within one standard deviation of the mean and roughly 95% of the data would fall within two standard deviations of the mean in the desired vehicle speed distribution.

Desired Deceleration Rate Function

Because of data deficiency, vehicle deceleration rate was estimated with available data, such as rainfall intensity, road geometry, traffic volume, and speed. The estimation methodology is explained in the section on the car-following distance factor. Relating the estimated vehicle deceleration rates to V-SPOC traffic detector speed on the rainy weather weekday sample, a linear regression between the vehicle deceleration rate and traffic speed was fitted.

Headway Time

VISSIM assumes that freeway car-following behavior follows the Wiedemann 99 car-following model. To estimate headway time, necessary because of the data deficiency, average spacing was calculated by dividing 5,280 ft/mi by lane density. The lane density was computed by dividing hourly flow rate by space mean speed. Since the hourly flow rate and space mean speed were collected, headway time on each rainy-weather weekday was estimated by dividing average spacing by average vehicle speed.

Simulation Performance Measures

During simulation, two criteria were needed to stop traffic parameter calibration and to judge the similarity between simulated and observed traffic data.

To stop the traffic parameter calibration, the Wisconsin DOT system performance measure was used in this research. The Wisconsin DOT measure includes the number of observations that meet the difference between simulated and observed flows falls within 15% of the observed traffic flow is greater than 85% of the number of total observations (*35*). Additionally, the simulated speed and occupancy are visually acceptable, which indicates that the simulated measurements in traffic speed and occupancy are consistent with stable traffic flow–speed–occupancy relationships (*36*).

TABLE 7 Weather-Sensitive Parameter Adjustments

Parameter	Adjustment	a.m.	p.m.
Desired speed distribution	Mean (mph)	50	45
	Standard deviation	6	10
	Maximum	62	65
	Minimum	38	25
	17th percentile	44	35
	84th percentile	56	55
Desired deceleration rate function	Deceleration rate^2 (ft/s^2) = 17.2 – 0.06 ∗ vehicle speed (mph)		
Headway time	1.1 s (Wiedemann 99 car-following model default = 0.9 s)		

As the simulation similarity criteria, the root-mean-square percent error (RMSPE) was used in this research to quantify overall error of the simulator, which is shown in the following (*37*):

$$\mathrm{RMSPE} = \left[\frac{1}{N}\sum_{1}^{n}\frac{(Y\mathrm{sim} - Y\mathrm{obs})^2}{(Y\mathrm{obs})^2}\right]^{1/2} \tag{6}$$

where

Ysim = simulated traffic performance estimates,
Yobs = observed traffic performance estimates, and
N = total number of observations.

TRAFFIC SIMULATION RESULTS

For rainy weather conditions, 56 weekdays were collected with 15-min traffic volume, speed, and occupancy data. For robust simulated results, 10 automatic runs were conducted to produce average simulated traffic data measurements on each weekday.

Compared with traffic speeds observed in dry weather, in rainy weather a reduction in traffic speed and increments of traffic occupancy were clearly found, and the standard deviation of traffic speed was approximately twice as high during peak periods. Key weather-sensitive parameters were adjusted for Scenarios 1 through 4, as shown in Tables 7 and 8. The simulation results with weather-sensitive parameter adjustments are also provided in Tables 7 and 8.

TABLE 8 Weather-Sensitive Simulation Performance

	RMSPE				Ratio of Weekdays with Acceptable Performance			
Scenario	V^a	SPD[b]	OCC[c]	Average[d]	V	SPD	OCC	Average[e]
1	0.02	0.13	0.21	0.12	1	0.70	0.45	0.72
2	0.02	0.13	0.20	0.12	1	0.75	0.50	0.75
3	0.02	0.14	0.20	0.12	1	0.63	0.45	0.69
4	0.02	0.14	0.20	0.12	1	0.71	0.46	0.72
5	0.02	0.24	0.34	0.20	1	0.50	0.15	0.55

[a]Hourly traffic volume.
[b]Hourly traffic speed.
[c]Hourly traffic occupancy.
[d](sum of RMSPE for each traffic data including V, SPD, and OCC)/3.
[e](sum of ratio of weekdays with acceptable performance for each traffic data)/3.

The average RMSPE in Scenario 5 was clearly found to be the greatest of that of all the scenarios. This implies that VISSIM did not efficiently simulate rainy weather traffic under the dry pavement assumption. In a comparison of Scenarios 1 through 4, however, the best scenario apparently was not identified by RMSPE. Therefore, the ratio of the number of weekdays with acceptable performance to total number of weekdays was calculated for each simulated traffic datum to measure the system performance of each scenario. If RMSPE of each traffic data measurement on a weekday was less than 0.15, the weekday was counted as the day with acceptable performance of simulation.

As a result, the average ratio was 0.75 in Scenario 2, the greatest of all the scenarios. This implies that the adjustments in both desired speed distribution and deceleration rate function were effective to particularly replicate traffic operations observed in rainy weather through microsimulation.

CONCLUSIONS AND FUTURE EXTENSION OF RESEARCH

Road surface conditions and visibility during rainfall have not been sufficiently characterized in previous studies. This study, therefore, estimated several novel data, including 15-min rainfall intensity, water film depth, and deficiency of car-following distance, to microscopically reflect rainy weather conditions at the time of crash. The weather data estimation method was used to comprehensively examine rainfall-derived factor impact on highway safety and applied to microsimulation of traffic operations in rainy weather.

In multivehicle crash frequency estimation, daily rainfall in winter, off ramp, and change of left shoulder width comparatively increased the likelihood of crash occurrence. Particularly, daily rainfall in winter was likely to increase the crash frequency by the highest factor. This result can be caused by low visibility from rainfall and less sunlight in winter.

In multivehicle crash severity estimation, the backward sequential logistic regression model was found to perform better in predicting the most severe crashes with fatal, incapacitating, and nonincapacitating injuries. As a weather determinant, strong wind was found to be likely to decrease crash severity. Strong wind combined with rainfall can also reduce driver vision, resulting in cautious driving in rainy weather. This is a possible explanation for the decrease in the most severe crashes with strong wind. In addition, negotiating curve by at-fault driver and roadside fixed objects were found to significantly increase the likelihood of the most severe vehicle crashes in rainy weather conditions. These findings imply that there is a need to implement countermeasures to avoid skidding on curves and off-road collisions during rainfall.

In traffic simulation, following weather-sensitive traffic parameters were adjusted to reflect rainy weather effects on traffic operations through microsimulation: desired speed distribution, desired deceleration rate, and headway time. Simultaneous adjustment of the first two parameters was found to best replicate traffic in rainy weather. To improve traffic safety in rainy weather, possible strategies based on statistical model results will be efficiently tested by safety surrogate measures in simulations adjusting for weather-sensitive traffic parameters.

Findings related to weather, particularly rain-derived factor and road geometry effects, in highway safety estimation models will provide quantitative support on comprehensive safety improvement strategies in rainy weather. Implementing weather warning, lighting, and antiskid systems are possible strategies. Correspondingly, the strategies can be examined through traffic simulation with weather-sensitive parameter adjustments conducted in this research.

For future research, databases on a regional and national scale are needed to explore and validate the impact of rainy weather on highway safety more comprehensively. Furthermore, the use of weather-based adjustment factors in the core simulation models should be expanded so that these microsimulation models can effectively consider adverse weather impact, because researchers need better models to evaluate safety decisions and identify the best techniques of safety surrogate measures.

ACKNOWLEDGMENTS

This work was supported indirectly by a National Research Foundation of Korea grant, funded by the government of South Korea.

REFERENCES

1. Blair, T. A. *Climatology: General and Regional,* Prentice-Hall, New York, 2007.
2. *Definition of Terms.* Florida Department of Highway Safety and Motor Vehicles. Tallahassee. http://www.flhsmv.gov/hsmvdocs/cf2004/PG6.htm. Accessed July 24, 2010.
3. *MV 4000 Crash Database Query Tools.* Traffic Operations and Safety Laboratory, University of Wisconsin–Madison. http://transportal.cee.wisc.edu/applications. Accessed July 25, 2010.
4. Shankar, V., F. Mannering, and W. Barfield. Effect of Roadway Geometrics and Environmental Factors on Rural Freeway Accident Frequencies. *Accident Analysis and Prevention*, Vol. 27, No. 3, 1995, pp. 371–389.
5. Shankar, V. N., S. Chayanan, S. Sittikariya, M.-B. Shyu, N. K. Juvva, and J. C. Milton. Marginal Impacts of Design, Traffic, Weather, and Related Interactions on Roadside Crashes. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1897,* Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 156–163.
6. Abdel-Aty, M., R. Pemmanaboina, and L. Hsia. Assessing Crash Occurrence on Urban Freeways by Applying a System of Interrelated Equations. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1953*, Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 1–9.
7. Caliendo, C., M. Guida, and A. A. Parisi. Crash Prediction Model for Multilane Roads. *Accident Analysis and Prevention,* Vol. 39, No. 4, 2007, pp. 657–670.
8. Shankar, V., F. Mannering, and W. Barfield. Statistical Analysis of Accident Severity on Rural Freeways. *Accident Analysis and Prevention,* Vol. 28, No. 3, 1996, pp. 391–401.
9. Duncan, C. S., A. J. Khattak, and F. M. Council. Applying the Ordered Probit Model to Injury Severity in Truck–Passenger Car Rear-End Collisions. In *Transportation Research Record 1635,* TRB, National Research Council, Washington, D.C., 1998, pp. 63–71.
10. Khorashadi, A., D. Niemeier, V. Shankar, and F. Mannering. Differences in Rural and Urban Driver-Injury Severities in Accidents Involving Large-Trucks: An Exploratory Analysis. *Accident Analysis and Prevention,* Vol. 37, No. 5, 2005, pp. 910–921.
11. Hill, J., and L. Boyle. Assessing the Relative Risk of Severe Injury in Automotive Crashes for Older Female Occupants. *Accident Analysis and Prevention*, Vol. 38, No. 1, 2006, pp. 148–154.
12. Zhang, L., P. Holm, and J. Colyar. *Identifying and Assessing Key Weather-Related Parameters and Their Impacts on Traffic Operations Using Simulation.* FHWAHRT-04-131. FHWA, U.S. Department of Transportation, 2004.
13. Lieu, H. C., and S.-M. Lin. Benefit Assessment of Implementing Weather-Specific Signal Timing Plans by Using CORSIM. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1867,* Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 202–209.

14. Tantillo, M., and M. Demetsky. *Investigating the Impacts of Rainy Weather at Isolated Signalized Intersections.* UVACTS-15-13-90. Center for Transportation Studies, University of Virginia, Charlottesville, 2006.
15. Abdel-Aty, M. Analysis of Driver Injury Severity Levels at Multiple Locations Using Ordered Probit Models. *Journal of Safety Research,* Vol. 34, No. 5, 2003, pp. 597–603.
16. Milton, J. C., V. Shankar, and F. Mannering. Highway Accident Severities and the Mixed Logit Model: An Exploratory Empirical Analysis. *Accident Analysis and Prevention,* Vol. 40, No. 1, 2008, pp. 260–266.
17. Wang, X., and M. Abdel-Aty. Analysis of Left-Turn Crash Injury Severity by Conflicting Pattern Using Partial Proportional Odds Models. *Accident Analysis and Prevention,* Vol. 40, No. 5, 2008, pp. 1674–1682.
18. Eluru, N., C. Bhat, and D. Hensher. A Mixed Generalized Ordered Response Model for Examining Pedestrian and Bicyclist Injury Severity Level in Traffic Crashes. *Accident Analysis and Prevention,* Vol. 40, No. 3, 2008, pp. 1033–1054.
19. Peterson, B., and F. E. Harrell, Jr. Partial Proportional Odds Model for Ordinal Response Variables. *Applied Statistics,* Vol. 39, No. 2, 1990, pp. 205–217.
20. Jung, S., X. Qin, and D. Noyce. Rainfall Effect on Single-Vehicle Crash Severities Using Polychotomous Response Models. *Accident Analysis and Prevention,* Vol. 42, No. 1, 2010, pp. 213–224.
21. Kleinbaum, D., and M. Klein. *Logistic Regression: A Self-Learning Text.* Springer, New York, 2002.
22. Agresti, A. *An Introduction to Categorical Data Analysis.* John Wiley and Sons, New York, 1996.
23. *Wisconsin Weather Station Data.* Weather Underground. http://www.wunderground.com/US/WI. Accessed July 25, 2009.
24. Russam, K., and N. Ross. *The Depth of Rain Water on Road Surfaces.* Road Research Laboratory, Ministry of Transport, Crowthorne, Berkshire, United Kingdom, 1968.
25. Roess, R., E. Prassas, and W. Mcshane. *Traffic Engineering.* Pearson Education, Upper Saddle River, N.J., 2004.
26. *Green Book: A Policy on Geometric Design of Highways and Streets,* 5th ed. AASHTO, Washington, D.C., 2004.
27. Kokkalis, A., and O. Panagouli. Factual Evaluation of Pavement Skid Resistance Variation: Surface Wetting. *Chaos, Solitons and Fractals,* Vol. 9, No. 11, 1998, pp. 1875–1890.
28. *Reference Tables-Coefficient of Friction.* Engineer's Handbook. http://www.engineershandbook.com/Tables/frictioncoefficients.htm. Accessed July 24, 2009.
29. Patrick, N., and D. Stephenson. Spatial Variation of Rainfall Intensities for Short Duration Storms. *Hydrological Sciences,* Vol. 35, No. 6, 1990, pp. 667–680.
30. Press, W., S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes: The Art of Scientific Computing.* Cambridge University Press, New York, 2007.
31. Chatterjee, S., A. Hadi, and B. Price. *Regression Analysis by Example.* John Wiley and Sons, New York, 2000.
32. Kweon, Y., and K. Kockelman. Overall Injury Risk to Different Drivers: Combining Exposure, Frequency, and Severity Models. *Accident Analysis and Prevention,* Vol. 35, No. 3, 2003, pp. 414–450.
33. Gettman, D., and L. Head. *Surrogate Safety Measures from Traffic Simulation Models: Final Report.* FHWA-RD-03-050. FHWA, U.S. Department of Transportation, 2003.
34. *Facilities Development Manual.* Wisconsin Department of Transportation, Madison, 2008.
35. Dowling, R., A. Skabardonis, and V. Alexiadis. *Traffic Analysis Toolbox Volume III: Guidelines for Applying Traffic Microsimulation Software.* FHWA-HRT-04-040. FHWA, U.S. Department of Transportation, 2004.
36. May, A. *Traffic Flow Fundamentals.* Prentice-Hall, Englewood Cliffs, N.J., 1990.
37. Pindyck, R., and D. Rubinfeld. *Econometric Models and Economic Forecasts,* 4th ed. Irwin McGraw-Hill, Boston, Mass., 1997.

The Safety Data, Analysis, and Evaluation Committee peer-reviewed this paper.