Exploring weather-related factors affecting the delay caused by traffic incidents
Mitigating the negative effect of traffic incidents

Su, Xiangtong; Zhi, Danyue; Song, Dongdong; Tian, Le; Yang, Yitao

DOI
10.1016/j.scitotenv.2023.162938

Publication date
2023

Document Version
Final published version

Published in
Science of the Total Environment

Science of the Total Environment 877 (2023) 162938

Contents lists available at ScienceDirect

Science of the Total Environment

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

Exploring weather-related factors affecting the delay caused by traffic incidents: Mitigating the negative effect of traffic incidents

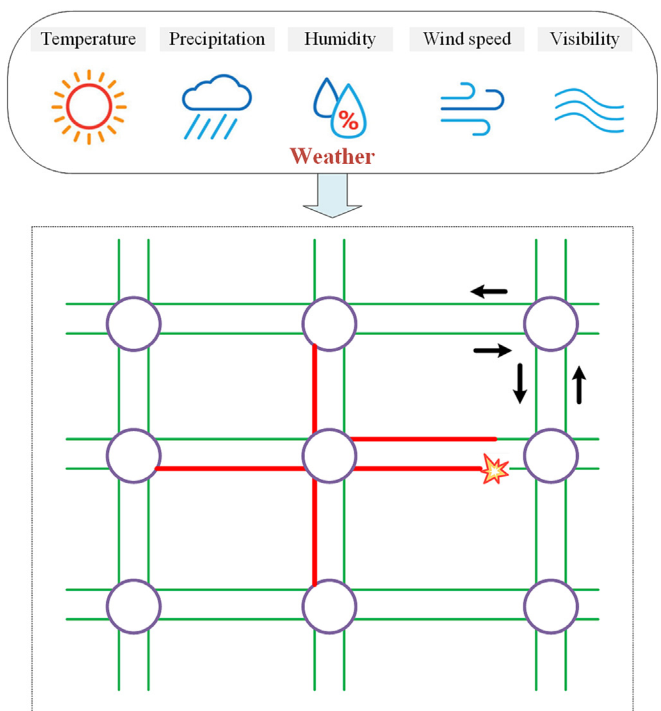
Xiangtong Su [a], Danyue Zhi [a,*], Dongdong Song [a,*], Le Tian [b], Yitao Yang [c]

[a] MOT Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
[b] China Communications Information Technology Group Co., Ltd, Beijing 100088, China
[c] Department of Transport & Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg1, Delft 2628 CN, the Netherlands

HIGHLIGHTS

- This study examines the influence of weather conditions on the delay caused by traffic incidents.
- The RPHDHM model is introduced to determine the relationship between weather conditions and the duration of traffic delay.
- The RPLHM model is introduced to explore the relationship between weather conditions and the delay severity.
- Substantial differences in probabilities of traffic delay severity across various time models are observed.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Pavlos Kassomenos

Keywords:
Weather condition
Traffic delay
Hazard-based duration model

ABSTRACT

Background: Existing studies mainly focus on the relationship between real-time weather and traffic crash injury severity, while few scholars have investigated the operation risk levels caused by traffic incidents. Identifying weather-related factors that affect the incident-induced delay is helpful for estimating the delay levels when an incident occurs. Accordingly, the present study profoundly explores the relationship between weather conditions and traffic delays caused by traffic incidents.

* Corresponding authors.
E-mail addresses: danyue-zhi@bjtu.edu.cn (D. Zhi), dongdongsong@bjtu.edu.cn (D. Song).

http://dx.doi.org/10.1016/j.scitotenv.2023.162938
Received 3 January 2023; Received in revised form 21 February 2023; Accepted 14 March 2023
Available online 17 March 2023

Random parameters logit model
Heterogeneity in the means

Methods: The traffic incident and weather datasets from January 1 to December 31, 2020, in New York State are used. To that end, the hazard-based duration and multinomial logit modeling frameworks are employed to determine the effect of weather conditions on the duration of traffic delay and the delay severity, respectively. More importantly, to account for multiple layers of unobserved heterogeneity, a random parameter with heterogeneity in means approach is introduced into the above two models.
Results: (1) The strong breeze (wind speed over 8 m/s) and low visibility (visibility under 5 km) significantly affect the duration of delay. (2) Hot day (between 20 and 30 °C) has a 344.03 % greater probability of minor delay. A strong breeze has a higher probability of severe delay. The low visibility is found to increase the estimated odds of moderate delay and severe delay by 51.15 % and 13.39 %, respectively. In comparison, the normal visibility (between 10 and 20 km) significantly decreases the estimated odds of severe delay by 119.17 %.
Conclusions: Compared with other weather factors, wind speed, temperature, and visibility have the greatest impact on the traffic delay levels after a traffic accident, and there are significant differences in the impact under different delay severity. Findings from this study will help policymakers to establish comprehensive differentiating security measures to resolve traffic delays.

1. Introduction

In recent years, severe weather conditions have been one of the most critical factors affecting the safe operation of the road network, which caused frequent traffic accidents and resulted in serious injuries and fatalities. From 2014 to 2019, nearly 30 % of road accidents in Canada occurred each year under severe weather conditions, leading to over 910,000 injuries and >11,000 fatalities (Transport Canada, 2021). In the US, there are over 5.8 million vehicle collisions yearly, approximately 21 % of which are weather-related. On average, almost 5000 people are killed in weather-related crashes, and over 418,000 are injured yearly (FHWA, 2022).

Many scholars have studied the impact of severe weather on traffic safety, such as examining how weather-related factors affect traffic accident likelihood or injury severity (Ghasemzadeh and Ahmed, 2019; Liu et al., 2017; Seeherman and Liu, 2015; Yu et al., 2015; Yu and Abdel-Aty, 2014a,b; Zhan et al., 2020). Table 1 presents a detailed description of some representative literature in the field mentioned above in recent years.

In addition to studying the impact of severe weather on the crash likelihood or injury severity, investigating how weather affects the duration of the incident-induced delay can also be relevant, helping traffic regulators develop improved incident response countermeasures. Accordingly, some scholars have examined the weather-related factors influencing traffic incident durations. For example, Vlahogianni and Karlaftis (2013) proposed a survival neural network model for predicting the duration of events affected by congestion, finding that rainfall intensity is an essential feature. Dimitriou and Vlahogianni (2015) developed a fuzzy rule-based system to study the duration of highway incidents and revealed that higher rainfall is associated with longer durations. Javid (2018) developed a robust regression model to estimate incident-caused travel time variability based on traffic and weather data. Precipitation (including rainfall and snowfall) can lead to congestion and delays, especially during peak hours (Bi et al., 2022; Koetse and Rietveld, 2009; Tsapakis et al., 2013). According to Giang et al. (2014), the timely transfer of patients by ambulance was delayed by 8 % - 10 % due to precipitation.

It is noted that more academics have analyzed the factors influencing the traffic incident duration based on survival analysis models. Specifically, Alkaabi et al. (2011) found that weather-related factors such as windy and rainy conditions significantly increase highway clearance time based on the Weibull AFT (accelerated failure time) model. Li and Shang (2014) studied incident duration based on flexible parametric hazard-based models, reporting that the incident duration is longer in Winter. This phenomenon may be related to poor weather conditions during Winter, such as heavy snowfall and low visibility (Yu et al., 2013). Furthermore, to account for

Table 1
Summary of research on the impact of weather conditions on traffic safety.

Author (year)	Study area	Study period	Weather-related variables	Dependent variables	Model
Abdel-Aty et al., 2011	Central Florida (US)	2007–2009	Rainfall; visibility	Crash likelihood	Bayesian matched case-control logistic
Hassan and Abdel-Aty, 2013	Central Florida (US)	2007–2009	rainfall; visibility	Crash likelihood	Random Forests and matched case-control logistic regression models
Xu et al., 2013a	California (US)	2008;2010	Rainfall; visibility	Accident likelihood	Bayesian random intercept logistic regression models
Xu et al., 2013b	San Francisco Bay (US)	2008	Rain; fog	Crash severity level	Sequential logit model
Yu et al., 2013	Colorado (US)	2010–2011	Visibility, precipitation, and temperature	Crash-frequency	Bayesian random effect models
Yu and Abdel-Aty, 2014a	Colorado and Orlando (US)	2007–2011	Temperature; snow season;	Crash injury severity level	Fixed parameter logit model; support vector machine; correlated random parameter logit model
Yu and Abdel-Aty, 2014b	Colorado (US)	2007–2011	Snow; visibility	Crash injury severity level	Hierarchical Bayesian binary probit models
Seeherman and Liu, 2015	California (US)	2007–2013	Snowfall; rainfall	Crashes and incidents frequencies	Negative binomial regression model
Yu et al., 2015	Colorado (US)	2008–2010	Visibility; precipitation	Crash rate	Correlated random parameter Tobit model
Liu et al., 2017	Maryland (US)	2000–2012	Temperature; precipitation	Risk of motor vehicle collisions	Time-stratified case-crossover analysis; conditional logistic regression models
Ghasemzadeh and Ahmed, 2019	Washington state (US)	2009–2013	Clear; rain; snow; fog; hail; severe crosswind; blowing snow	Crash severity level	Ordered probit model
Zhan et al., 2020	Shenzhen (China)	2010–2016	Temperature; precipitation	Road traffic casualty	Time-stratified case-crossover analysis; conditional quasi-Poisson regression

the unobserved heterogeneity, Hojati et al. (2013) applied the Weibull AFT models with random parameters to examine the effects of air temperature, wind speed, and rainfall on the crash-induced traffic incident duration. Islam et al. (2022) evaluated the factors influencing the incident clearance duration using the random parameters hazard-based duration model and discovered that rainy conditions increase the clearance time.

However, traditional random parameters models assume that the means of random parameters are the same for all observations, resulting in failure to explore the effects of explanatory variables on the means of estimated parameters and possibly leading to biased estimates (Mannering et al., 2016; Washington et al., 2020). Therefore, it is necessary to explore the effect of weather on the delay duration caused by traffic incidents using the advanced random parameters hazard-based duration model with heterogeneity in means.

Further, to provide additional insights into how weather affects delay severity, we examine the impact of weather conditions on multiple levels of delay caused by traffic incidents. Since the delay severity is a discrete output variable, it can be modeled analogously to the crash injury severity. In general, discrete choice models, such as logit or probit models, are widely used to investigate injury severities (Ghasemzadeh and Ahmed, 2019; Theofilatos, 2017; Ye and Lord, 2014; Yu and Abdel-Aty, 2014a,b). In recent years, a proliferation of studies has considered the effect of heterogeneity and allowed parameter estimates to be different across observations. Thus, random parameters are introduced (Chen et al., 2019; Gong et al., 2022; Hosseinzadeh et al., 2021; Lee et al., 2021; Wang et al., 2020). Moreover, to account for the effect of explanatory variables on the random parameters and provide more insight, we apply random parameters with heterogeneity in means and variances (Ijaz et al., 2022; Li et al., 2021; Washington et al., 2020; Yu et al., 2020; Zamani et al., 2021). Noteworthy, few authors, to the authors' knowledge, have used the random parameters model with heterogeneity in means to analyze the operational risk levels after traffic incidents, such as the incident-induced delay severity.

To sum up, evidence suggests that the existing studies appear to have the following gaps.

First, few researchers have considered operation risk levels caused by traffic incidents, especially the impact of real-time weather factors. However, studying the risk level of operation after a traffic incident could also be relevant, which help traffic managers develop improved incident response countermeasures.

Second, existing studies generally used traditional hazard-based duration models to investigate time-related traffic crash data, ignoring the effect of unobserved multilayer heterogeneity, which leads to biased estimates and not practically indicative results (Mannering et al., 2016).

Third, the traffic operational risk level is a discrete output variable that can be studied with discrete choice models. However, to the authors' knowledge, no researcher has employed advanced discrete choice models to study this problem.

Accordingly, we summarize the contributions of this research as follows.

(1) This study focuses on the impact of weather factors on traffic delays caused by traffic incidents, which will provide technical support for real-time risk control and incident emergency response under different weather conditions.
(2) We use a hazard-based duration model that considers unobserved heterogeneity to study the multilayer heterogeneous influence of real-time weather factors on traffic delay durations, which is new to the literature.
(3) We classify the delay severity into four types and use random parameters logit model with heterogeneity in means approach (i.e., RPLHM) to provide additional insights into how weather conditions affect different delay severities, which is an innovation for the literature.

The traffic incident-induced delay and weather dataset in New York State, US, from Jan to Dec 2020, are selected (Moosavi et al., 2019a,b). Firstly, we consider the incident-induced delay as a continuous duration variable and develop a random parameter hazard-based duration model with heterogeneity in means (i.e., RPHDHM) to investigate the effects of potential weather factors on traffic delay duration. Furthermore, to help policymakers to establish a comprehensive differentiating security policy, a random parameter logit model with heterogeneity in means approach (i.e., RPLHM) is introduced to determine the relationship between weather conditions and the delay severity (i.e., minor delay, moderate delay, severe delay, and extreme delay). Finally, considering the difference in traffic and human characteristics on weekdays and weekends, day and night, the models related to the day of week and time of day are used to examine further the relationship between the weather conditions and traffic delay severity. The detailed framework of this paper is displayed in Fig. 1.

The next structure of the current study is organized as follows: Section 2 describes the data used for this study, followed by Section 3 on the methodological approach. Section 4 gives the model results. And the last section describes the conclusions and discussions of this study and the outlook for future research directions.

2. Data

2.1. Source of the data

The dataset is a publicly available crash database (Moosavi et al., 2019a, b), which covers 49 states of the US. The accident data are collected from February 2016 to Dec 2021, using multiple APIs that provide streaming traffic incident (or event) data. Currently, there are about 2.8 million accident records in this dataset.

In this study, we selected the latest data from January 1 and December 31, 2020, a total of 3440 traffic incidents in the State of New York. Traffic incident information includes the date and start-end time, delay-severity,[1] and weather-related variables. Fig. 2 shows the proportion of various delay severity and the number of accidents in each month, respectively.

2.2. Data processing

Temperature, wind, visibility, and precipitation are widely considered in the relevant studies about the weather-delay relationship (Chen and Wang, 2019; Giang et al., 2014; Schuldt et al., 2021). In this study, two kinds of real-time weather variables are extracted from the weather data, including basic and dummy weather variables. This section elaborates on the data processing for these weather variables.

(1) Basic weather variables

We select temperature (°C), visibility (km), humidity (%), wind speed (m/s), and precipitation (mm) as weather-related continuous variables. Fig. 3 shows the distribution of some weather variables at different delay severities.

(2) Dummy weather variables

Based on the study of Bi et al. (2022), we classify the continuous weather variables in the present study as follows,

Firstly, for temperature, we classify the temperature values into four levels, namely cold (under 0 °C), normal (between 0 and 20 °C), hot (between 20 and 30 °C), and torrid (over 30 °C). As shown in Table 2, the normal temperature has the highest percentage at 73.1 %, followed by hot days at 21.5 %. For visibility, there are also four levels, bad (under 5 km), poor (between 5 and 10 km), normal (between 10 and 20 km), and unlimited (over 20 km). In terms of precipitation, it is divided into 0 mm, between 0 and 1 mm, and over 1 mm, where the proportion of no precipitation is 92.5 %. Finally, we divide the wind speed by the unit interval of 2 m/s

[1] Note that delay-severity refers to the traffic incident-induced delay-severity, expressed as the number between I and IV, where type I - IV indicates minor delay, moderate delay, severe delay, and extreme delay, respectively.

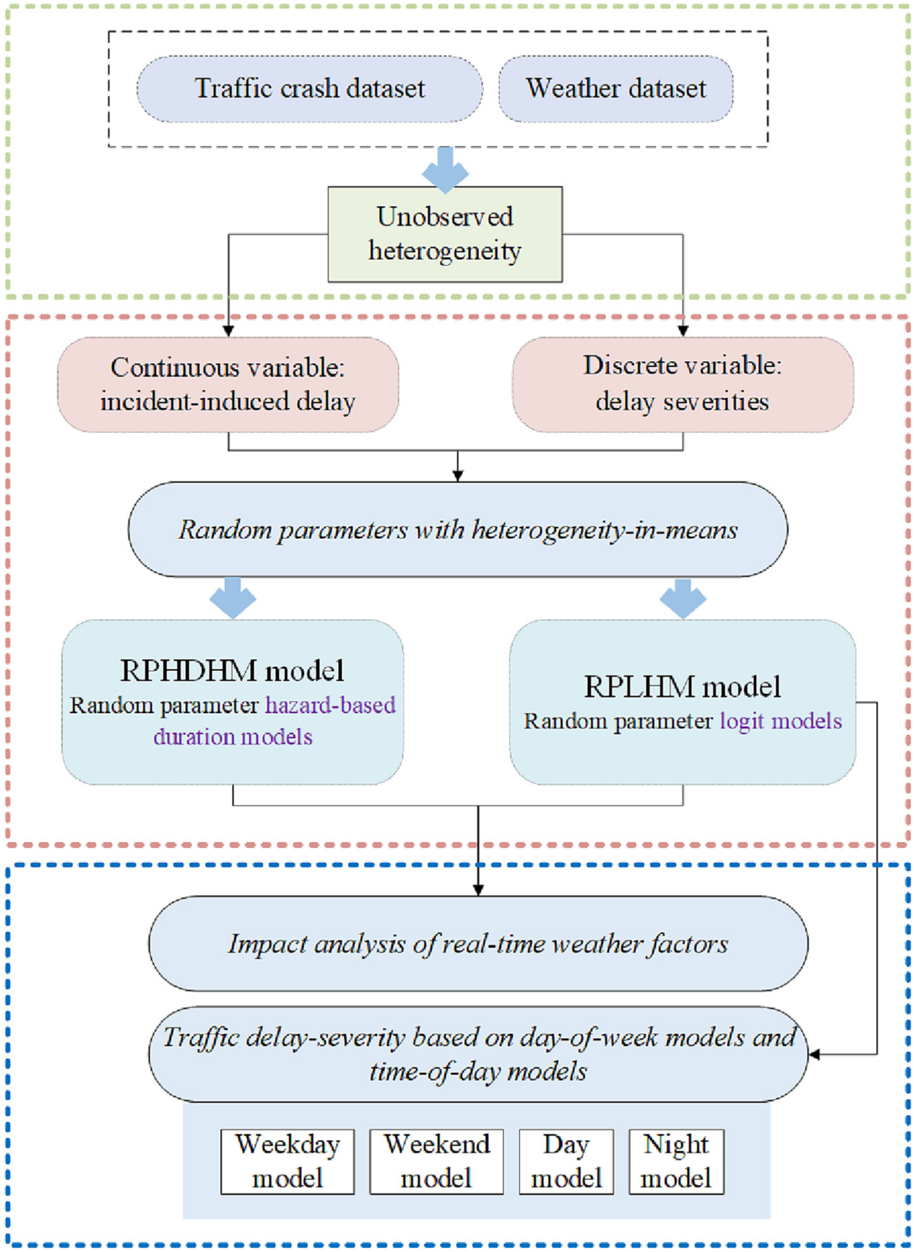


Fig. 1. The study framework layout.

and define a strong breeze as a wind speed exceeding 8 m/s, in line with the study of Wu and Liao (2020). Table 2 shows details of the descriptive statistics of each variable. Further, Fig. 4 respectively shows the proportion and number of the four types of delay-severity among the temperature, wind speed, visibility, precipitation, and sky conditions dummy variables. The figure shows that the proportion of moderate delay is the highest among all the weather dummy variables. In addition, the proportion of extreme delay in cold weather (under 0 °C) is the highest.

3. Methodology

3.1. Random parameters hazard-based duration models with heterogeneity in means

Firstly, we consider the incident-induced delay, the time from the beginning to the end of the delay, as a duration variable. The random parameter hazard-based duration model with heterogeneity in means approach (i.e., RPHDHM) is introduced to determine the relationship between critical influencing factors (such as temperature, wind speed, visibility, and precipitation) and the duration of traffic delay.

To begin with, the cumulative distribution function *F(t)* is the probability that the duration of the incident-induced delay does not exceed a given moment *t*, which can be expressed as (Washington et al., 2020):

$$F(t) = P\{T \le t\} = \int_0^t f(u)du \tag{1}$$

where *T*, as a continuous nonnegative random variable, represents the delay duration caused by a traffic incident. The probability density function *f(t)* of the random variable *T* is the derivative of *F(t)* and can be denoted as follows:

$$f(t) = \frac{dF(t)}{dt} = \lim_{\Delta t} \frac{P\{t < T < t + \Delta t\}}{\Delta t} \tag{2}$$

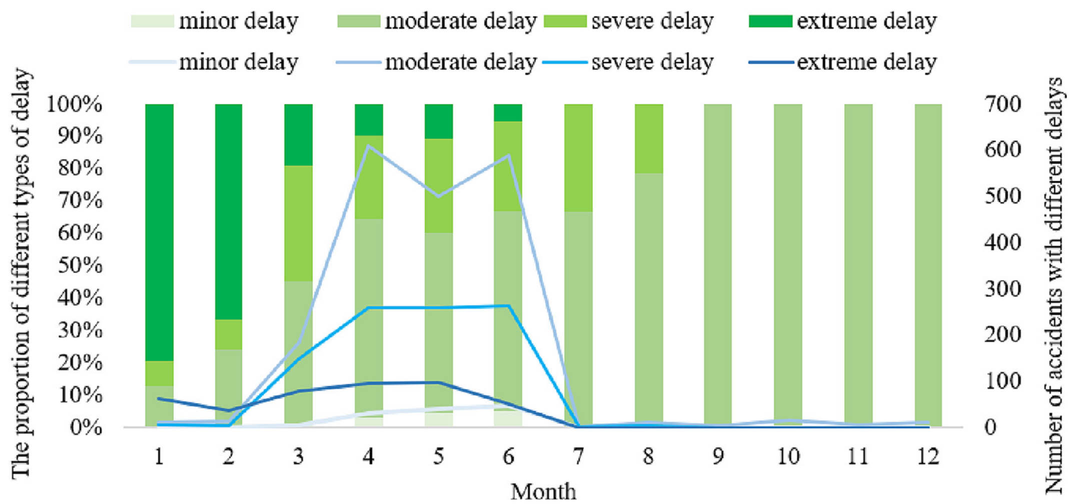


Fig. 2. The statistics of monthly traffic incidents in 2020.

The survival function $S(t)$ specifies the probability that the delay duration is longer than the given moment t and is presented as follows:

$$S(t) = 1 - F(\mathrm{t}) = P\{T > t\} = \int_{\mathrm{t}}^{+\infty} f(u)du \tag{3}$$

The hazard function represents the conditional probability that the delay does not end before the moment t but ends within $(t, t + \Delta t)$, where Δt is a very short time interval. The hazard function $h(t)$ is expressed as follows:

$$h(t) = \lim_{\Delta t \to 0} \frac{P\{t < T < t + \Delta t | T \geq t\}}{\Delta t} = \frac{f(\mathrm{t})}{S(\mathrm{t})} \tag{4}$$

Several approaches exist to incorporate explanatory variables in the hazard-based duration models, including semi- and fully-parametric models. The latter allows us to make assumptions for the distribution of the duration and the functional form.

Therefore, a fully parametric AFT (accelerated failure time) model is employed in this study. The natural logarithm of the delay duration, log T, is expressed as a linear function of the covariates, with the basic form shown below:

$$\log T = \beta' X + \varepsilon \tag{5}$$

where X is a vector of covariates, β is a vector of coefficients, and ε is the error term.

Various parametric distributions have been applied in full parametric models, such as the exponential distribution, the Weibull distribution, and the log-normal distribution (Washington et al., 2020). Among these, the Weibull distribution is flexible because it allows positive or negative duration dependence, i.e., it enables the modeling of data whose probability of the end of duration rises or falls monotonically over time (Ali et al., 2022; Pang et al., 2022; Washington et al., 2020). Therefore, this study is modeled based on the Weibull AFT model with the following hazard function (Washington et al., 2020):

$$h(t) = (\lambda P)(\lambda t)^{P-1} \tag{6}$$

where P is the shape parameter and $P > 0$; λ is the scale parameter and $\lambda > 0$.

3.2. Random parameters logit models with heterogeneity in means

To further identify the influencing factors of delay severities, four delay-severity categories are determined as discrete outcome variables: type-I, type-II, type-III, and type-IV (Moosavi et al., 2019a,b). The random parameters logit model with heterogeneity in means (i.e., RPLHM) is used to study the effect of weather variables on delay severity.

A traffic delay-severity function, Y_{in}, that determines the delay-severity i in crash n, is specified as follows (Ahmed et al., 2021; Fountas et al., 2021, 2018; Washington et al., 2020):

$$Y_{in} = \beta_i X_{in} + \varepsilon_{in} \tag{7}$$

where X_{in} are vectors of explanatory variables that affect traffic delay-severity i (I, II, III, and IV) in crash n, β_i is a vector of corresponding estimable parameters, and ε_{in} is an error term that is assumed to follow an independent and identical distribution with zero mean and variance σ^2.

Further, to account for the unobserved heterogeneity, we introduce random parameters with heterogeneity in means to our model, which not only

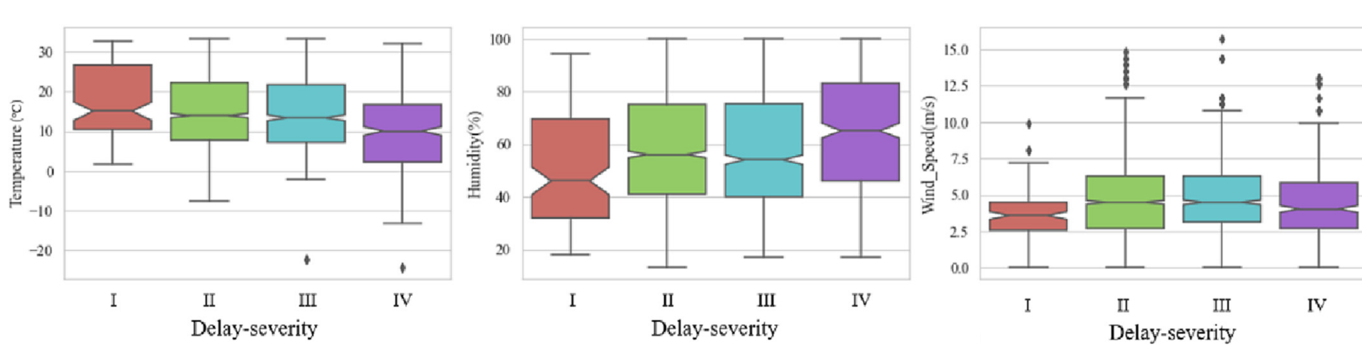


Fig. 3. The distribution of weather variables on the severity of traffic delay.

Table 2
Descriptive statistics of variables.

Variable	Mean	Std.	Min	Max
Severity	2.484	0.753	1	4
Distance (km)	0.289	0.933	0	15.9
Duration (min)	50.325	13.801	9	60
Continuous variables				
Temperature (°C)	14.190	9.000	−24	33
Humidity (%)	58.127	21.012	13	100
Wind speed (m/s)	4.622	2.625	0	16
Visibility (km)	15.257	3.396	0	32
Precipitation (mm)	0.078	0.459	0	7.37
Dummy variables				
Temperature (°C)				
<0 °C	0.029	0.167	0	1
0–20 °C	0.731	0.444	0	1
20–30 °C	0.215	0.411	0	1
>30 °C	0.026	0.160	0	1
Wind speed (m/s)				
0 m/s	0.066	0.249	0	1
0–2 m/s	0.049	0.217	0	1
2–4 m/s	0.287	0.452	0	1
4–6 m/s	0.308	0.462	0	1
6–8 m/s	0.176	0.381	0	1
>8 m/s	0.113	0.317	0	1
Visibility (km)				
<5 km	0.072	0.259	0	1
5–10 km	0.032	0.176	0	1
10–20 km	0.872	0.334	0	1
>20 km	0.023	0.151	0	1
Precipitation (mm)				
0 mm	0.925	0.263	0	1
0–1 mm	0.052	0.222	0	1
>1 mm	0.023	0.150	0	1
Sky conditions				
Clear	0.276	0.447	0	1
Cloud	0.588	0.492	0	1
Rain	0.099	0.299	0	1
Snow	0.026	0.159	0	1
Fog	0.011	0.103	0	1
Season				
Spring	0.670	0.470	0	1
Summer	0.282	0.450	0	1
Autumn	0.007	0.083	0	1
Winter	0.041	0.199	0	1
Week				
Weekday	0.813	0.390	0	1
Weekend	0.187	0.390	0	1
Time				
Morning peak (6–9 a.m.)	0.224	0.417	0	1
Evening peak (5–8 p.m.)	0.297	0.457	0	1
Off-peak	0.480	0.500	0	1
Day	0.873	0.333	0	1
Night	0.127	0.333	0	1

allows the parameter estimates to vary across observations but also enables us to observe the effect of the explanatory variables on the means of observation-specific parameter estimates. The random parameters can be expressed as follows (Washington et al., 2020):

$$\beta_n = \beta + \Theta Z_n + \xi_n \tag{8}$$

where β is the vector of mean parameter estimate across all observations, $Z_n$is a vector of explanatory variables from observation n that influence the mean of β_n, Θ is a vector of estimable parameters and $\xi_n$is a vector of randomly distributed terms.

The simulated maximum likelihood method was applied for model estimation (Train, 2009), and 1000 Halton draws were used to achieve stable parameter estimates (McFadden and Train, 2000). In terms of the distribution of the random parameters, the normal distribution is used to achieve the best goodness of fit (Behnood and Mannering, 2019; Fountas et al., 2018, 2021).

4. Result analysis

Firstly, fixed and random parameters models are also estimated. Each hazard-based duration framework and multinomial logit framework is evaluated by comparing the Akaike Information Criterion (AIC) value, the McFadden R-Squared, and the log-likelihood value at convergence. Smaller AIC values, higher McFadden R-Squared values, and higher log-likelihood values at convergence indicate a better model fit (Washington et al., 2020). The goodness-of-fit of both models suggests that the random parameters models with heterogeneity-in-means are superior (see Table 3).

Therefore, in the following subsections, we investigate how weather-related factors affect the delay duration and severity based on the estimates of these two superior models (see Tables 4–5 for RPHDHM and RPLHM, respectively). Specifically, Subsection 4.1 explains both model results of random parameters and heterogeneity-in-means. Subsection 4.2 explores the effect of weather variables on delay durations and levels after traffic incidents. And Subsection 4.3 discusses the effects of weather factors on the delay severity under different times (i.e., Weekday/Weekend, Day/Night).

4.1. Random parameters with heterogeneity-in-means results

For the RPHDHM model (see Table 4), four explanatory variables produce statistically significant standard deviations and heterogeneity in means, including cold temperature (under 0 °C), hot temperature (between 20 and 30 °C), fog, and weekend.

Specifically, as shown in Fig. 5 (a), the cold temperature produces a normally distributed random parameter following $N(-0.143, 0.222^2)$, demonstrating that traffic incident delay time is lower in 74.03 % of the crashes that occurred under cold weather. The presence of a lower percentage of traffic volume on cold days compared to warm days is the possible reason behind this. Moreover, the normal visibility (between 10 and 20 km) and evening peak-hour variables reduce the mean of the cold temperature variable, indicating that long delay durations are less likely to occur under cold and evening peak hours. Conversely, a strong breeze (over 8 m/s) increases its mean, which suggests long delays usually happen on windy days.

Similarly, hot temperature (between 20 and 30 °C) results in a random parameter, with a low probability of delay for the vast majority of the observations (99.71 %). However, strong breeze, normal visibility, and evening peak hours increase the mean of hot temperatures, making longer delays more likely. In contrast, precipitation (over 1 mm) decreases the mean value under hot temperatures, indicating that longer delays are less likely to happen in hot, rainy weather.

In terms of foggy weather, while the duration of delays is reduced by 29.60 %, a strong breeze significantly increases its mean, which means that windy conditions are associated with longer delays.

As shown in Fig. 5 (b), the weekend indicator produces a significant random parameter, with 79.68 % of traffic incidents decreasing delay durations. The evening peak reduces weekends' mean value, while precipitation increases its mean, implying shorter delays during weekend evening rush hours and longer ones on rainy weekends.

For the RPLHM model results in Table 6, two variables have statistically significant random parameters: hot temperature and normal visibility.

Concretely, the probability of minor delay after traffic accidents rises dramatically on hot days, up to 344.03 %. Calm conditions (wind speed = 0 m/s) reduce the mean of hot temperatures, indicating that the probability of minor delay in windless weather is lower.

The estimated parameter of normal visibility obeys a normal distribution, with a mean of −0.216 and a standard deviation of 0.111. It means 83.47 % of the likelihood of extreme delay is reduced (see Fig. 6). Humidity decreases the mean of normal visibility, while rainfall increases its mean, illustrating that the probability of extreme delay is slightly lower in wet weather and higher in rainy weather.

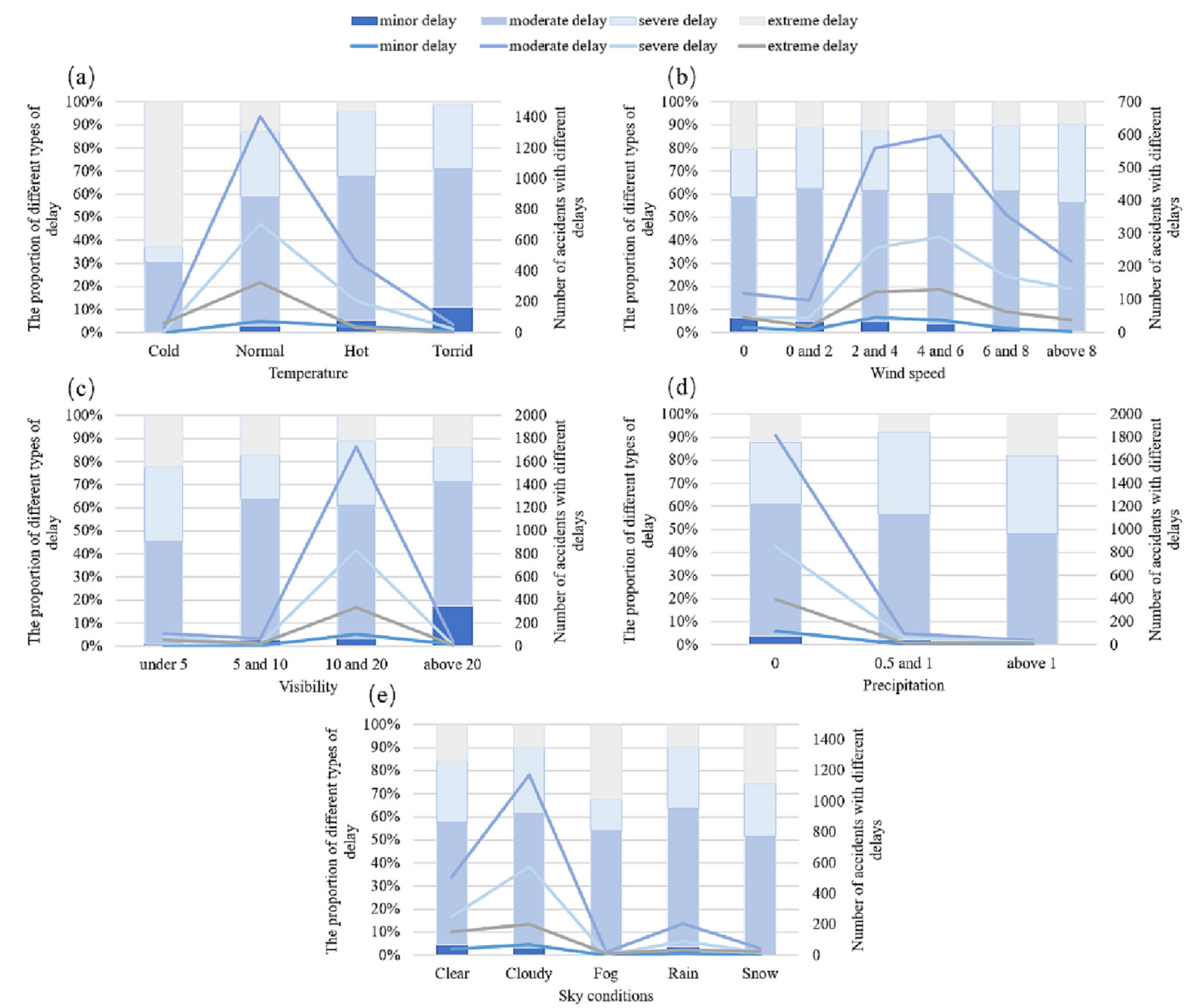


Fig. 4. The statistics of traffic incidents in dummy weather variables.

4.2. Impact analysis of real-time weather factors

(1) Temperature

The temperature has a significant effect on traffic delays. Specifically, cold days (under 0 °C) and hot days (between 20 and 30 °C) reduce the duration of delays by 13.32 % and 33.37 %, respectively (see Table 4). This negative association can be explained by previous studies showing a significant decrease in travel demand under high and low temperature conditions, resulting in lower traffic volumes and accident possibilities, thus reducing delays (Roh et al., 2013).

Remarkably, most of the existing literature focuses on analyzing the impact of a single weather factor on traffic conditions. However, this study introduces heterogeneity in means, thus providing more insights by exploring the effects of multiple weather factor interactions on traffic delays in depth. For example, a strong breeze (over 8 m/s) is found to have a significant positive interaction effect with cold days (under 0 °C), and hot days (between

Table 3

Goodness-of-fit measures for the estimated models.

Goodness-of-fit measures	Hazard-based duration model			Logit model		
	HD	RPHD	RPHDHM	MNL	RPL	RPLHM
Number of observations	3440	3440	3440	3440	3440	3440
No. of parameters	20	26	36	29	31	35
Log-likelihood at convergence	−180.27	−175.95	−155.93	−3350.7	−3342.3	−3338.0
McFadden R-Squared	0.134	0.174	0.268	0.076	0.299	0.300
Akaike information criterion (AIC)	400.54	403.90	383.86	6759.40	6746.60	6746.00

MNL: Multinomial logit model; RPL: Random parameters logit model; RPLHM: Random parameters logit model with heterogeneity-in-means. HD: Hazard-based duration model; RPHD: Random parameters hazard-based duration model; RPHDHM: Random parameters hazard-based duration model with heterogeneity-in-means.

Table 4
Estimation results of significant variables for RPHDHM model.

Variable	Parameter estimate	t-stat.	Hazard ratio
Constant	3.839***	153.19	
Continuous variables			
Humidity (%)	0.028***	2.76	2.84 %
Dummy variables			
Temperature (°C)			
<0 °C	−0.143***	−4.46	−13.32 %
Random parameters (normally distributed)	0.222***	9.57	–
Heterogeneity in the means of the random parameters			
<0 °C: >8 m/s	0.021***	5.24	–
<0 °C: 10–20 km	−0.231*	−1.69	–
<0 °C: Evening peak	−0.260***	−4.28	–
0–20 °C	0.030*	1.81	3.05 %
20–30 °C	−0.406***	−4.96	−33.37 %
Random parameters (normally distributed)	0.147***	4.19	
Heterogeneity in the means of the random parameters			
20–30 °C: >8 m/s	0.012**	2.47	–
20–30 °C: 10–20 km	0.312*	1.78	–
20–30 °C: >1 mm	−0.329*	−1.89	–
20–30 °C: Evening peak	0.159*	1.84	–
Wind speed (m/s)			
>8 m/s	0.003***	3.22	0.30 %
Visibility (km)			
<5 km	0.007***	4.68	0.70 %
10–20 km	−0.031*	−1.67	−3.05 %
Precipitation (mm)			
>1 mm	0.051***	4.32	5.23 %
Sky conditions			
Clear	−0.040***	−3.89	−3.92 %
Snow	0.056*	1.83	5.76 %
Fog	−0.351***	−7.09	−29.60 %
Random parameters (normally distributed)	0.075**	2.23	–
Heterogeneity in the means of the random parameters			
Fog: >8 m/s	0.014***	3.91	–
Week			
Weekend	−0.225***	−5.61	−20.15 %
Random parameters (normally distributed)	0.271***	14.74	–
Heterogeneity in the means of the random parameters			
Weekend: >1 mm	0.108***	2.85	–
Weekend: Evening peak	−0.198***	−5.31	–
Time			
Evening peak	−0.034***	−3.16	−3.34 %
Day	0.059***	5.17	6.08 %
Shape parameter	0.184***	66.70	

Note: ***, **, * = => Significance at 0.99, 0.95, and 0.90 level of confidence, respectively.

20 and 30 °C), i.e., the strong breeze significantly increases the mean of cold days and hot days, suggesting that longer delays often happen on windy days with extreme temperatures. The result is well understood: strong winds can affect vehicle controllability, further increasing the likelihood of crashes, which in turn can lead to severe congestion (Hou et al., 2018). Moreover, normal visibility (between 10 and 20 km) and evening peaks also have a significant positive interaction effect with hot days (between 20 and 30 °C), increasing its mean and making the probability of longer delays more likely. This result indicates that longer delay duration generally happens during hot, clear evening peak hours. It may be because drivers are exhausted and crashes are more frequent during the hot evening rush hour, resulting in long delays (Cabrera-Arnau et al., 2020).

(2) Wind

Strong breeze (wind speed over 8 m/s) significantly increases the duration of traffic delays and the probability of severe delay. The conclusion is intuitively easy to understand: the strong breeze will increase the difficulty

Table 5
Estimation results of significant variables for the RPLHM model.

Variable	Parameter estimate	t-stat.	Pseudo-elasticities (%)[a]			
			I	II	III	IV
[II] Constant	6.048***	6.04				
[III] Constant	6.778***	6.54				
[IV] Constant	7.509***	7.33				
Continuous variables						
[III] Humidity (%)	−0.006**	−2.46	8.74	9.08	−23.97	7.37
Dummy variables						
Temperature (°C)						
[II] < 0 °C	−0.327***	−3.46	4.01	−3.46	4.18	3.91
[I] 0–20 °C	0.711***	2.76	10.68	−0.61	−0.61	−0.51
[II] 0–20 °C	0.335***	2.74	−3.39	2.02	−3.62	−2.93
[IV] 0–20 °C	1.782***	6.76	−2.82	−2.93	−2.93	21.96
[I] 20–30 °C	0.232***	4.01	**344.03**	−13.50	−13.50	−111.6
Random parameters (normally distributed)	0.064*	1.83				
Heterogeneity in the means of the random parameters						
20–30 °C: 0 m/s	−0.001*	−1.73				
[III] 20–30 °C	0.995**	2.29	−4.43	−4.53	**10.05**	−2.92
[IV] 20–30 °C	3.630***	6.72	−5.74	−5.81	−5.81	**30.67**
[I] > 30 °C	1.015**	2.29	2.14	−0.26	−0.26	−0.22
[IV] > 30 °C	2.089***	4.12	−1.29	−1.31	−1.31	**18.97**
Wind speed (m/s)						
[I] 0 m/s	−0.261***	−4.97	−113.55	3.16	3.16	2.59
[III] > 8 m/s	0.206*	1.71	−0.79	−0.80	**1.53**	−0.67
Visibility (km)						
[II] < 5 km	0.081***	5.42	−69.24	**51.15**	−73.11	−57.78
[IV] < 5 km	1.781**	2.36	−0.52	−0.53	−0.53	**13.39**
[IV] 10–20 km	−0.216***	−5.80	12.00	12.50	12.50	**−119.17**
Random parameters (normally distributed)	0.111***	4.13				
Heterogeneity in the means of the random parameters						
10–20 km: Rain	0.053**	2.12				
10–20 km: Humidity	−0.001*	−1.79				
Precipitation (mm)						
[III] 0 mm	−0.721***	−4.08	17.53	18.35	−48.36	14.51
[II] > 1 mm	0.820*	1.91	−6.66	**5.22**	−6.80	−4.41
Sky conditions						
[II] Clear	−1.084***	−5.13	15.59	−13.32	16.65	12.24
[II] Cloud	−0.789***	−4.66	26.20	−18.93	27.43	21.98
[I] Rain	1.971	1.34	**18.28**	−0.66	−0.66	−0.59
[III] Fog	−1.077**	−2.07	0.16	0.16	−1.00	0.13
Week						
[IV] Weekday	−0.0390**	−2.42	2.92	2.98	2.98	**−22.92**
Time						
[I] Morning peak	−0.888***	−2.79	−18.87	0.41	0.41	0.34
[I] Day	1.235**	2.39	**98.97**	−3.96	−3.96	−3.25
[IV] Day	−1.608***	−9.47	9.94	10.22	10.22	**−104.02**

Note: 1) ***, **, * = => Significance at 0.99, 0.95, and 0.90 level of confidence, respectively.
2) I - minor delay, II - moderate delay, III- severe delay, and IV - extreme delay.
[a] The pseudo-elasticities quantify the change in outcome probability when an explanatory variable changes from "0" to "1" (Washington et al., 2020).

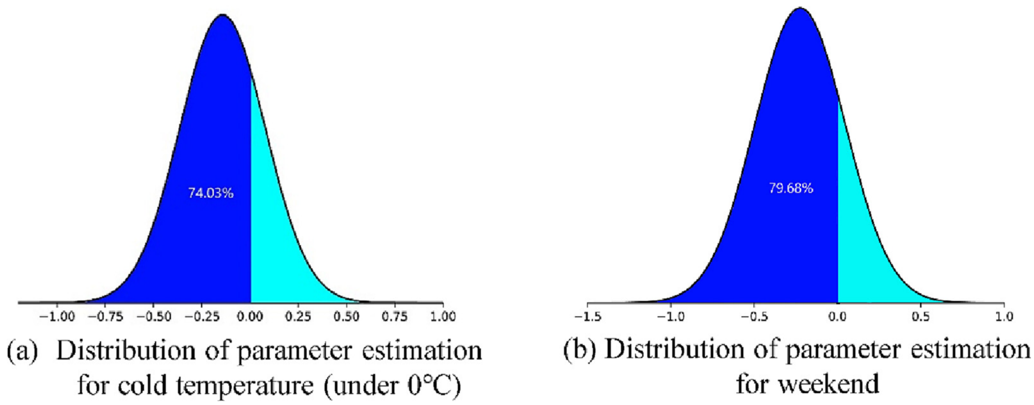


Fig. 5. Distribution of parameters estimation for representing variables in the RPHDHM model.[a]
[a]According to Table 4, hot temperature (between 20 °C and 30 °C) and fog make 99.71 % and 99.99 % of the incident-induced delays lower, respectively.

of handling large vehicles and reduce vehicle stability, increasing the frequency of rollover crashes, which induces severe traffic delays (Hou et al., 2018; Young and Liesman, 2007). Moreover, the strong breeze can blow up dust on the road, thus affecting visibility (FHWA, 2022).

An interesting finding of this paper is that there is a significant interaction between the strong breeze and foggy days, thus making the probability of longer delays more likely. It provides a new idea for road safety management by producing a joint warning system for different severe weather conditions. Precisely, real-time speed limit values should be measured when two or more severe weather conditions are present at the same time, in this case, a joint warning for strong winds and foggy weather. Concretely, real-time speed limit values should be recalculated when two or more severe weather conditions exist simultaneously. And traffic management departments should set up real-time speed limit warning signs by the roadside, which also broadcast the road conditions ahead to provide drivers with more accurate driving guidance.

Table 6
The estimation results of significant variables for the Weekday model based on the RPLHM model.

Variable	Parameter estimate	t-stat.	I	II	III	IV
[II] Constant	5.056***	7.18				
[III] Constant	5.351***	6.99				
[IV] Constant	4.716***	6.53				
Continuous variables						
[III] Humidity (%)	−0.005*	−1.94	7.63	7.63	−19.49	5.55
Dummy variables						
Temperature (°C)						
[II] < 0 °C	−0.383***	−3.76	4.78	−4.15	4.78	4.20
[I] 0–20 °C	0.643**	2.45	9.81	−0.57	−0.57	−0.40
[II] 0–20 °C	0.374***	2.73	−3.95	2.07	−3.95	−2.76
[IV] 0–20 °C	1.723***	5.55	−2.31	−2.31	−2.31	17.81
[I] 20–30 °C	0.197***	4.83	289.85	−12.7	−12.70	−8.94
[III] 20–30 °C	0.236*	1.83	−1.12	−1.12	2.39	−0.58
[IV] 20–30 °C	2.911***	6.78	−3.79	−3.79	−3.79	20.48
[IV] > 30 °C	1.688***	3.41	−0.91	−0.91	−0.91	13.23
Wind speed (m/s)						
[I] 0 m/s	−0.197***	−4.44	−90.94	2.81	2.81	2.00
Visibility (km)						
[II] < 5 km	0.053***	3.04	−48.17	33.25	−48.17	−33.98
[IV] 10–20 km	−0.302***	−6.62	9.92	9.92	9.92	−100.77
Random parameters (normally distributed)	0.152***	5.26				
Heterogeneity in the means of the random parameters						
10–20 km: Rain	0.060**	1.96				
Precipitation (mm)						
[III] 0 mm	−0.698***	−3.33	18.49	18.49	−46.84	13.15
Sky conditions						
[II] Clear	−1.111***	−4.53	17.01	−13.29	17.01	11.02
[III] Clear	−0.334**	−2.02	2.64	2.64	−6.47	1.72
[II] Cloud	−0.712***	−3.49	26.15	−17.54	26.15	18.93
[I] Rain	1.209***	2.96	9.23	−0.41	−0.41	−0.30
[III] Fog	−0.883*	−1.67	0.17	0.17	−0.94	0.12
Time						
[I] Morning peak	−0.970***	−3.06	−21.25	0.42	0.42	0.30

Note: 1) ***, **, * = => Significance at 0.99, 0.95, and 0.90 level of confidence, respectively.
2) I - minor delay, II - moderate delay, III- severe delay, and IV - extreme delay.

(3) Precipitation

The daily precipitation over 1 mm first increases the duration of traffic delays by 5.23 % and the probability of moderate delay by 5.22 %. Similarly, Giang et al. (2014) also confirmed that rainfall causes 8.6 % longer travel time compared to no precipitation. It might be due to the fact that precipitation can cause slippery roads and reduced visibility, leading to crashes and triggering delays (FHWA, 2022). Consequently, in rainy or snowy weather, drivers should be reminded to drive carefully and slowly and encouraged to install anti-skid chains on their tires if necessary to reduce traffic accidents caused by slipping.

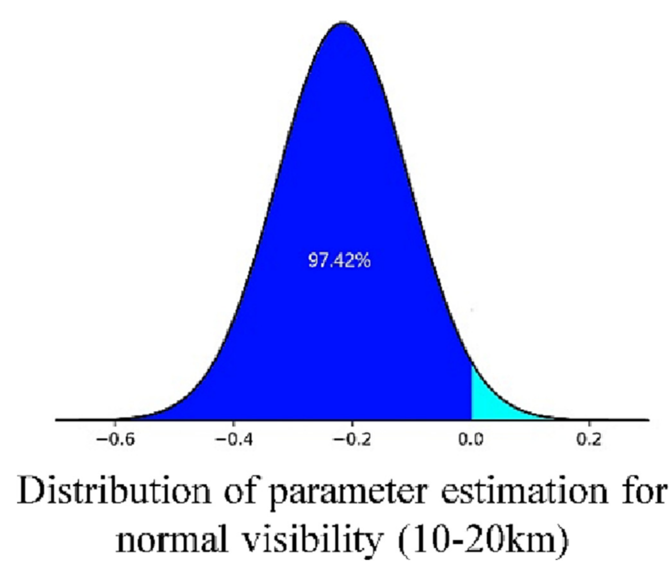


Fig. 6. Distribution of parameters estimation for representing variables in the RPLHM model.

It is worth noting that precipitation increases the mean of the weekend indicator, and precipitation during the weekend makes the likelihood of longer delays more likely. In addition, the higher the relative humidity, the long delays are caused. Conversely, higher relative humidity results in a 23.97 % reduction in the probability of severe delay in the RPLHM model.

(4) Visibility

Low visibility conditions cause serious crashes and, consequently, traffic delays by causing, for example, increased speed variance (Abdel-Aty et al., 2011; Hassan and Abdel-Aty, 2013; Yu and Abdel-Aty, 2014b). According to Table 4, low visibility (under 5 km) significantly increases the traffic delay duration, while normal visibility (between 10 and 20 km) decreases the traffic delay duration by 3.05 %. In terms of the delay severity, low visibility (under 5 km) increases the probability of moderate delay and extreme delay by 51.15 % and 13.39 %, respectively, while normal visibility (between 10 and 20 km) reduces the likelihood of extreme delay by 119.17 %. A potential interpretation is that drivers who suddenly encounter relatively high traffic density under lower visibility will be too late to reduce speed, resulting in crashes such as rear-end collisions (Abdel-Aty et al., 2012). Therefore, drivers must comply with the applicable speed limits in low visibility conditions.

Sky conditions also reflect the effects of visibility. The clear condition reduces the duration of traffic delays by 3.92 % and the probability of moderate delay by 13.32 %. This result is consistent with common sense, as drivers have a better view in clear weather and are less likely to have an accident, hence lower delay levels. Rain conditions increase the likelihood of minor delays by 18.28 %, and snow conditions increase the duration of traffic delays by 5.76 %.

Finally, the weekday reduces the probability of extreme delay by 22.92 %, and the day indicator also reduces the likelihood of extreme delay by 104.02 % but increases the probability of minor delay by 98.97 %. According to Abdel-Aty et al. (2011), drivers may have a clearer perception of the external environment during daytime compared to nighttime, which reduces the likelihood of traffic accidents and then eases delays.

Table 7
The estimation results of significant variables for the Weekend model based on the RPLHM model.

Variable	Parameter estimate	t-stat.	I	II	III	IV
[II] Constant	7.518***	4.78				
[III] Constant	9.867***	5.96				
[IV] Constant	8.220***	5.02				
Dummy variables						
Temperature (°C)						
[I] 0–20 °C	1.024**	2.09	19.27	−1.27	−1.27	−1.27
[IV] 0–20 °C	1.315***	3.95	−7.16	−7.16	−7.16	16.56
[I] 20–30 °C	0.429***	4.53	619.65	−21.89	−21.89	−21.89
[III] 20–30 °C	−1.347*	−1.73	0.42	0.42	−3.77	0.42
[IV] 20–30 °C	−2.334***	−2.72	0.73	0.73	0.73	−40.35
Wind speed (m/s)						
[III] 0 m/s	0.129***	3.04	−14.44	−14.44	37.14	−14.44
[III] 6–8 m/s	−0.671**	−2.02	1.77	1.77	−6.88	1.77
Visibility (km)						
[II] < 5 km	0.157***	4.96	−132.29	102.30	−132.29	−132.29
[IV] 10–20 km	2.215**	2.54	−1.38	−1.38	−1.38	3.44
Precipitation (mm)						
[III] 0 mm	−1.021**	−2.36	21.27	21.27	−68.41	21.27
[II] 0.5–1 mm	1.231***	2.63	−5.55	2.49	−5.55	−5.55
Sky conditions						
[IV] Clear	0.646**	2.36	−4.12	−4.12	−4.12	14.78
[II] Rain	0.934***	3.32	−11.04	6.25	−11.04	−11.04
[III] Snow	−1.582***	−2.61	1.72	1.72	−5.41	1.72
Time						
[IV] Morning peak	−0.787*	−1.93	1.10	1.10	1.10	−16.52
[IV] Night	1.509***	5.34	−12.91	−12.91	−12.91	15.72
[IV] Summer	−0.723**	−2.14	1.80	1.80	1.80	−21.38
[IV] Winter	1.267***	2.76	−4.73	−4.73	−4.73	1.97

Note: 1) ***, **, * ==> Significance at 0.99, 0.95, and 0.90 level of confidence, respectively.
2) I - minor delay, II - moderate delay, III- severe delay, and IV - extreme delay.

4.3. Analysis of traffic delay-severity based on day-of-week models and time-of-day models

Traffic and human characteristics may vary by day of week and time of day; day-of-week models (i.e., Day model and Night model) and time-of-day models (i.e., Weekday model and Weekend model) are used to examine further the relationship between the weather conditions and traffic delay-severity caused by traffic incidents.

Tables 6–7 show estimation results separately based on Weekday and Weekend models. We found that more significant weather-related variables are observed in the Weekday model than in the Weekend model, indicating that the overall effect of weather conditions on weekdays is more sensitive and stronger than that on weekends. The same indicator also impacted the traffic delay severity differently between weekdays and weekends. Specifically, for example, the probability of extreme delays during hot weekdays (between 20 and 30 °C) increases by 20.48 %. This finding is exciting and reasonable because people need to commute on weekdays. And they tend to travel by car during muggy weather to improve their comfort, which increases the traffic density and consequently leads to extreme delays (Badshah et al., 2022).

Conversely, the probability of extreme delays on hot weekends decreases by 40.35 %, which could be attributed to the fact that people generally reduce unnecessary trips during hot and muggy weather, decreasing traffic and easing congestion. During clear, windless weekends (i.e., wind speed = 0 m/s), people's travel demand increases, and the probability of severe delays increases by 14.78 % and 37.14 %, respectively. Moreover, the probability of extreme delays increases by 15.72 % at night on weekends. The potential reason for this could be that most public transportation stops operating at night. Therefore, people have to take cabs or private cars to return home, and the darkness at night makes drivers less aware of their surroundings, which increases the traffic flow and the possibility of crashes and induces extreme congestion and delay. Thus, one suggestion is to install intelligent heat-sensitive street lights to illuminate the road, reduce drivers' blind spots, and thus improve road safety.

To further verify the transferability of influencing factors, this study adopts the out-of-sample simulation, namely, adopting the Weekend model to predict Weekday data. Then the prediction accuracy is finally obtained by comparing the difference with the predicted probability of their influencing factors. It should be noted that these out-of-sample forecasts do not simply use the mean of the random parameters, which would result in obviously biased predictions. For details regarding this technique and how to interpret the results, a reader may refer to recent studies on injury severity (Alnawmasi and Mannering, 2022; Alogaili and Mannering, 2022, 2020; Hou et al., 2022; Islam et al., 2020; Se et al., 2022).

Firstly, we used the Weekend model to predict Weekday data. The result of this out-of-sample simulation is presented in Fig. 7. Specifically, minor and severe delay predictions are underestimated by 0.0002 and 0.0001, respectively. Moderate and extreme delay predictions are overestimated by 0.0001 and 0.0002, respectively. Generally, it seems that the influence factors of traffic delay levels in the Weekend model can be used to predict the Weekday data. We also found that many individuals showed prediction precision with a significant deviation; namely, the mean value of prediction accuracy is likely to be the positive and negative balance between the high estimate and the low estimate in individual prediction. Therefore, judging the transferability of influencing factors only from the forecast mean value is defective. In order to express individual differences in the

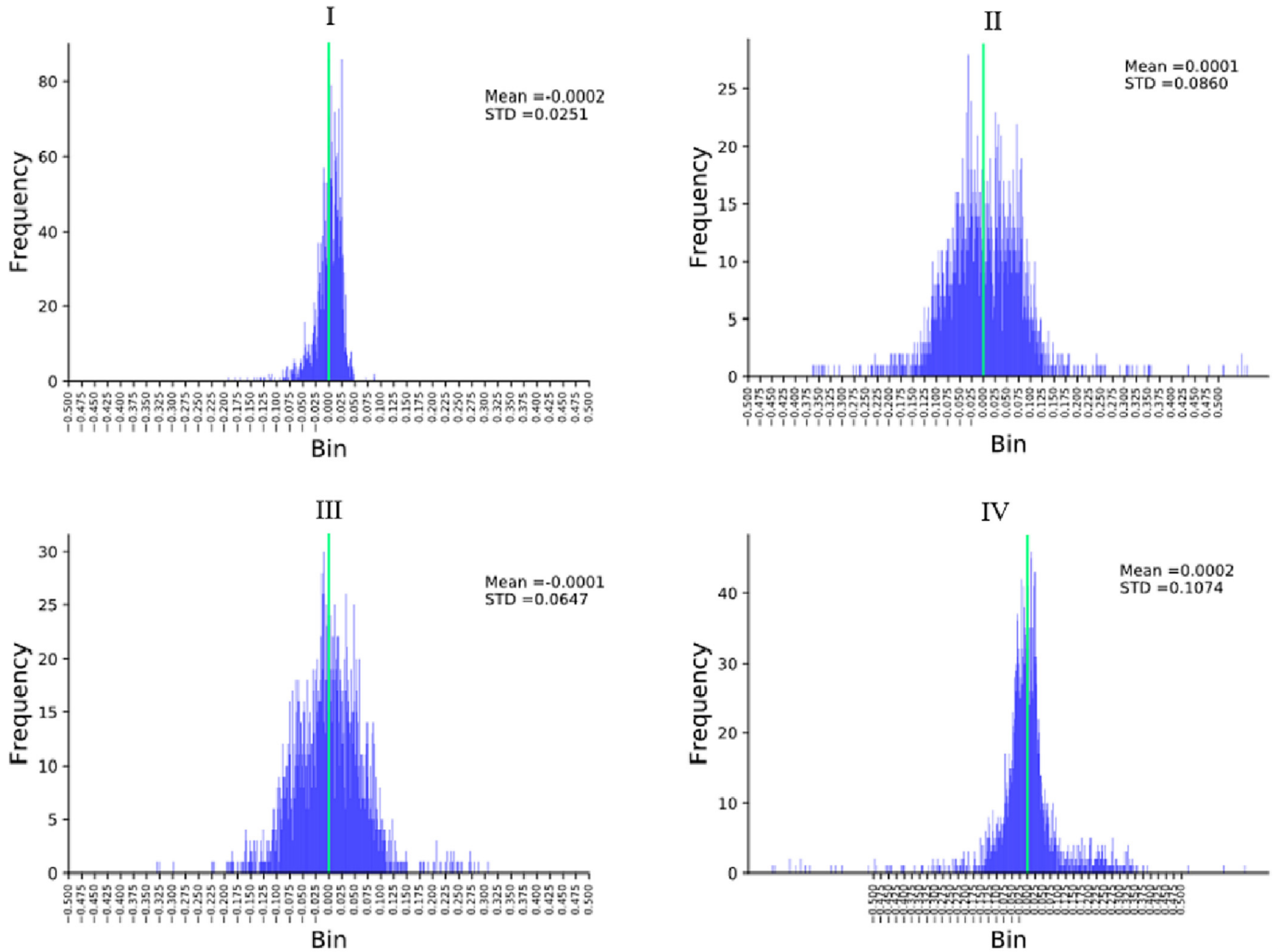


Fig. 7. Using the Weekend model to predict Weekday data.

forecasting process more intuitively, this study constructed the frequency distribution map to show individual prediction accuracy. See Fig. 7 for details.

Tables 8–9 show estimation results based on the Day and Night models. This study also finds that more weather-related variables are observed in the Day model than in the Night model. The same significant factor also significantly affects the impact of day and night traffic delay severity. For example, in the Night model, the probability of minor delay increases significantly by 73.81 % for cold (under 0 °C) weather conditions, while in the day model, the probability of moderate delay increases by 11.93 % and 359.42 % for the normal day (between 0 and 20 °C) and hot day (between 20 and 30 °C), respectively. The no-wind condition increases the probability of severe delay by 41.42 % in the Night model but reduces the probability of minor delay by 112.11 % in the Day model. In addition, the morning peak increases the probability of severe delay by 10.74 % and decreases the probability of minor delay by 112.11 % in the Night model. However, the morning peak decreases the probability of minor delay in the Day model by 25.91 %.

Secondly, we used the Night model to predict day's data. The result of this out-of-sample simulation is presented in Fig. 8. Specifically, minor delay, moderate delay, and severe delay predictions are overestimated by 0.0003, 0.0017, and 0.0011, respectively. The extreme delay predictions are underestimated by 0.0032. It is worth noting that other studies also illustrated significant underestimation and overestimation in the out-of-sample predictions (Alnawmasi and Mannering, 2022; Alogaili and Mannering, 2022; Wang et al., 2022). However, Yan et al. (2022) concluded that the prediction deviation is relatively small based on random parameters models. The overall differences in the integration of the sample size and contributing variables might cause the contradictory phenomenon. Please see Mannering (2018) and Mannering et al. (2020) for a detailed discussion of the issues.

5. Conclusions and discussion

Using traffic delay data that occurred between January 1 and December 31, 2020, in the state of New York, the present study explores deeply the relationship between weather conditions and traffic delay caused by traffic incidents, while the potential influencing factors, including temperature, wind speed, visibility, and precipitation are statistically analyzed. To that end, to account for multiple layers of unobserved heterogeneity, a random parameters hazard-based duration model with heterogeneity in means approach is introduced to determine the relationship between key influencing factors and the duration of traffic delay. Further, following the original dataset reports, the delay-severity outcomes are categorized as follows: type-I delay (i.e., minor delay), type-II delay (i.e., moderate delay), type-III delay (i.e., severe delay), and type-IV delay (i.e., extreme delay), a random parameters logit model with heterogeneity in means approach is introduced to determine the relationship between weather conditions and the delay-severity. The key findings are summarized as follows:

(1) Wind speed, temperature, and visibility significantly impact the incident-induced delay levels, and their impact varies across the severity of delays. Specifically, strong breeze (wind speed over 8 m/s), precipitation over 1 mm, and low visibility (visibility under 5 km)

Table 8
The estimation results of significant variables for the Day model based on the RPLHM model.

Variable	Parameter estimate	t-stat.	I	II	III	IV
[II] Constant	5.380***	5.46				
[III] Constant	5.619***	5.59				
[IV] Constant	4.865***	4.84				
Dummy variables						
Temperature (°C)						
[II] < 0 °C	−0.267***	−2.63	3.16	−2.48	3.37	2.75
[I] 0–20 °C	0.805***	2.95	11.93	−0.76	−0.76	−0.48
[II] 0–20 °C	0.398***	3.11	−4.22	−4.22	−4.69	−2.75
[IV] 0–20 °C	1.558***	4.41	−1.41	−1.49	−1.49	15.03
[I] 20–30 °C	0.196***	3.24	359.42	−15.16	−15.16	−9.48
Random parameters (normally distributed)	0.076**	2.10				
[III] 20–30 °C	0.912**	2.06	−4.11	−4.27	9.24	−1.70
[IV] 20–30 °C	3.771***	5.45	−3.48	−3.50	−3.50	20.92
[I] > 30 °C	1.144**	2.45	2.66	−0.31	−0.31	−0.16
[IV] > 30 °C	1.442**	2.57	−0.61	−0.61	−0.61	8.59
Wind speed (m/s)						
[I] 0 m/s	−0.255***	−4.38	−112.11	3.39	3.39	2.08
Visibility (km)						
[II] < 5 km	0.077***	4.82	−67.35	44.44	−73.11	−42.68
[IV] 10–20 km	−0.369***	−5.24	5.41	5.88	5.88	−62.40
Random parameters (normally distributed)	0.215***	4.91				
Heterogeneity in the means of the random parameters						
10–20 km: Rain	0.073**	2.27				
Precipitation (mm)						
[III] 0 mm	−0.777***	−4.18	19.07	20.47	−51.25	12.03
[II] > 1 mm	0.809*	1.84	−6.96	4.74	−7.24	−2.89
Sky conditions						
[II] Clear	−1.714***	−6.75	24.34	−18.83	27.18	13.46
[III] Clear	−0.628***	−3.29	4.42	4.89	−11.96	2.43
[II] Cloud	−0.987***	−5.28	34.29	−22.11	36.60	21.97
Time						
[I] Morning peak	−1.081***	−3.29	−25.91	0.57	0.57	0.36

Note: 1) ***, **, * ==> Significance at 0.99, 0.95, and 0.90 level of confidence, respectively.
2) I - minor delay, II - moderate delay, III- severe delay, and IV - extreme delay.

significantly affect the duration of delay; a hot day (between 20 and 30 °C) has a 344.03 % greater probability of being minor delay. Strong breeze has a higher likelihood of severe delay, and precipitation over 1 mm increases the probability of moderate delay. The low visibility is found to increase the estimated odds of moderate delay and severe delay by 51.15 % and 13.39 %, respectively. In comparison, the normal visibility (between 10 and 20 km) significantly decreases the estimated odds of severe delay by 119.17 %.

(2) The RPHDHM and RPLHM models provide superior statistical fit and offer additional insights by accommodating variations of the explanatory variables across the observations and factors affecting the means of the parameter density functions of the random parameters. For the RPHDHM model, there are four statistically significant variables as random parameters, including the cold temperature (under 0 °C), the hot temperature (between 20 and 30 °C), the fog weather, and the weekend. For the RPLHM model, there are two statistically significant variables as random parameters, including the hot temperature (between 20 and 30 °C) and the normal visibility (between 10 and 20 km).

(3) The out-of-sample predictions undertaken in this study further confirm non-transferability by adopting Weekend model parameters to predict the data of the Weekday and using the Night model to predict the day's data. The overall differences in the integration of the sample size, statistics of traffic delay levels, and contributing variables might cause the contradictory phenomenon.

Table 9
The estimation results of significant variables for the Night model based on the RPLHM model.

Variable	Parameter estimate	t-stat.	I	II	III	IV
[II] Constant	5.148***	5.02				
Random parameters (normally distributed)	2.271*	1.73				
[III] Constant	4.287***	3.54				
[IV] Constant	5.211***	4.99				
Continuous variables						
[III] Humidity (%)	−0.025***	−2.97	48.42	21.16	−132.94	48.42
Dummy variables						
Temperature (°C)						
[III] < 0 °C	−1.308*	−1.74	1.01	0.52	−16.36	1.01
[I] < 0 °C	2.328**	2.06	73.81	−1.24	−2.89	−2.89
[IV] 0–20 °C	1.223***	2.75	−8.29	−3.69	−8.29	8.23
[IV] 20–30 °C	0.935*	1.85	−5.07	−1.93	−5.07	7.56
Wind speed (m/s)						
[III] 0 m/s	0.190***	2.91	−24.82	−11.38	41.42	−24.82
Visibility (km)						
[IV] 10–20 km	−0.096***	−2.83	28.22	10.75	28.22	−51.43
Precipitation (mm)						
[IV] 0.5–1 mm	1.299*	1.68	−1.78	−0.83	2.68	−1.78
Time						
[III] Evening peak	0.925**	2.48	−7.46	−3.71	10.74	−7.46
[III] Spring	1.478***	4.05	−34.87	−15.56	58.81	−34.87
[I] Winter	−1.897**	−2.24	4.33	−17.18	4.33	4.33

Note: 1) ***, **, * ==> Significance at 0.99, 0.95, and 0.90 level of confidence, respectively.
2) I - minor delay, II - moderate delay, III- severe delay, and IV - extreme delay.

(4) The findings from this analysis also offer a number of practical implications. First, in addition to exploring the duration of traffic delay, we examine the impact of weather conditions on multiple levels of delay caused by traffic incidents. Accurate identification of weather-related influencing factors that correspond to different delay severity is expected to help policymakers to establish a comprehensive differentiating security policy to resolve traffic congestion, e.g., installing intelligent heat-sensitive street lights to illuminate the road and reduce drivers' blind spots. Second, findings from the study indicate that weather-related variables significantly affect traffic delay; more importantly, the interactions between variables (obtained by random parameters with heterogeneity in means approach) are found to increase the delay duration, and accordingly, a road safety management system with joint warnings for multiple severe weather conditions need to be proposed. Third, more weather-related variables are observed in the Weekday model than in the Weekend model. So, integrating diverse driver behaviors of weekday and weekend perspectives into consideration at all stages of resolved policy.

This study also has some limitations. Firstly, due to the lack of crash characteristics in original datasets, this study only explores the relationship between weather conditions and traffic delays caused by traffic incidents, so more comprehensive data can be obtained for further research. In addition, future work in this area can explore the spatial transferability of the proposed models across states and compare the differences among states. It is noted that the utilization of Intelligent Transportation Systems (ITS) easing traffic congestion, a good understanding of what are the Critical Success Factors to support ITS in reducing traffic jams will be potential future research (Çaldağ and Gökalp, 2020). Moreover, the effect of weather on traffic safety during the COVID-19 lockdown requires further exploration. With the advent of COVID-19, a unique and unprecedented period of slower traffic occurred due to the outbreak control policy. Existing research found that crash frequency decreased during the earlier "Lockdown" period while severity increased (Sekadakis et al., 2021; Dong et al., 2022; Shaik and Ahmed, 2022). There also has study to investigate its impacts on traffic

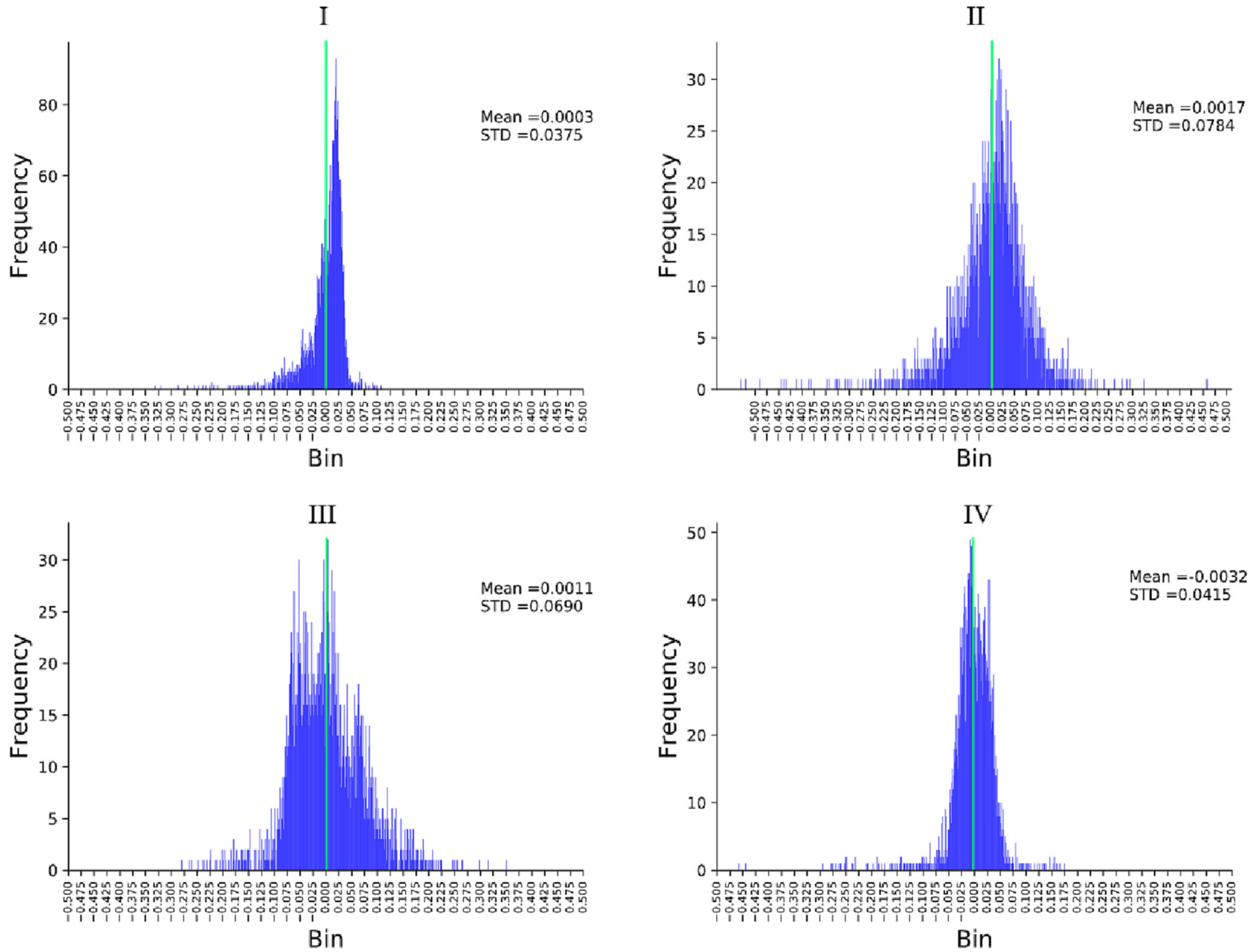


Fig. 8. Using the Night model to predict Day data.

safety during the later stage of the pandemic (Gong et al., 2023). They discovered that the crash frequency was significantly lower throughout the pandemic. However, it significantly increases during the later stage due to the relaxed restrictions. In future work, longer periods of data to study the impact of weather on traffic safety from the COVID-19 "Lockdown" to the "New Normal" would be a promising direction.

CRediT authorship contribution statement

Xiangtong Su: Conceptualization, Formal analysis, Methodology, Project administration, Validation, Writing – original draft, Writing – review & editing. **Danyue Zhi:** Formal analysis, Investigation, Methodology, Validation, Funding acquisition, Resources, Supervision, Visualization, Writing – original draft. **Dongdong Song:** Conceptualization, Supervision, Data curation, Writing – review & editing. **Le Tian:** Investigation, Data curation. **Yitao Yang:** Software, Validation, Writing – review & editing.

Data availability

Data will be made available on request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank Dr. Sobhan Moosavi for providing the data used in this study. The authors also thank the anonymous reviewers for their comments and suggestions that helped improve the manuscript.

References

Abdel-Aty, M., Ekram, A.-A., Huang, H., Choi, K., 2011. A study on crashes related to visibility obstruction due to fog and smoke. Accid. Anal. Prev. 43, 1730–1737.

Abdel-Aty, M.A., Hassan, H.M., Ahmed, M., Al-Ghamdi, A.S., 2012. Real-time prediction of visibility related crashes. Transport. Res. Part C: Emerging Technol. 24, 288–298.

Ahmed, S.S., Cohen, J., Anastasopoulos, P.Ch., 2021. A correlated random parameters with heterogeneity in means approach of deer-vehicle collisions and resulting injury-severities. Analytic Methods in Accident Research. 30.

Ali, Y., Haque, Md.M., Zheng, Z., Afghari, A.P., 2022. A Bayesian correlated grouped random parameters duration model with heterogeneity in the means for understanding braking behaviour in a connected environment. Analytic Methods in Accident Research. 35.

Alkaabi, A.M.S., Dissanayake, D., Bird, R., 2011. Analyzing clearance time of urban traffic accidents in Abu Dhabi, United Arab Emirates, with Hazard-based duration modeling method. Transp. Res. Rec. 2229, 46–54.

Alnawmasi, N., Mannering, F., 2022. The impact of higher speed limits on the frequency and severity of freeway crashes: accounting for temporal shifts and unobserved heterogeneity. Analytic Methods in Accident Research. 34.

Alogaili, A., Mannering, F., 2022. Differences between day and night pedestrian-injury severities: accounting for temporal and unobserved effects in prediction. Analytic Methods in Accident Research. 33.

Alogaili, A., Mannering, F., 2020. Unobserved heterogeneity and the effects of driver nationality on crash injury severities in Saudi Arabia. Accid. Anal. Prev. 144, 105618.

Badshah, I., Khan, Z.H., Gulliver, T.A., Khattak, K.S., Saad, S., 2022. Modeling sustainable traffic behavior: avoiding congestion at a stationary bottleneck. Civ. Eng. J. 8 (11), 2378–2390.

Behnood, A., Mannering, F., 2019. Time-of-day variations and temporal instability of factors affecting injury severities in large-truck crashes. Anal. Methods Accid. Res. 23, 100102.

Bi, H., Ye, Z., Zhu, H., 2022. Data-driven analysis of weather impacts on urban traffic conditions at the city level. Urban Clim. 41, 101065.

Cabrera-Arnau, C., Prieto Curiel, R., Bishop, S.R., 2020. Uncovering the behaviour of road accidents in urban areas. R. Soc. Open Sci. 7 (4), 191739.

Çaldağ, M.T., Gökalp, E., 2020. Exploring critical success factors for blockchain-based intelligent transportation systems. Emerg. Sci. J. 4, 27–44.

Chen, F., Song, M., Ma, X., 2019. Investigation on the injury severity of drivers in rear-end collisions between cars using a random parameters bivariate ordered probit model. Int. J. Environ. Res. Public Health 16, 2632.

Chen, Z., Wang, Y., 2019. Impacts of severe weather events on high-speed rail and aviation delays. Transp. Res. Part D: Transp. Environ. 69, 168–183.

Dimitriou, L., Vlahogianni, E.I., 2015. Fuzzy modeling of freeway accident duration with rainfall and traffic flow interactions. Analytic Methods in Accident Research. 5–6, pp. 59–71.

Dong, X., Xie, K., Yang, H., 2022. How did COVID-19 impact driving behaviors and crash Severity? A multigroup structural equation modeling. Accid. Anal. Prev. 172, 106687.

FHWA, 2022. How Do Weather Events Impact Roads? - FHWA Road Weather Management. https://ops.fhwa.dot.gov/weather/q1_roadimpact.htm.

Fountas, G., Anastasopoulos, P.Ch., Abdel-Aty, M., 2018. Analysis of accident injury-severities using a correlated random parameters ordered probit approach with time variant covariates. Analytic Methods in Accident Research. 18, pp. 57–68.

Fountas, G., Fonzone, A., Olowosegun, A., McTigue, C., 2021. Addressing unobserved heterogeneity in the analysis of bicycle crash injuries in Scotland: a correlated random parameters ordered probit approach with heterogeneity in means. Analytic Methods in Accident Research. 32.

Ghasemzadeh, A., Ahmed, M.M., 2019. Exploring factors contributing to injury severity at work zones considering adverse weather conditions. IATSS Res. 43, 131–138.

Giang, W.C.W., Donmez, B., Ahghari, M., MacDonald, R.D., 2014. The impact of precipitation on land interfacility transport times. Prehosp. Disaster Med. 29, 593–599.

Gong, H., Fu, T., Sun, Y., Guo, Z., Cong, L., Hu, W., Ling, Z., 2022. Two-vehicle driver-injury severity: a multivariate random parameters logit approach. Analytic Methods in Accident Research. 33.

Gong, Y., Lu, P., Yang, X.T., 2023. Impact of COVID-19 on traffic safety from the "lockdown" to the "new normal": a case study of Utah. Accid. Anal. Prev. 106995.

Hassan, H.M., Abdel-Aty, M.A., 2013. Predicting reduced visibility related crashes on freeways using real-time traffic flow data. J. Saf. Res. 45, 29–36.

Hojati, A.T., Ferreira, L., Washington, S., Charles, P., 2013. Hazard based models for freeway traffic incident duration. Accid. Anal. Prev. 52, 171–181.

Hosseinzadeh, A., Moeinaddini, A., Ghasemzadeh, A., 2021. Investigating factors affecting severity of large truck-involved crashes: comparison of the SVM and random parameter logit model. J. Saf. Res. 77, 151–160.

Hou, Q., Huo, X., Leng, J., Mannering, F., 2022. A note on out-of-sample prediction, marginal effects computations, and temporal testing with random parameters crash-injury severity models. Analytic Methods in Accident Research. 33.

Hou, Q., Tarko, A.P., Meng, X., 2018. Investigating factors of crash frequency with random effects and random parameters models: new insights from Chinese freeway study. Accid. Anal. Prev. 120, 1–12.

Ijaz, M., Lan, L., Usman, S.M., Zahid, M., Jamal, A., 2022. Investigation of factors influencing motorcyclist injury severity using random parameters logit model with heterogeneity in means and variances. Int. J. Crashworthiness 27, 1412–1422.

Islam, M., Alnawmasi, N., Mannering, F., 2020. Unobserved heterogeneity and temporal instability in the analysis of work-zone crash-injury severities. Analytic Methods in Accident Research. 28.

Islam, N., Adanu, E.K., Hainen, A.M., Burdette, S., Smith, R., Jones, S., 2022. Evaluating the impact of freeway service patrol on incident clearance times: a spatial transferability test. J. Adv. Transp. 2022, 1–10.

Javid, R., 2018. A framework for travel time variability analysis using urban traffic incident data. IATSS Res. 42, 30–38.

Koetse, M.J., Rietveld, P., 2009. The impact of climate change and weather on transport: an overview of empirical findings. Transp. Res. Part D: Transp. Environ. 17.

Lee, J., Mao, S., Abdel-Aty, M., Fu, W., 2021. Use of bivariate random-parameter probit model to analyze the injury severity of highway traffic crashes involving school-age children. Transp. Res. Rec. 2675, 530–537.

Li, R., Shang, P., 2014. Incident duration modeling using flexible parametric Hazard-based models. Comput. Intell. Neurosci. 2014, 1–10.

Li, Y., Song, L., Fan, W.(David), 2021. Day-of-the-week variations and temporal instability of factors influencing pedestrian injury severity in pedestrian-vehicle crashes: a random parameters logit approach with heterogeneity in means and variances. Analytic Methods in Accident Research. 29, p. 100152.

Liu, A., Soneja, S.I., Jiang, C., Huang, C., Kerns, T., Beck, K., Mitchell, C., Sapkota, A., 2017. Frequency of extreme weather events and increased risk of motor vehicle collision in Maryland. Sci. Total Environ. 580, 550–555.

Mannering, F., 2018. Temporal instability and the analysis of highway accident data. Analytic Methods in Accident Research. 17, pp. 1–13.

Mannering, F., Bhat, C.R., Shankar, V., Abdel-Aty, M., 2020. Big data, traditional data and the tradeoffs between prediction and causality in highway-safety analysis. Analytic Methods in Accident Research. 25.

Mannering, F.L., Shankar, V., Bhat, C.R., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Analytic Methods in Accident Research. 11, pp. 1–16.

McFadden, D., Train, K., 2000. Mixed MNL models for discrete response. J. Appl. Econ. 15, 447–470.

Moosavi, S., Samavatian, M.H., Parthasarathy, S., Ramnath, R., 2019. A Countrywide Traffic Accident Dataset.

Moosavi, S., Samavatian, M.H., Parthasarathy, S., Teodorescu, R., Ramnath, R., 2019. Accident risk prediction based on heterogeneous sparse data: new dataset and insights. Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, SIGSPATIAL '19. Association for Computing Machinery, New York, NY, USA, pp. 33–42.

Pang, J., Krathaus, A., Benedyk, I., Ahmed, S.S., Anastasopoulos, P.Ch, 2022. A temporal instability analysis of environmental factors affecting accident occurrences during snow events: The random parameters hazard-based duration model with means and variances heterogeneity. Analytic Methods in Accident Research. 34, p. 100215.

Roh, H.J., Datla, S., Sharma, S., 2013. Effect of snow, temperature and their interaction on highway truck traffic. J. Transport. Technol. 3, 24–38.

Schuldt, S.J., Nicholson, M.R., Adams, Y.A., Delorit, J.D., 2021. Weather-related construction delays in a changing climate: a systematic state-of-the-art review. Sustainability 13, 2861.

Se, C., Champahom, T., Jomnonkwao, S., Kronprasert, N., Ratanavaraha, V., 2022. The impact of weekday, weekend, and holiday crashes on motorcyclist injury severities: Accounting for temporal influence with unobserved effect and insights from out-of-sample prediction. Analytic Methods in Accident Research. 36, p. 100240.

Seeherman, J., Liu, Y., 2015. Effects of extraordinary snowfall on traffic safety. Accid. Anal. Prev. 81, 194–203.

Sekadakis, M., Katrakazas, C., Michelaraki, E., Kehagia, F., Yannis, G., 2021. Analysis of the impact of COVID-19 on collisions, fatalities and injuries using time series forecasting: the case of Greece. Accid. Anal. Prev. 162, 106391.

Shaik, M.E., Ahmed, S., 2022. An overview of the impact of COVID-19 on road traffic safety and travel behavior. TransportationEngineering 100119.

Theofilatos, A., 2017. Incorporating real-time traffic and weather data to explore road accident likelihood and severity in urban arterials. J. Saf. Res. 61, 9–21.

Train, K.E.E., 2009. Discrete Choice Methods With Simulation. Cambridge University Press, Cambridge.

Transport Canada, 2021. National Collisions Database Online, National Collision Database Online 1.0 - Transport Canada. https://wwwapps2.tc.gc.ca/Saf-Sec-Sur/7/NCDB-BNDC/p.aspx?l=en (accessed 11.14.2022).

Tsapakis, I., Cheng, T., Bolbol, A., 2013. Impact of weather conditions on macroscopic urban travel times. J. Transp. Geogr. 28, 204–211.

Vlahogianni, E.I., Karlaftis, M.G., 2013. Fuzzy-entropy neural network freeway incident duration modeling with single and competing uncertainties. Comput. Aided Civ. Inf. Eng. 28, 420–433.

Wang, C., Ijaz, M., Chen, F., Zhang, Y., Cheng, J., Zahid, M., 2022. Evaluating gender differences in injury severities of non-helmet wearing motorcyclists: accommodating temporal shifts and unobserved heterogeneity. Analytic Methods in Accident Research. 36.

Wang, J., Huang, H., Xu, P., Xie, S., Wong, S.C., 2020. Random parameter probit models to analyze pedestrian red-light violations and injury severity in pedestrian–motor vehicle crashes at signalized crossings. J. Transport. Saf. Secur. 12, 818–837.

Washington, S., Karlaftis, M., Mannering, F., Anastasopoulos, P., 2020. Statistical and econometric methods for transportation data analysis. Chapman and Hall/CRC.

Wu, J., Liao, H., 2020. Weather, travel mode choice, and impacts on subway ridership in Beijing. Transp. Res. A Policy Pract. 135, 264–279.

Xu, C., Tarko, A.P., Wang, W., Liu, P., 2013a. Predicting crash likelihood and severity on freeways with real-time loop detector data. Accid. Anal. Prev. 57, 30–39.

Xu, C., Wang, W., Liu, P., 2013b. Identifying crash-prone traffic conditions under different weather on freeways. J. Saf. Res. 46, 135–144.

Yan, X., He, J., Wu, G., Zhang, C., Wang, C., Ye, Y., 2022. Differences of overturned and hit-fixed-object crashes on rural roads accompanied by speeding driving: accommodating potential temporal shifts. Analytic Methods in Accident Research. 35.

Ye, F., Lord, D., 2014. Comparing three commonly used crash severity models on sample size requirements: multinomial logit, ordered probit and mixed logit models. Analytic Methods in Accident Research. 1, pp. 72–85.

Young, R.K., Liesman, J., 2007. Estimating the relationship between measured wind speed and overturning truck crashes using a binary logit model. Accid. Anal. Prev. 39 (3), 574–580.

Yu, M., Zheng, C., Ma, C., 2020. Analysis of injury severity of rear-end crashes in work zones: a random parameters approach with heterogeneity in means and variances. Analytic Methods in Accident Research. 27.

Yu, R., Abdel-Aty, M., 2014a. Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data. Saf. Sci. 7.

Yu, R., Abdel-Aty, M., 2014b. Using hierarchical bayesian binary probit models to analyze crash injury severity on high speed facilities with real-time traffic data. Accid. Anal. Prev. 62, 161–167.

Yu, R., Abdel-Aty, M., Ahmed, M., 2013. Bayesian random effect models incorporating real-time weather and traffic data to investigate mountainous freeway hazardous factors. Accid. Anal. Prev. 50, 371–376.

Yu, R., Xiong, Y., Abdel-Aty, M., 2015. A correlated random parameter approach to investigate the effects of weather conditions on crash risk for a mountainous freeway. Transport. Res. Part C: Emerg. Technol. 50, 68–77.

Zamani, A., Behnood, A., Davoodi, S.R., 2021. Temporal stability of pedestrian injury severity in pedestrian-vehicle crashes: new insights from random parameter logit model with heterogeneity in means and variances. Analytic Methods in Accident Research. 32.

Zhan, Z.-Y., Yu, Y.-M., Chen, T.-T., Xu, L.-J., Ou, C.-Q., 2020. Effects of hourly precipitation and temperature on road traffic casualties in Shenzhen, China (2010–2016): a time-stratified case-crossover study. Sci. Total Environ. 720, 137482.