Accident Analysis and Prevention 99 (2017) 321–329

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Predicting drowsy driving in real-time situations: Using an advanced driving simulator, accelerated failure time model, and virtual location-based services

Junhua Wang[a], Shuaiyi Sun[a], Shouen Fang[a,*], Ting Fu[b], Joshua Stipancic[b]

[a] College of Transportation Engineering, Tongji University, China
[b] Department of Civil Engineering and Applied Mechanics, McGill University, Canada

ARTICLE INFO

Article history:
Received 17 July 2016
Received in revised form
24 November 2016
Accepted 20 December 2016
Available online 27 December 2016

Keywords:
Drowsy driving
Duration prediction
Accelerate failure time
Location based service
Drowsy driving warning system

ABSTRACT

This paper aims to both identify the factors affecting driver drowsiness and to develop a real-time drowsy driving probability model based on virtual Location-Based Services (LBS) data obtained using a driving simulator. A driving simulation experiment was designed and conducted using 32 participant drivers. Collected data included the continuous driving time before detection of drowsiness and virtual LBS data related to temperature, time of day, lane width, average travel speed, driving time in heavy traffic, and driving time on different roadway types. Demographic information, such as nap habit, age, gender, and driving experience was also collected through questionnaires distributed to the participants. An Accelerated Failure Time (AFT) model was developed to estimate the driving time before detection of drowsiness. The results of the AFT model showed driving time before drowsiness was longer during the day than at night, and was longer at lower temperatures. Additionally, drivers who identified as having a nap habit were more vulnerable to drowsiness. Generally, higher average travel speeds were correlated to a higher risk of drowsy driving, as were longer periods of low-speed driving in traffic jam conditions. Considering different road types, drivers felt drowsy more quickly on freeways compared to other facilities. The proposed model provides a better understanding of how driver drowsiness is influenced by different environmental and demographic factors. The model can be used to provide real-time data for the LBS-based drowsy driving warning system, improving past methods based only on a fixed driving.

1. Introduction

Drowsy driving is a common threat that endangers the lives of drivers and nearby road users alike. Studies have repeatedly shown that drowsy driving plays a critical role in traffic collisions, disproportionately affecting severe and fatal crashes (Pack et al., 1995; Maycock 1996; Lyznicki et al., 1998; Williamson et al., 2001; Vanlaar et al., 2008; Accidentes 2015; Zhang et al., 2016). Despite this evidence, drivers continue this risky behavior and the number of collisions and fatalities caused by drowsy driving have remained high throughout the past decades. Horne and Reyner (1995) found that 23% of accidents occurring on monotonous motorways were related to drowsy driving. Using a multiple imputation methodology, Tefft (2014) estimated that drowsy driving accounted for 7% of all crashes and 16.5% of fatal crashes in the US from 2009 to 2013. Furthermore, a nationally representative telephone survey in the US found that 41% of drivers admit to having "fallen asleep or nodded off" while driving (Suite 2010). Smith et al. (2005) conducted a 4-week follow-up study finding that young adult drivers felt drowsy while driving in more than 23% of the cases.

Drowsy driving may be caused by limited sleep, long periods of driving, and monotonous environments, among other possibilities (Saccomanno et al., 1970; Arnold and Hartley 1998; Hu et al., 2010). Several countermeasures have been adopted in an attempt to mitigate these causes. In some cases, collisions related to drowsy driving can be avoided by alerting drivers to their potential drowsiness, allowing them to take a break from driving before an incident occurs (Kulmala 1997; Ferdinands 1999; Regan et al., 2001; Young et al., 2003). Maximum continuous driving times have been proposed to allow drivers to take breaks during their trip. Drowsy driving warning systems relying on facial recognition techniques have been developed and used to warn drivers who are falling asleep (Grace and Steward 2001). Meanwhile, warning signs and

* Corresponding author.
E-mail addresses: benwjh@163.com (J. Wang), sshuaiyi@126.com (S. Sun), fangshouen@126.com (S. Fang), ting.fu@mail.mcgill.ca (T. Fu), joshua.stipancic@mail.mcgill.ca (J. Stipancic).

http://dx.doi.org/10.1016/j.aap.2016.12.014

indicators of drowsy driving have been installed in monotonous environments to alert drivers to the possibility of drowsiness. However, such systems may be difficult to implement in practice, as drowsy driving is a gradual behavior which occurs without the drivers' knowledge. Modeling the drowsy driving time, the time spent driving before signs of drowsiness are detected, and understanding the impact of different factors on drowsy driving time will help to understand how built environment and personal habits are related to the probability of drowsy driving. Additionally, facial recognition warning systems can be developed according to the probability of drowsiness, to either better alert drivers to take breaks before dangerous levels of drowsiness set in, or when to stop driving if dangerous levels of drowsiness are reached.

Driving data is difficult to obtain in real-world traffic environments because of the potential danger to participants. This is especially true of drowsy driving studies which require participants to be drowsy in order to collect meaningful data. Selecting proper testing environments may be difficult, as continually changing road environments make it is impossible to isolate specific built environment variables. Therefore, driving simulation experiments are preferred and should be expected. However, because of the challenges in simulating experiences which closely mirror real-world expectations, the reliability of driving simulators in studying drowsy driving behavior has been challenged. Nevertheless, researchers at universities and other institutions have invested time and money in developing and building advanced driving simulators which provide experiences much closer to what would be expected in the field (Slob 2008; Fischer et al., 2010; Lin et al., 2015). With these advanced simulators, drowsing driving in real-world scenarios can be better explored. In this study, a driving simulator was used to collect virtual Location Based Service (LBS), services provided by applications that offer geographic location data and provide rich information of the built environment including geometric road data. This virtual LBS data, including drowsy driving time and other driving information, can be easily collected in the field using LBS devices.

The objective of this study is to build a real-time drowsy driving probability model for predicting continuous driving time before the onset of driver drowsiness, considering different factors such as driver gender, age, experience, and sleeping habits, weather condition and temperature, road type, time of day, and traffic condition. For this purpose, an Accelerated Failure Time (AFT) model was developed based on 258 experimental drowsiness records collected from an advanced driving simulator with 8° of freedom. This model could be used as the basis for a drowsy warning system based on LBS.

2. Literature review

The key task explored in past research of drowsy driving warning systems is the detection and recognition of drowsiness. Based on past literature, approaches in identifying drowsy driving can be classified into two categories: (a) identifying drowsy driving by setting a fixed driving time threshold, with which drowsy driving would be identified if the continuous driving time exceeds such threshold; (b) detecting and identifying drowsy driving through multiple aspects of the driver or the vehicle in addition to the driving time factor, such as cognitive distraction, physiological reaction, facial expression, or vehicle operation condition. Eriksson and Papanikolopoulos (2001) recognized drowsy driving based on the analysis of facial expressions including eye blink times, eye closure duration, and eyelid movement. Yeo et al. (2009) developed a method of Automatic Electroencephalographic (EEG) for detecting drowsiness with the application of a Support Vector Machine (SVM). Wang et al. (2015) also proposed an EEG-based detection method to determine driving drowsiness in real-time by analyzing drivers' neural mechanisms. Such methods, though gaining their attention due to the awareness of the hazards of drowsy driving, have still not been widely utilized and remained to be improved. Considering driving time, as one of the key factors associated to drowsy driving, modeling driving time and use a calibrated driving time threshold from the model potentially helps increase the performance of drowsiness detection. In practice, traditional drowsy driving warning systems based on time spent driving adopt a simple fixed driving time as the upper threshold for alerting drivers. Such thresholds have been investigated in several studies. For example, Yanli et al. (2009) used psychological tests and subjective drowsiness investigation, concluding that continuous driving time should not exceed 3.5 h. Other studies have detected drowsy driving through driver behavior. In one example, Mcdonald et al. (2013) identified drowsy driving based lane departure degree and steering wheel angle.

Several studies have investigated the relationship between driving duration or distance, factors of the built environment or driver, and drowsiness. Friswell and Williamson (2013) studied the impacts of vehicle type and driving distance and duration on drowsiness and found that, on average, drivers of short haul light vehicles felt drowsy after 6 h of driving, while drivers of long distance heavy vehicles felt drowsy after 11 h. Sang and Li (2012) tested the psychological fatigue of bus drivers using a Psychology Fatigue Measurement System. The authors found that drivers' operation capability decreased after 4 h of continuous driving due to drowsiness. Many studies investigated the impact of driving environment on drowsy driving. Pilcher and Huffcutt (1996) identified that complex road environments and traffic conditions made drivers more vulnerable to the effects of drowsiness. However, a study conducted by Liu and Wu (2009) illustrated that complex driving environments did not necessarily increase the likelihood of drowsiness compared to monotonous environment. In their study, both environments induced drowsy symptoms after 60 min of driving. Others have explored the impact of driver working time on their sleep conditions (Jones et al., 2005). Despite the progress made in this past research, these studies fall short to some extent, as drowsy driving is explained by a combination of different factors from the built environment and individual driver. Modeling the impact of these factors on drowsy driving, specifically on the duration/distance of continuous driving before detection of drowsiness, should be considered.

In the past several decades, different methods have been investigated for modeling driving duration. Such models include models based on variance analysis, regression, non-parametric regression, hazard-based methods, decision tree, fuzzy logic, Artificial Neural Network and Bayesian Networks. Among these models, hazard-based duration models have recently gained popularity. Hazard-based duration models study the conditional probability of an event ending given that the event has lasted up until a specified time (Washington et al., 2010). As the time variable is connected with a conditional probability, hazard-based duration modeling allows for the explicit study of the relationship between drowsy driving duration and the explanatory variables. Accelerated Failure Time (AFT) models are effective and easily interpretable parametric approaches that can incorporate the effect of external covariates on the hazard function (Greene 2003), improving prediction for cases with missing data. As a result, the AFT model has been applied extensively in a number of transportation fields such including incident duration (Junhua et al., 2013) and pedestrian waiting duration (Yang et al., 2015).

Driver warning systems with fixed driving durations can be improved with a dynamic driving duration threshold dependent on different driver and built environment factors. LBS are a good source for built environment data, and can provide the necessary environ-

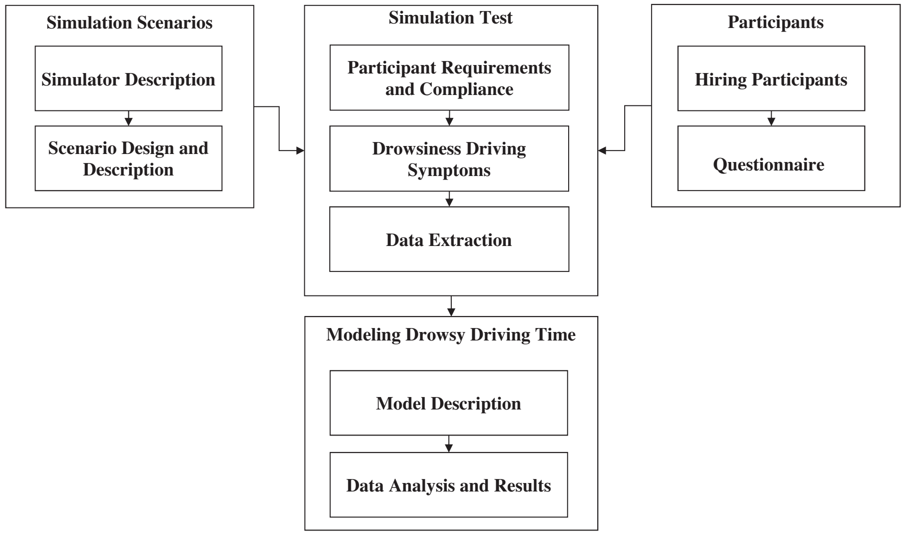


Fig. 1. Methodology framework.

a) Outside the simulator *b) Scenario projection* *c) Driver view inside the simulator*

Fig. 2. Driving simulator with eight-degree of freedom.

mental variables for the development of dynamic driving duration thresholds. In this study, the LBS system was implemented virtually and related data were extracted with the driving simulator.

3. Methodology

The study used an advanced driving simulator in investigating the impact of different factors on drowsiness driving time. The framework of the methodology is presented in Fig. 1. Details are provided in the following sections.

3.1. Simulation scenarios

3.1.1. Simulator description

The experiment was conducted using the OKTAL driving simulator, an advanced simulator with eight degrees of freedom, at Tongji University, China, as shown in Fig. 2. As shown in Fig. 2a, the simulator is one of the world's largest full motion driving simulators, provides scenarios which are extremely similar to real road situations, and uses a real vehicle within the simulator. The simulated environment is projected 250° around the driver on the curved screen of the dome which houses the car (Fig. 2b).

3.1.2. Scenario design and description

Scenarios were designed by varying parameters including weather, time-of-day, and road network in order to explore the impact of the factors on drowsy driving. With regards to the environment, the study considered three factors: precipitation (sunny or raining), daylight (day or night), and routes (three different route network designs). Each scenario was a unique combination of these three factors, resulting in 12 different test scenarios. For each of the twelve scenarios, 24 individual tests were designed, resulting in 288 tests total. More details of these factors are presented in Table 2.

For routes in the simulation, to collect observations from diverse road environments, three different networks were designed and used, as presented in Fig. 3. Driving simulation tests were run based on these maps. Table 1 presents the description of the networks. To avoid the impact of route repetition on drowsiness driving time, routes for each network were designed before the experiment. In the simulation test, drivers follow the voice guidance to keep on the designed route. The guidance system automatically guide the driver back to the designed route when a mistaken of turning into a wrong direction is made.

3.2. Participants

Experiment participants included 32 drivers with different driving experience. Before the experiment, demographic data including participants' gender, age, driving experience, and nap habit were recorded through questionnaires. Details of the participants are summarized in Table 2. Participants were instructed not to nap

Table 1
Description of the Networks.

Maps	Unit	Freeway	Arterial Road	Urban Street
1	Meter	25670	16689	10849
2	Meter	10489	16975	18800
3	Meter	3047	10665	16760

a) Map 1 b) Map 2 c) Map 3

Fig. 3. Maps of the Road Networks.

Table 2
Description Statistics of the Factors that Affect Drowsy Driving Time.

Fields	Type of variables	Values and Explanations	Percentage (%)				
			1	2	3	4	5
Gender	Discrete	1- Male, 2–Female	51.94	48.06	–	–	–
Age	Discrete	1- (20–29) years, 2- (30–39) years, 3- (40–49) years, 4- (50–59) years	36.05	29.84	18.60	15.50	
Nap habit	Discrete	1 – Yes, 2 – No	57.36	42.64	–	–	–
Driving experience	Discrete	1 – (1–3) years, 2 – (4–6) years, 3 – (7–9) years, 4 – (>9) years	36.43	43.80	10.85	8.91	
Weather	Discrete	1 – Sunny, 2 – Rainy	59.30	40.70	–	–	–
Time of Day	Discrete	1 – Day, 2 – Night	56.20	43.80	–	–	–
Temperature	Discrete	1 – (10–15) °C, 2 – (15–20) °C, 3 – (20–25) °C, 4 – (25–30) °C	25.20	41.86	28.68	4.26	
Lane width	Discrete	1 – 3.5*m*, 2 – 3.75*m*	48.83	51.16	–	–	–
Average travel speed	Discrete	1 – (0–40) *km/h*, 2 – (40–60) *km/h*, 3 – (60–80) *km/h*, 4–(80–100) *km/h*, 5–(100–120) *km/h*	5.43	36.82	14.73	15.50	27.52
		Unit	**Mean**	**Variance**	**Std. Dev**	**Min**	**Max**
Percentage of low speed driving time	Continuous	%	26.73	1015.88	31.87	0	100
Driving time on freeway (8 lanes)	Continuous	minute	12.81	209.17	14.46	0	51
Driving time on arterial road (4 lanes)	Continuous	minute	14.55	162.47	12.75	0	67
Driving time on urban street (2 lanes)	Continuous	minute	12.26	123.27	11.10	0	65

Note: – indicates that the value was not assigned with any meaning for the related indicator.

during the day of the experiment. Each participant completed 9 of the 12 possible scenarios. Both the scenarios and their order were selected randomly to avoid the impact of the learning process and participant tiredness. Due to delays and scheduling issues, several participants were unable to complete all nine tests. Therefore, 258 observations were collected in total of the possible 288.

3.3. Simulation test

During the simulation test, participants were instructed to obey all traffic regulations and not to listen to music, make phone calls, or chat with observers. The observer sitting next to the participants could not only record relevant information but also ensure compliance with these instructions. Each participant drove in the simulator once a day, requiring 8 or 9 days to complete all required tests.

Tests were conducted between 12:00 PM and 10:00 PM, with participants instructed to keep their normal daily routine while avoiding taking naps. For tests in the afternoon, participants were awake for 4–8 h prior to participation, while for tests in the evening, participants were awake between 9 and 14 h. Referring to (Qiong et al., 2006; Fan et al., 2007; Nordbakke and Sagberg 2007), driver drowsiness was detected by manual observation if more than one of the following symptoms occurred and lasted for more than 5 min: a) head nodding and the inability to keep the eyes open; b) constant yawning; c) different reactions to fight drowsiness, such as deep breaths, shaking head, rubbing eyes, moving body, and; d) poor concentration or trouble focusing. The first three symptoms could be observed externally, while the last one was estimated and reported by the participants themselves.

LBS can provide speed, road type, geographic location, and weather, which are potentially useful for modeling the probability of drowsiness. Though LBS data is not difficult to obtain in real-world traffic environments, the simulated nature of this study required LBS data to be extracted from the simulator itself. Therefore, all LBS data obtained in the experiment is therefore considered "virtual" LBS data. Three different driving routes with various combinations of freeway, arterial, and urban streets were designed. Each test scenario utilized a different combination of weather and time of day variables. The participant was instructed to drive on the predetermined route until drowsiness was detected. Meanwhile, the corresponding virtual LBS data were recorded by the simulator.

3.4. Modelling drowsy driving time

3.4.1. Accelerated failure time (AFT) model

Referring to Washington (Washington et al., 2010), an AFT model for predicting drowsy driving time can be developed and used to investigate the impact of different factors. The relationship between drowsy driving time and the influencing factors can be described using a general linear model;

$$ln(t_i) = \beta x_i' + e_i \quad (1)$$

where t_i is the drowsy driving time, $x_i' = (x_{i1}, \ldots, x_{in})'$ is the vector of variables influencing drowsy driving time, β is the vector of regression coefficients, and e_i is the residual of the model. Based on AFT theory, set $e_i = ln(t_{i0})$;

$$t_i = t_{i0} e^{\beta x_i'} \quad (2)$$

where t_{i0} is the actual drowsy driving time without influence of the covariates, which becomes $t_{i0} e^{\beta x_i'}$ with the impact of driver and environmental covariates.

In the AFT model, a drowsy driving time is considered a continuous variable, T, with a probability density function $f(t)$ and a cumulative distribution function $F(t)$.The survival function, $S(t)$, used commonly in duration analysis, is presented in the following equation to describe the probability of the drowsy driving time being larger than a specific time t:

$$S(t) = Pr(T \geq t) = \int_t^\infty f(s)\,ds = 1 - F(t), 0 < t < \infty \quad (3)$$

The hazard function, $h(t)$, another important element in duration analysis which predicts the probability of drowsiness within a very short interval $(t, t + \Delta t]$ after driving for a specific time interval, t, can be expressed as:

$$h(t) = \lim_{\Delta t} \frac{\Pr(t \langle T \leq t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)} = \frac{f(t)}{1 - F(t)} \quad (4)$$

3.4.2. AFT model for the log-logistic distribution

When a log-logistic distribution is used in the AFT model, the cumulative distribution function $F(t)$ and the probability density function $f(t)$ can be expressed by the following equations:

$$F(t) = \frac{1}{1 + (t/\alpha)^{-\rho}} \quad (5)$$

$$f(t) = \frac{\left(\rho/\alpha\right)\left(t/\alpha\right)^{\rho-1}}{\left(1 + \left(t/\alpha\right)^{\rho}\right)^2} \quad (6)$$

where α (> 0) is a scale parameter representing the median of the distribution and ρ (> 0) is a shape parameter. Setting $u_i = \rho ln\left(t_i/\alpha\right)$, the survival and hazard functions of u_i are presented as:

$$S(u_i) = \frac{1}{1 + e^{u_i}} \quad (7)$$

$$h(u_i) = \frac{e^{u_i}}{1 + e^{u_i}} \quad (8)$$

As $du_i/dln(t_i) = \rho$, and $\alpha = ln\left(\beta x'\right)$, the survival and hazard functions of $ln(t_i)$ can be generated based on the log-logistic distribution:

$$S(ln(t_i)) = S(u_i) = S\left(\rho \times \left(ln(t_i) - \beta x_i'\right)\right) \quad (9)$$

$$h(\ln(t_i)) = \rho \times \mathrm{h}(u_i) = \rho \times \mathrm{h}\left(\rho \times \left(ln(t_i) - \beta x_i'\right)\right) \quad (10)$$

Based on the AFT model, the impact of different factors on drowsy driving time can be explored.

4. Data description

A total number of 258 samples with 13 variables were obtained from the simulation experiment. Variables included driver characteristics of gender, age, nap habit, and driving experience. Environmental variables of weather condition, temperature, and time-of-day were recorded by the observer according to the experimental scenario. An additional six variables were statistics based on the data extracted from simulator which recorded the location and speed of the vehicle, the road type, and the lane width. Lane width was either 3.5 m or 3.75m, selected based on the most common lane width over the duration of the route. Traffic volumes fluctuated within a specified range, at times causing heavy traffic which require slow travel speeds. The low speed variable is defined as the speed lower than the half of design speed which is 120 km/h, 80 km/h, and 40 km/h, for freeways, arterials, and urban streets respectively. Low speed driving is used as part of simulating situation when driving in congested traffic. The percentage of low speed driving time is the ratio of driving time at the low speed to the cumulative time in once test. The explanations and statistical descriptions of the variables are provided in Table 2.

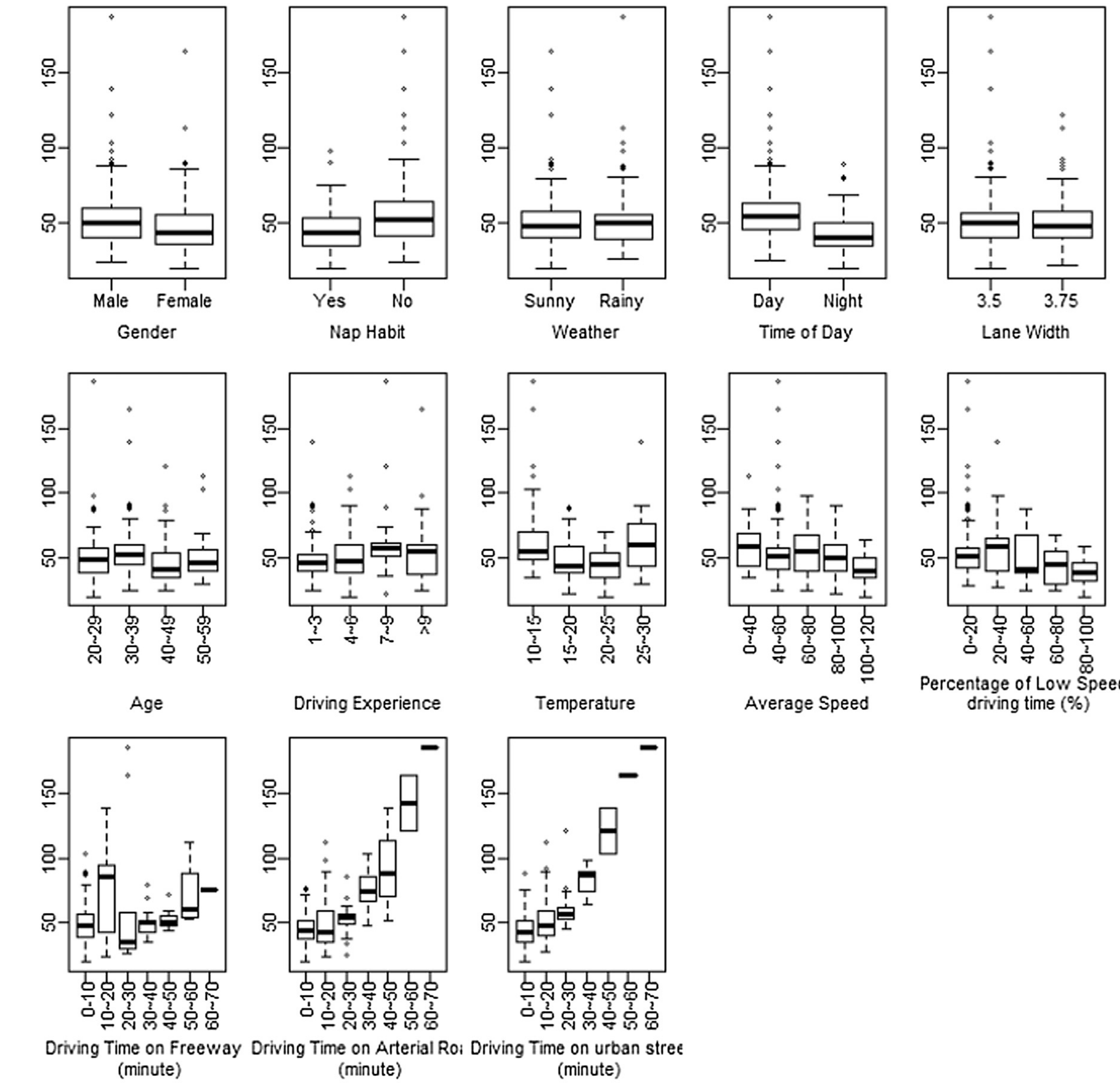
5. Data analysis and results

5.1. Relationship between different factors and drowsy driving time

The relationships between different factors and drowsy driving time are visualized as box-plots in Fig. 4. It is observed that male drivers, drivers with no nap habit, and more experienced drivers tend to have an increased drowsy driving time. Rainy weather, daytime, and lower average travel speed increase the continuous driving time before drowsiness. Temperature does not positively or negatively relate to driving duration, though results show that driving duration slightly increases in tests with both low temperature (10–15 °C) and high temperature (25–30 °C). Similarly, age does not generally relate to driving duration. Furthermore, congested conditions postpone the occurrence of drowsiness.

Further investigation of the impact of different factors on driving duration was completed using an Analysis of Variance (ANOVA) test. The significance of the examined factors are presented in Table 3. The significance level was set to 5% (95% confidence). From the ANOVA results, it was concluded that factors including nap habit, time of day, temperature, average travel speed, percentage of low speed driving time, and driving time on all the three roadway types have significant effects on driving duration before drowsiness. However, the impacts of gender, age, driving experience, weather, and lane width are not significant.

To quantify the relationships between variables, a correlation table is given in Table 4. Most variables are not significantly correlated with one another, with the most lower than 0.3 correlation. The only significant correlation is between lane width and average travel speed (0.560).



Note: the vertical axes represent drowsy driving time (minutes)

Fig. 4. Box-plots for different factors on Drowsy driving time.

Table 3
ANOVA Results of different Factors.

Variables	Sum Sq	Mean Sq	F value	Pr(>F)
Gender	2	1.58	0.05	0.829
Age	79	78.8	0.19	0.665
Nap Habit	360	359.60	11.19	0.001
Driving experience	25	24.61	0.74	0.392
Weather	51	50.58	1.52	0.219
Time of Day	256	255.93	7.86	0.005
Temperature	101	101.37	3.06	0.042
Lane width	66	66.03	1.98	0.160
Average travel speed	482	481.70	15.21	0.000
Percentage of lower speed driving time	8584	8584	22.21	0.000
Driving time on freeway	1088	1088.20	37.10	0.000
Driving time on arterial road	6	5.66	0.17	0.038
Driving time on urban street	178	177.85	5.41	0.021

5.2. Distribution of drowsy driving time

Traditional duration models were mainly developed on the assumption that the survival distribution is homogeneous among individuals. In this study, the survival distribution was the distribution of drowsy driving times, of which summary statistics are presented in Table 5. To describe drowsy driving time, four different alternatives for survival distributions, including the normal distribution, the log-normal distribution, the log-logistic distribution, and the Weibull distribution are compared in Fig. 5. Fitting parameters for the four typical distribution alternatives and the result of Kolmogorov Smirnov (K-S) test and Anderson Darling (A-D) test are presented in Table 6. Only the log-logistic distribution passed both tests with *p*-values larger than 0.05, indicating that log-logistic distribution can properly describe the distribution of drowsy driving time.

5.3. Modeling results: general log-linear model vs AFT model

Considering the high goodness-of-fit of the log-normal and log-logistic distributions in describing driving duration, the general log-linear model (GLM) and the AFT can be used in modeling driving duration based on different factors, though their performance

Table 4
Correlation Table of Variables.

	Gender	Age	Nap habit	Driving experience	Weather	Time of day	Temperature	Lane width	Average speed	Percentage of low speed driving time	Driving time on freeway	Driving time on arterial road	Driving time on urban street
Gender	1.000	0.015	0.123	0.099	−0.044	−0.003	−0.096	0.105	0.116	0.047	0.031	−0.096	−0.170
Age		1.000	−0.009	−0.033	0.013	0.012	−0.036	0.059	0.005	−0.017	−0.064	−0.031	−0.008
Nap habit			1.000	0.173	−0.008	−0.152	−0.187	0.022	−0.050	−0.072	0.046	0.180	0.169
Driving Experience				1.000	0.010	−0.130	−0.028	−0.015	0.062	0.065	−0.032	0.083	0.112
Weather					1.000	0.048	−0.100	0.004	−0.130	−0.042	0.140	−0.021	0.055
Time of day						1.000	0.013	0.003	0.129	0.182	−0.093	−0.215	−0.236
Temperature							1.000	−0.175	−0.046	0.133	−0.271	−0.036	−0.020
Lane width								1.000	0.560	0.412	0.428	−0.502	−0.492
Average speed									1.000	0.499	0.226	−0.503	−0.483
Percentage of low Speed driving time										1.000	−0.238	−0.439	−0.394
Driving time on freeway											1.000	−0.268	−0.220
Driving time on arterial road												1.000	0.462
Driving time on urban street													1.000

Table 5
Summary Statistics of Drowsy Driving Time.

Statistic	Value
Sample	258
Min	20
Max	186
Range	166
Mean	51.40
Variance	418.33
Std. Deviation	20.45
Coef. of Variation	0.39
Std. Error	1.27
Skewness	2.52
Excess Kurtosis	11.08

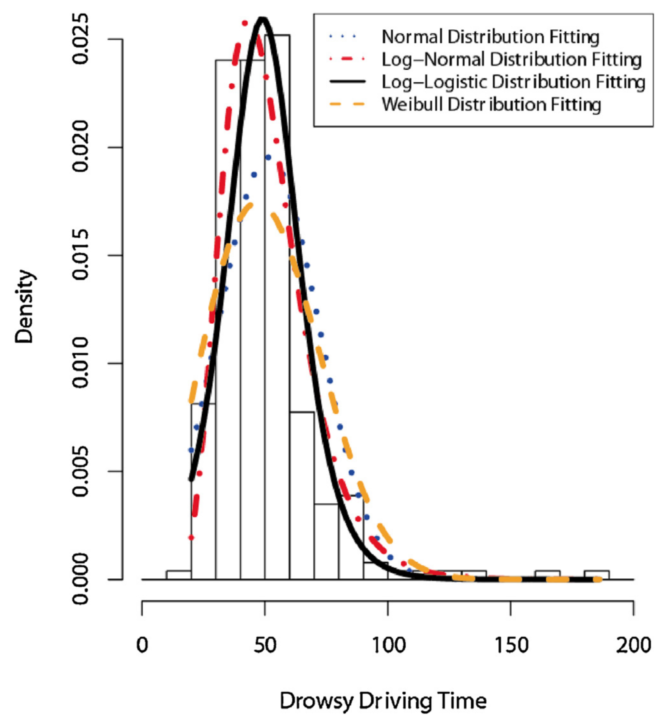


Fig. 5. Fitting distributions of drowsy driving time.

Table 6
Estimated Parameters of Driving Duration Distribution.

Distribution	Position	Scale	K-S test		A-D test	
			P Value	Stat	P Value	Stat
Log-Logistic	48.99	9.63	0.176	0.079	0.255	3.026
Log-Normal	3.89	0.34	0.073	0.068	0.027	1.234
Normal	51.40	20.45	0	0.155	0	8.949
Weibull	2.49	57.68	0	0.150	0	9.933

should be validated and compared. The significance of the factors on drowsy driving time was first examined. For this purpose, Akaike Information Criterion (AIC), a popular measure for model selection, was used. A smaller AIC value indicates a better goodness-of-fit of the model. Based on the AIC test, factors including gender, age, driving experience and weather condition are insignificant variables with relative high AIC values and p-values much greater than 0.05. These variables are therefore eliminated from the model. Results of the AIC test and remarks for the eliminated factors are presented in Table 7.

After eliminating insignificant variables, the GLM and the AFT model were estimated and results were compared as shown in Table 8. Based on the results, the AFT model has a better goodness-of-fit with smaller AIC compared to the GLM. Additionally, using

Table 7
AIC to decide the best group of factors for drowsy model.

Step	Deleted variable	AIC Step Value	Correlation coefficient with duration	Remark
1	Gender	1296.98	0.003	Gender data is almost averagely distributed
2	Age	1295.08	−0.027	Age data also is almost averagely distributed
3	Driving experience	1293.32	0.004	The box-plot shows no rhythm
4	Weather	1291.70	−0.017	The simulator is not able to simulate the actual weather condition

Table 8
Results of AFT model vs General log-linear model.

Explanatory variables	AFT			GLM		
	AIC – 1541			AIC – 1292		
	Coef.	Std.Dev	P value	Coef.	Std.Dev	P value
(Intercept)	3.085	0.064	0.000	3.076	0.114	0.000
Nap habit	0.010	0.004	0.009	–	–	–
Temperature	−0.002	0.000	0.000	−0.002	0.001	0.048
Time of day	–	–	–	–	–	–
Lane width	0.047	0.000	0.000	–	–	–
Average travel speed	–	–	–	0.000	0.000	0.021
Percentage of low speed driving time	−0.019	0.000	0.000	−0.017	0.003	0.000
Driving time on freeway	0.005	0.001	0.000	0.002	0.001	0.006
Driving time on arterial road	0.027	0.002	0.000	0.014	0.001	0.000
Driving time on urban street	0.009	0.001	0.000	0.013	0.001	0.000
Log(scale)	−2.136	0.013	0.000			

Note: in this table, – indicates that the variable is not significant and is eliminated from the model.

a confidence interval of 95%, 3 variables in the GLM (nap habit, time of day, and lane width) were found to be insignificant. The AFT model had only two insignificant variables, indicating that the AFT model may better explain the relationship between the variables and drowsy driving time. One additional advantage of the AFT model is that it can be applied to different fitting distribution models of driving. Interestingly, the impact of temperature was shown to be significantly linear, which is inconsistent with the results of the earlier box plots in Fig. 4. However, in the box plots, a low sample size for specific temperature ranges (25–30 °C had only 13 observations for 4.26% of the total number of observations) may skew the true relationship between temperature and drowsy driving time. The modeling results suggest that the relationship is, in fact, generally linear.

Factors including nap habit and lane width had a positive impact on drowsy driving time. This indicates that habitual nappers became drowsy more quickly compared to other participants. Increased lane width slightly increased the drowsy driving time. The temperature was found to be negative. Interestingly, lower temperature reduced the chance of drowsiness, which using cold air conditioning as a countermeasure to drowsy driving for long trips. The negative impact of percentage of low speed driving time indicates that drivers become drowsy more quickly in congested traffic, likely due to monotonous driving.

Driving time on all different roadway types and drowsy driving time have a positive correlation. Obviously, if drivers spend more time on any roadway type, then they are spending more time driving without becoming drowsy. However, comparing the magnitude of the coefficients can reveal more about the effect of roadway type. Freeway driving had the smallest coefficient (0.005) indicating that drowsy driving time is shortest in freeway environments, perhaps because they tend to be more monotonous than other facilities. Drivers are more resistant to drowsiness on arterial roads mainly due to the complex driving environment which has more stimuli and requires increased driver attentiveness.

6. Conclusion

This paper presents a hazard-based duration model, the Accelerated Failure Time Model, to estimate the drowsy driving time based on virtual LBS and demographic data. The model was built based on data from a driving simulation experiment conducted using the advanced driving simulator in Tongji University, China, with 8 ° of freedom and 250 ° of view. In the experiment, 258 observations from driving tests by 32 participants were extracted and investigated. For this study, the observed drowsy driving time is best bit by the log-logistic distribution. Accordingly, an AFT model based on the log-logistic distribution was developed to investigate factors on drowsy driving time. A GLM, which is typically used in drowsiness prediction modeling, was also estimated and compared to the performance of the AFT model. Model results show that AFT model works reliably with a better goodness-of-fit than the GLM model, and better explains the impact of different factors from built environment and individual drivers on drowsy driving time. Reduced drowsy driving time was associated with high temperature, driving at night, decreased lane width, having a nap habit but not napping, and high average driving speed potentially resulting from reduced traffic in monotonous driving environments. Driving on freeways with more monotonous conditions was found to increase the chance of drowsiness. in contrast, driving conditions on arterial roads, with more complicated situations, extended the drowsy driving time.

By using virtual LBS data extracted from one of the most advanced driving simulators in the world, this paper provides an example of using LBS data for predicting drowsy driving in real-time situations. In practical applications, similar variables would be extracted from the LBS terminators and applications. Together with the LBS data, the AFT model can be integrated in drowsy driving warning systems to alert drivers of drowsiness in real time. Dynamic driving duration thresholds, compared to the fixed driving time thresholds used in most past systems, provide more are more time-sensitive and reliable.

The advanced driving simulator provides driving scenarios highly similar to real world situations, making drivers' perception and reaction highly close to real world situations. Still, the experiments are still affected by various factors, such as projection resolution and lighting conditions among others. It is acknowledged that drivers may behave differently in a simulation compared to real driving situations. Therefore, before implementation in practice, the model must be further validated and improved based on data from real-world situations, which is currently a challenge for transportation engineers and researchers due to the difficulties of measuring drowsy driving time in the field.

Acknowledgement

The authors would like to acknowledge the financial support by National Natural Science Foundation (71301119) and Fundamental Research Funds for the Central Universities.

References

Accidentes, P., 2015. Road Safety Annual Report.2015.
Arnold, P., Hartley, L., 1998. It's not just hours of work; ask the drivers. In: International Conference on Fatigue and Transportation, 3rd, 1998, Fremantle, Western Australia.
Eriksson, M., Papanikolopoulos, N.P., 2001. Driver fatigue: a vision-based approach to automatic diagnosis. Transp. Res. Part C: Emerg. Technol. 9 (6), 399–413.
Fan, X., Yin, B.C., Sun, Y.F., 2007. Yawning detection for monitoring driver fatigue. 2007 International Conference on Machine Learning and Cybernetics.
Ferdinands, A., 1999. Intelligent transport systems and road safety. In: Insurance Commission of Western Australia Conference on Road Safety, 1999, Perth, Western Australia.
Fischer, M., Sehammer, H., Palmkvist, G., 2010. Motion Cueing for 3-, 6- and 8-degrees-of-freedom Motion Systems the Driving Simulation–Conference , Europe, 2010.
Friswell, R., Williamson, A., 2013. Comparison of the fatigue experiences of short haul light and long distance heavy vehicle drivers. Saf. Sci. 57, 203–213.
Grace, R., Steward, S., 2001. Drowsy driver monitor and warning system. In: International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design.
Greene, W., 2003. Econometric analysis. J. Am. Stat. Assoc. 89 (89), 182–197.
Horne, J.A., Reyner, L.A., 1995. Sleep related vehicle accidents. Br. Med. J., 656–667.
Hu, X., Eberhart, R., Foresman, B., 2010. Modeling drowsy driving behaviors. vehicular electronics and safety (ICVES). 2010 IEEE International Conference On, IEEE.
Jones, C.B., Dorrian, J., Rajaratnam, S.M.W., Dawson, D., 2005. Working hours regulations and fatigue in transportation: a comparative analysis. Saf. Sci. 43 (4), 225–252.
Junhua, W., Haozhe, C., Shi, Q., 2013. Estimating freeway incident duration using accelerated failure time modeling. Saf. Sci. 54, 43–50.
Kulmala, R., 1997. The potential of ITS to improve safety on rural roads. In: Proceedings of the 4th World Congress on Intelligent Transport Systems, Berlin, Germany Oct. 21424.
Lin, H., Wang, X., Wu, Z., 2015. DSC 2015 Europe Driving Simulation Conference & Exibition.
Liu, Y.-C., Wu, T.-J., 2009. Fatigued driver's driving behavior and cognitive task performance: effects of road environments and road environment changes. Saf. Sci. 47 (8), 1083–1089.
Lyznicki, J.M., Doege, T.C., Davis, R.M., Williams, M.A., 1998. Sleepiness driving, and motor vehicle crashes. JAMA J. Am. Med. Assoc. 279 (23), 1908–1913.
Maycock, G., 1996. Sleepiness and driving: the experience of UK car drivers. J. Sleep Res. 5 (4), 229–231.
Mcdonald, A.D., Lee, J.D., Schwarz, C., Brown, T.L., 2013. Steering in a random forest: ensemble learning for detecting drowsiness-related lane departures. Hum. Factors: J. Hum. Factors Ergon. Soc. 56 (5), 986–998.
Nordbakke, S., Sagberg, F., 2007. Sleepy at the wheel: knowledge, symptoms and behaviour among car drivers. Transp. Res. Part F 10 (1), 1–10.
Pack, A.I., Pack, A.M., Rodgman, E., Cucchiara, A., Dinges, D.F., Schwab, C.W., 1995. Characteristics of crashes attributed to the driver having fallen asleep? Accid. Anal. Prev. 27 (6), 769–775.
Pilcher, J.J., Huffcutt, A.I., 1996. Effects of sleep deprivation on performance: a meta-analysis. Sleep: J. Sleep Res. Sleep Med. 19 (4), 318–326.
Qiong, W., Jingyu, Y., Mingwu, R., Yujie, Z., 2006. Driver fatigue detection: a survey. 2006 6th World Congress on Intelligent Control and Automation.
Regan, M.A., Oxley, J., Godley, S., Tingvall, C., 2001. Intelligent Transport Systems: Safety and Human Factors Issues.
Saccomanno, F.F., Yu, M., Shortreed, J., 1970. Effect of driver fatigue on truck accident rates. WIT Trans. Built Environ., 18.
Sang, Y., Li, J., 2012. Research on beijing bus driver psychology fatigue evaluation. Procedia Eng. 43, 443–448.
Slob, J. (2008). State-of-the-Art Driving Simulators a Literature Survey. Eindhoven.
Smith, S., Carrington, M., Trinder, J., 2005. Subjective and predicted sleepiness while driving in young adults. Accid. Anal. Prev. 37 (6), 1066–1073.
Suite, 2010. Asleep at the wheel: the prevalence and impact of drowsy driving. Drivers 279 (7), 80.
Tefft, B.C., 2014. Prevalence of Motor Vehicle Crashes Involving Drowsy Drivers, United States, 2009–2013. AAA Foundation for Traffic Safety, Washington DC, US.
Vanlaar, W., Simpson, H., Mayhew, D., Robertson, R., 2008. Fatigued and drowsy driving: a survey of attitudes: opinions and behaviors. J. Safety Res. 39 (3), 303–309.
Wang, H., Zhang, C., Shi, T., Wang, F., Ma, S., 2015. Real-time EEG-based detection of fatigue driving danger for accident prediction? Int. J. Neural Syst. 25 (2), 1550002–1550002.
Washington, S.P., Karlaftis, M.G., Mannering, F., 2010. Statistical and Econometric Methods for Transportation Data Analysis. CRC Press.
Williamson, A.M., Feyer, A.-M., Mattick, R.P., Friswell, R., Finlay-Brown, S., 2001. Developing measures of fatigue using an alcohol comparison to validate the effects of fatigue on performance. Accid. Anal. Prev. 33 (3), 313–326.
Yang, X., Abdel-Aty, M., Huan, M., Peng, Y., Gao, Z., 2015. An accelerated failure time model for investigating pedestrian crossing behavior and waiting times at signalized intersections. Accid. Anal. Prev. 82, 154–162.
Yanli, M.A., Wang, Y., Pei, Y., 2009. Experimental psychology study on relationship between fatigue and driving time? J. Southw. Jiaotong Univ. 44 (4), 535–540.
Yeo, M.V.M., Li, X., Shen, K., Wilder-Smith, E.P.V., 2009. Can SVM be used for automatic EEG detection of drowsiness during car driving? Saf. Sci. 47 (1), 115–124.
Young, K., Regan, M., Mitsopoulos, E., Haworth, N., 2003. Acceptability of In-vehicle Intelligent Transportation Systems to Young Novice Drivers in New South Wales.
Zhang, G., Yau, K.K.W., Zhang, X., Li, Y., 2016. Traffic accidents involving fatigue driving and their extent of casualties. Accid. Anal. Prev. 87, 34–42.