Exploring Discretionary Lane-changing Behaviors of Autonomous Vehicles Using the Waymo Open Dataset

Xiao Wen[1], Chunxi Huang[2], Sisi Jian[3*], Dengbo He[4*]

[1] Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR; E-Mail: xwenan@connect.ust.hk

[2] Robotics and Autonomous Systems, Interdisciplinary Programs Office, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR; Email: tracy.huang@connect.ust.hk

[3] (Corresponding Author) Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR; E-Mail: cesjian@ust.hk

[4] (Corresponding Author) Intelligent Transportation Thrust, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China; Email: dengbohe@ust.hk

ABSTRACT

The gradual deployment of automated vehicles (AVs) will lead to a transition period where AVs will share the roads with human-driven vehicles (HDVs), resulting into the so-called mixed traffic. Nevertheless, AV-HDV interactions, especially how the discretionary lane-changing (DLC) behaviors of AVs will affect the following vehicles (FVs) in the target lane, remains the key research gap. In this paper, the real-world Waymo Open Dataset is used to analyze AV DLC behaviors in comparison to HDV DLC behaviors. DLC characteristics, including driving volatility of FVs, are quantified and compared. Additionally, the block maxima (BM) model in extreme value theory (EVT) is adopted to estimate the crash risks using the gap time (GT) as the surrogate safety measure (SSM). The results reveal that compared to HDV DLC, AV DLC has lower speed and yaw rate volatility of FVs, and smaller acceleration rates of FVs. The results of non-stationary BM models show that the crash risk in AV DLC is half of that in HDV DLC. These findings confirm the traffic and safety benefits of AVs and can be beneficial for AV tech companies to improve AV controllers.

INTRODUCTION

Because of the gradual deployment of autonomous vehicles (AVs), AVs and human-driven vehicles (HDVs) are expected to share the roads in the near future,

resulting into mixed traffic. Previous literature has identified that human drivers may behave significantly differently in mixed traffic as compared to when there are only HDVs in the traffic stream (Mahdinia et al., 2021; Wen et al., 2022a; Zhao et al., 2020). Thus, understanding how HDVs' behaviors change in mixed traffic is critical for evaluating the effects of AVs on traffic safety, traffic efficiency, energy consumption and exhaust emission (Hu et al., 2022). Further, modeling AV-HDV interactions can attain more insights into the improvement of AV controllers and propose appropriate public policies regarding AVs (Di and Shi, 2021).

Since currently the market penetration rates of AVs is low, empirical AV trajectory data is still lacking. When investigating AV-HDV interactions, former works mostly implemented two methods, i.e., traffic/numerical simulations (e.g., Dixit et al., 2019) and field experiments (e.g., Mahdinia et al., 2021; Zhao et al., 2020). However, traffic/numerical simulations simplify and even ignore important characteristics of mixed traffic, leading to questionable effects of AVs. Field experiments are often conducted with limited sample sizes (i.e., the number of driving events) and cannot reproduce the real-world driving scenarios with large speed fluctuations and complex interactions between road agents. In summary, the limitations of traffic/numerical simulations and field experiments can induce biased conclusions. In recent years, there have been more and more AVs being deployed on public roads and some AV tech firms (e.g., Waymo and nuScenes) have released large-scale real-world AV driving datasets (Wen et al., 2022b). These datasets include high-quality field data of movements of AV and surrounding road agents in the real world, and thus enable the transportation research community to analyze the effects of AVs on traffic flow, as well as HDVs' behavioral changes when interacting with AVs.

Previous literature on AV-HDV interactions usually analyzed car-following scenarios which involve with only longitudinal vehicle control (Mahdinia et al., 2021; Wen et al., 2022a; Zhao et al., 2020). On the contrary, lane-changing (LC) scenarios, which correspond to rear-end and sideswipe crashes (Ali et al., 2022a), are still under-studied. LC scenarios involve with both longitudinal and lateral movements of involving road agents. In LC scenarios, the lead vehicle changes the current lane into an neighboring lane, which may cause the following vehicle (FV) in the neighboring lane to decelerate or even stop, leading to stop-and-go oscillations and bottlenecks in traffic flow (Jiang et al., 2021). Based on the objectives of drivers, LC maneuvers are classified into two categories, i.e., mandatory lane-changing (MLC) and discretionary lane-changing (DLC). MLC is a required task when drivers must leave the current lane, e.g., to turn left when approaching an intersection. DLC is voluntary when drivers perceive that driving conditions in the target lane are better, e.g., to reach a desired speed. The latter is more complex and dangerous than the former (Ali et al., 2022b; Toledo et al., 2005). To this end, the effects of DLC behaviors of AVs on the

surrounding HDVs, especially the FV in the target lane, need to be studied.

This paper takes the first attempt to investigate the impacts of AVs' DLC behaviors on the FV in terms of driving volatility and crash risks and compare these metrics to those of HDV's DLC behaviors. The real-world AV trajectory data is retrieved from the Waymo Open Dataset (Ettinger et al., 2021). The contributions of this paper are listed as follows. First, it fulfills an in-depth analysis of HDVs' behavioral adaptations when AVs are changing lanes using the real-world autonomous driving dataset which provides more insights into complicated driving conditions. Second, the crash risks of AV and HDV DLC maneuvers are calculated and compared using the extreme value theory (EVT).

METHODOLOGY

Driving Volatility

The driving volatility is adopted to measure the variation of driving behaviors by extracting useful information from the longitudinal and lateral vehicle control. It has been identified that higher driving volatility is associated with higher driver instability, which corresponds to higher crash risks, more energy consumption, and increased exhaust emissions (Wen et al., 2022a). Three kinds of driving volatility measures are introduced and quantified for the selected DLC events: speed-based volatility, acceleration-based volatility, and yaw-rate-based volatility. The mathematical formulations of the driving volatility are given in Eqs. (1)-(6). Note that the driving volatility measure is computed for each FV involved with the selected DLC events.

Standard deviation (Std): Std is one of the most commonly-used variation measure, which can be calculated as follows:

$$Std = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}} \tag{1}$$

where x_i is the ith observation, $\bar{x}$ is the mean value of observations and n is the sample size. Std can be applied to measuring speed, acceleration and yaw rate.

Mean Absolute Deviation (D_{mean}): D_{mean} represents the average distance between the observations and the mean value and can be computed as follows:

$$D_{mean} = \frac{\sum_{i=1}^{n}|x_i-\bar{x}|}{n} \tag{2}$$

where x_i is the ith observation, $\bar{x}$ is the mean value of observations and n is the

sample size. Similar to Std, it can be applied to speed, acceleration and yaw rate.

Coefficient of Variation (C_v): C_v captures the dispersion of data using the ratio of the standard deviation to the mean value, which can be expressed as below:

$$C_v = \frac{Std}{|\bar{x}|} * 100 \tag{3}$$

where Std is the standard deviation and $\bar{x}$ is the mean value of observations. It can only be applied to speed.

Quartile Coefficient of Variation (Q_{cv}): Q_{cv} is another prevailing statistic for describing the dispersion of data:

$$Q_{cv} = \frac{Q_3 - Q_1}{Q_3 + Q_1} \tag{4}$$

where Q_1 and Q_3 are the first and third quartiles of data, respectively. It can only be applied to speed.

Time-Varying Stochastic Volatility (V_f): V_f reflects the fluctuation of data by computing the changes of the proportion of observations:

$$V_f = \sqrt{\frac{\sum_{i=1}^{n}(r_i - \bar{r})^2}{n-1}} \tag{5}$$

$$r_i = \ln\left(\frac{x_i}{x_{i-1}}\right) * 100 \tag{6}$$

where x_i and x_{i-1} are the ith and $i-1$th observations, ln is the natural logarithm, $\bar{r}$ is the mean value of r_i and n is the sample size. It can only be applied to speed.

Extreme Value Theory (EVT)

To estimate the crash risks of DLC events, this study opts for the block maxima (BM) approach in EVT to ensure that extreme events are sufficiently smooth to enable the extrapolation from observable traffic conflicts to unseen traffic crashes. The gap time (GT) which is a variant of time-to-collision (TTC) is adopted as the surrogate safety measure (SSM). GT represents the elapsed time between the expected completion time of LC for LCV and the expected time for FV to arrive at the conflict area. GT is negatively proportional to crash risks meaning that smaller GT values indicate higher crash risks.

Block maxima (BM) approach

In the BM approach, observations are aggregated into fixed intervals over time, and the maxima in each interval are treated as extremes. Suppose that a series of random observations $\{X_1, X_2, \dots, X_n\}$ are independently and identically distributed which follow an unknown distribution function $F(x) = Pr(X_i \leq x)$, and let maximum $M_n = max(X_1, X_2, \dots, X_n)$. When n is approaching to the infinity ($n \to \infty$), M_n will converge to a GEV distribution as shown in Eq. (7):

$$G(x) = \exp\{-[1 + \epsilon(\frac{x-\mu}{\sigma})]^{\frac{-1}{\epsilon}}\} \tag{7}$$

where μ is the location parameter, σ is the scale parameter, and ϵ is the shape parameter, and we have $-\infty < \mu < \infty$, $\sigma > 0$ and $-\infty < \epsilon < \infty$. When the shape parameter ϵ is 0, the GEV is a Gumbel distribution; when ϵ is positive, the GEV is a Frechet distribution; when ϵ is negative, the GEV is a Weibull distribution.

The tail behavior of an extreme value distribution should be focused on since the EVT enables the extrapolation of the fitted distributions to the unobserved events. When it comes to traffic safety, this is performed by using observable traffic conflicts to predict traffic crashes which are unobservable in a short time period. To measure the crash risks in DLC maneuvers, GT is chosen as a SSM. When $GT \leq 0$, there will be trajectory overlaps between LCV and FV, indicating the occurrence of traffic crashes. The negated values of GTs are used to fit the GEV distribution, and a crash can be identified if negated $GT \geq 0$. The crash risk is calculated based on the tail region of the GEV distribution as follows:

$$R = \Pr(Z \geq 0) = 1 - G(0) \tag{8}$$

where R is the crash risk and also the probability of negated $GT \geq 0$, Z represents the maximum negated GT, and $G(\cdot)$ represents the fitted GEV distribution.

DATA SOURCES

AV and HDV DLC events used in this study are extracted from the motion part of Waymo Open Dataset. Waymo is a leading AV tech firm and has been conducting road tests using SAE Level 4 AVs for more than 32 million km (kilometers) on public roads in the U.S. (Ettinger et al., 2021). 28,358 clips of 20-second scenes representing approximately 157.5 hours of driving data are retrieved from the motion part of Waymo Open Dataset. The motion part contains high-quality and continuous records of road agents' type, size (e.g., length, width and height), position and movement (e.g., speed and yaw angle) at 10-Hz frequency. Interesting readers are suggested to refer to Waymo

Open Dataset for more information, following the link https://waymo.com/open/data/.

All the 20-second clips are manually reviewed by the research team to detect possible DLC events. Note that since the sample size is limited to support studies related to DLC events on highways (40 events for AV DLC), only DLC events that occur on surface roads are used in this study. After screening, 180 AV DLC and 178 HDV DLC events have been extracted from the motion part.

RESULTS

Driving Volatility Analysis

Table 1 shows the summary statistics of driving volatility of FVs in different DLC modes. It should be noted that only the driving data collected within the DLC period is included in the computation. The column named "Difference (%)" represents the mean value changes in driving volatility of AV DLC with respect to HDV DLC. The positive (negative) values of "Difference (%)" represent the increase (decrease) in driving volatility of AV DLC relative to HDV DLC.

Table 1 Driving volatility comparison of FVs between AV DLC and HDV DLC

Metrics	LC mode								Difference (%)
	AV DLC ($n = 180$)				HDV DLC ($n = 178$)				
	Max	Min	Mean	Std	Max	Min	Mean	Std	
Speed volatility									
$Std(m/s)$	2.115	0.057	0.854	0.488	4.792	0.067	1.049	0.733	**-18.59%**
$D_{mean}(m/s)$	1.865	0.043	0.739	0.439	4.396	0.056	0.913	0.656	**-19.06%**
C_v	25.539	0.278	7.221	5.255	41.201	0.378	9.979	7.952	**-27.64%**
Q_{cv}	27.479	0.090	5.935	4.820	42.867	0.261	8.447	7.262	**-29.74%**
$V_f(m/s)$	1.434	0.045	0.397	0.225	2.838	0.053	0.490	0.427	**-18.98%**
Acceleration volatility									
$Std(m^2/s)$	0.807	0.015	0.376	0.185	1.305	0.045	0.389	0.250	**-3.34%**
$D_{mean}(m/s)$	0.709	0.013	0.327	0.164	1.218	0.032	0.338	0.226	**-3.25%**
Yaw rate volatility									
$Std(degree/s)$	3.549	0.455	1.001	0.519	6.035	0.366	1.159	0.800	**-13.63%**
D_{mean}(degree/s)	2.477	0.348	0.758	0.373	4.665	0.299	0.894	0.607	**-15.21%**

The results shown in Table 1 indicate that FVs show lower speed volatility in AV DLC compared to those in HDV DLC. The percentage changes of standard deviation (Std), mean absolute deviation (D_{mean}), coefficient of variation (C_v), quartile coefficient of variation (Q_{cv}), and time-varying stochastic volatility (V_f) of speed for AV DLC are -18.59%, -19.06%, -27.64%, -29.74%, and -18.98%, respectively.

All driving volatility measures except for V_f are found to be significantly different between these two DLC modes ($Std: p-value = .046$; $D_{mean}: p-value = .04$; $C_v: p-value = .003$; $Q_{cv}: p-value = .002$) by the Mann-Whitney U test. It may be explained by the precise motion control logic of AVs which enables them to handle complex driving scenarios on surface roads.

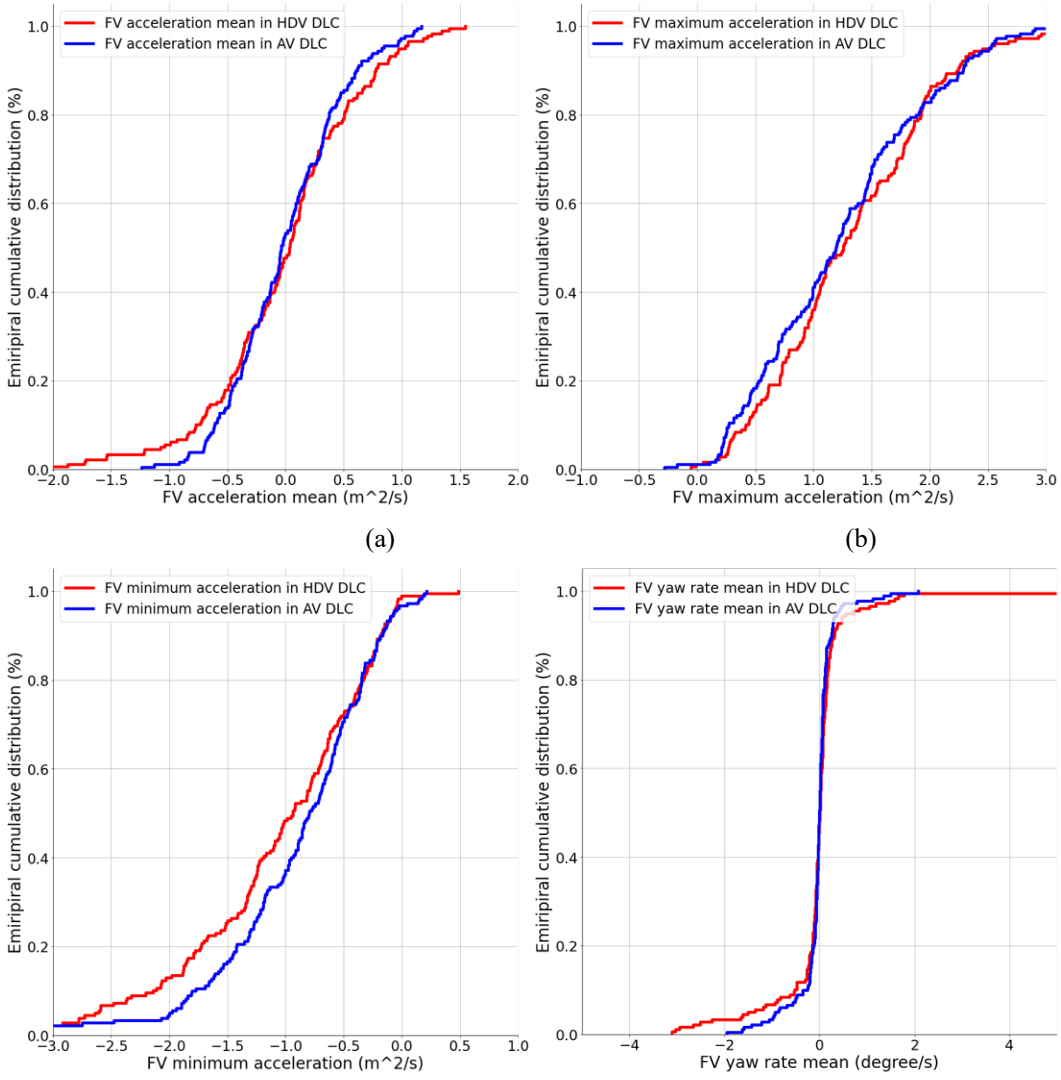


Fig. 1 Empirical cumulative distributions of (a) acceleration mean; (b) maximum acceleration; (c) minimum acceleration and (d) yaw rate mean

To further illustrate the impact of DLC on traffic flow, the speed change rate is computed:

$$speed\ change\ rate = \frac{v_E - v_S}{v_E} * 100\% \tag{9}$$

where v_E and v_S are the FV speed at the start and end points of the DLC, respectively. The FV is considered to be adversely affected only if the absolute value of speed change rate is higher than 20%. In our dataset, in 42.78% of AV DLC events and in 50% of HDV DLC events, the speed change rate of FV exceeded 20%. The comparisons of speed volatility and speed change rate suggest that the penetration of AVs in the mixed traffic can improve the driving smoothness of FVs, and thus, mitigate stop-and-go oscillations resulting from DLC behaviors.

As shown in Table 1, 3.34% and 3.25% reductions are found in the Std and D_{mean} of acceleration. The mean, maximum and minimum values of acceleration of FVs are assessed as displayed in Fig. 1(a)-(c). Although Mann-Whitney U test shows no significant differences in the acceleration volatility (, $Std: p - value = .5$; $D_{mean}: p - value = .5$), one can observe in Fig. 1(a)-(c) that FVs in HDV DLC are more likely to perform harsh acceleration and deceleration., indicating that AV DLC induces lower acceleration rates of FVs compared to the HDV counterpart, and thus lead to better driving comfort.

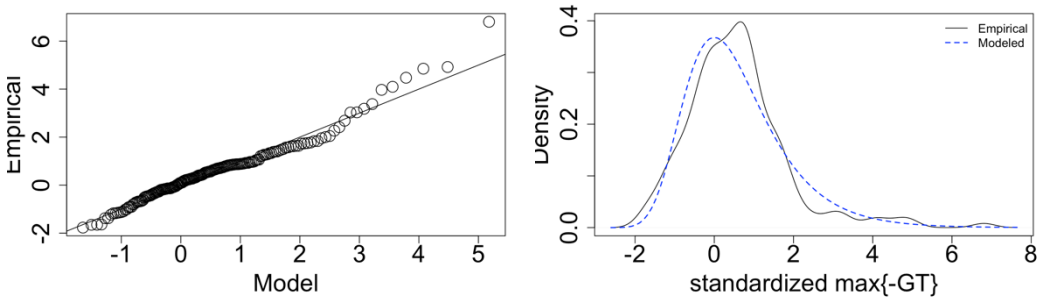
Yaw rate describes the angular speed of the forward direction of the vehicle, which plays a crucial role in vehicle lateral dynamics. From Table 1, one can observe that FVs in AV DLC events has smaller Std and D_{mean} of yaw rate than that in HDV DLC events. The differences are statistically significant based on the Mann-Whitney U test ($Std: p - value = .04$; $D_{mean}: p - value = .02$). The empirical cumulative distributions of the mean values of yaw rate are depicted in Fig. 1(d). It shows that FVs in AV DLC have smaller and more stable yaw rates and therefore more lateral stability compared to FVs in HDV DLC.

Crash Risk Analysis

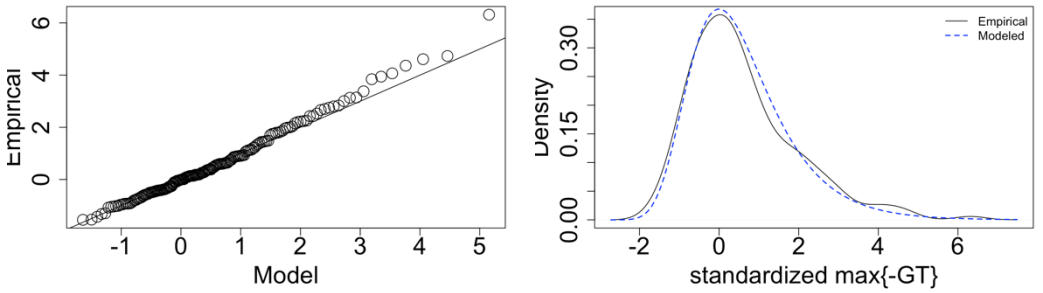
Note that since the aim of this study is to understand how DLC behaviors of AVs will affect FVs, only traffic conflicts between LCVs and FVs are focused on in this part. According to Ali et al. (2022b), each block represents a DLC maneuver where the duration of the block is the same as the duration of the corresponding DLC event. For each block, the minimum value of GT is chosen and used to develop the BM model. Only the GT value less than $3s$ is treated as an extreme event. Then, the GT data is filtered according to this approach, resulted in 177 maxima for AV DLC and 173 maxima for HDV DLC.

Table 2 presents the stationary and selected non-stationary BM models based on the maximum likelihood estimation (MLE) method. Two covariates are included in each non-stationary model: $lag\_spacing$ which means the distance (in meters) between LCV and FV at the start point of DLC and $relspd\_mean\_lcv\_fv$ which represents the average relative speed between LCV and FV during the LC period. One can see that incorporating the covariates into the location parameter can greatly reduce

the negative log-likelihood and thus improve the model fit. Fig. 2 shows the simulated quantile-quantile (Q-Q) plot and the probability density function of the empirical and modeled standardized maximum negated GT derived from the non-stationary BM models. For both DLC modes, the Kolmogorov–Smirnov (K-S) test is implemented, of which the null hypothesis is that the sample is drawn from the fitted GEV distribution. In both conditions, $p-values$ are significantly greater than 0.05 ($AV\ DLC: p-value = .9$; $HDV\ DLC: p-value = .99$), meaning that the null hypothesis cannot be rejected.



(a) AV DLC



(b) HDV DLC

Fig. 2 QQ-plot (left) and probability density function (right) for the non-stationary BM model

The crash risk suggesting the probability of a collision during an AV DLC or HDV DLC is computed using Eq. (8). The confidence intervals of crash risks are generated based on simulation where estimated parameters are assumed to follow the normal distribution. Based on the non-stationary BM model, the crash risk is computed as 0.010 with a 95% confidence interval (0.001,0.039) for AV DLC and 0.020 with a 95% confidence interval (0.003,0.078) for HDV DLC. Overall, the DLC maneuvers of AVs has been found to improve traffic safety significantly compared to those of HDVs, with a 2 times reduction in crash risk. This finding indicates the efficacy and potential of AVs in eliminating DLC crash risks.

Table 2 BM model estimation results

DLC mode	Model type	nllh	Location$(\mu)(standard\ error, SE)$			Scale$(\sigma)(SE)$	Shape$(\epsilon)(SE)$	AIC	BIC
			μ_0	$\mu_{lag\_spacing}$	$\mu_{relspd\_mean\_lcv\_fv}$				
AV DLC	Stationary	203.655	-1.391(NA)	--	--	1.612(NA)	-1.185(NA)	413.310	422.839
	Non-stationary	79.033	-0.242(0.069)	-0.043(0.003)	-0.338(0.027)	0.395(0.022)	-0.334(0.037)	168.065	183.946
HDV DLC	Stationary	143.023	-0.981(0.058)	--	--	0.690(0.052)	-0.679(0.075)	292.047	301.507
	Non-stationary	125.490	-0.739(0.052)	-0.009(0)	-0.126(0.009)	0.637(0.045)	-0.707(0.053)	260.979	276.746

Note: nllh: negative log-likelihood; NA: indicates that the standard error of the corresponding parameter does not exist.

CONCLUSIONS

We take the first attempt to investigate the attributes of AV LDC and its effects on FVs in the target lane using the real-world AV trajectory data. Driving volatility is measured and compared based on 180 AV DLC and 178 HDV DLC events collected on surface roads. Using the gap time (GT) as the SSM, the BM approach in EVT is employed to calculate the crash risks from the observed traffic conflicts in two DLC conditions.

For the driving volatility, in AV DLC, the FVs in the target lane show significantly lower speed and yaw rate volatility, indicating more longitudinal and lateral stability of FVs. Compared to HDV DLC, a smaller portion of FVs will be adversely affected in AV DLC events in terms of the speed change rate. Besides, smaller acceleration rates in AV DLC can be observed, which indicates more comfortable driving experiences. In short, inserting AVs into the traffic stream will have benefits in a variety of aspects, including the behavioral certainty and oscillation mitigation. The reasons may be that AVs with better speed management and route planning tend to have fewer speed and yaw angle fluctuations, and adopt more conservative driving strategies, and thus impose less interference on FVs.

Moreover, by developing the BM models, it is identified that in a AV DLC event, the crash risk is significantly lower compared to that in a HDV DLC event. This may be attributed to AVs' better ability to sense their surrounding environments and thus plan their routes in advance, which can minimize the uncertainty during the DLC decision-making process. For HDV DLC, higher crash risk is expected since most of the crashes happen due to human driver's risky behaviors caused by its uncertainty (Ali et al., 2022b).

REFERENCES

Ali, Y., Haque, M.M., and Zheng, Z. (2022a). "An Extreme Value Theory approach to estimate crash risk during mandatory lane-changing in a connected environment." *Analytic methods in accident research*, 33, 100193.

Ali, Y., Haque, M.M., and Zheng, Z. (2022b). "Assessing a Connected Environment's Safety Impact During Mandatory Lane-Changing: A Block Maxima Approach." *IEEE Transactions on Intelligent Transportation Systems*.

Di, X., and Shi, R. (2021). "A survey on autonomous vehicle control in the era of mixed-autonomy: From physics-based to AI-guided driving policy learning." *Transportation research part C: emerging technologies*, 125, 103008.

Dixit, V., Xiong, Z., Jian, S., and Saxena, N. (2019). "Risk of automated driving: Implications on safety acceptability and productivity." *Accident Analysis & Prevention*, 125, 257-266.

Ettinger, S., Cheng, S., Caine, B., Liu, C., Zhao, H., Pradhan, S., Chai, Y., Sapp, B., Qi,

C., Zhou, Y., and Yang, Z. (2021). "Large Scale Interactive Motion Forecasting for Autonomous Driving: The Waymo Open Motion Dataset." *Proceedings of the IEEE/CVF International Conference on Computer Vision*.

Hu, X., Zheng, Z., Chen, D., Zhang, X., and Sun, J. (2022). "Processing, assessing, and enhancing the Waymo autonomous vehicle open dataset for driving behavior research." *Transportation Research Part C: Emerging Technologies*, 134, 103490.

Jiang, L., Xie, Y., Wen, X., Chen, D., Li, T., and Evans, N.G. (2021). "Dampen the stop-and-go traffic with connected and automated vehicles–a deep reinforcement learning approach." *Proceedings of 7th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*.

Mahdinia, I., Mohammadnazar, A., Arvin, R., and Khattak, A. J. (2021). "Integration of automated vehicles in mixed traffic: Evaluating changes in performance of following human-driven vehicles." *Accident Analysis & Prevention*, 152, 106006.

Wen, X., Cui, Z., and Jian, S. (2022a). "Characterizing car-following behaviors of human drivers when following automated vehicles using the real-world dataset." *Accident Analysis & Prevention*, 172, 106689.

Wen, X., Jian S. and He, D. (2022b). "Modeling the Effects of Autonomous Vehicles on Human Driver Car-Following Behaviors using Inverse Reinforcement Learning." Working paper, Hong Kong University of Science and Technology, Hong Kong, China.

Zhao, X., Wang, Z., Xu, Z., Wang, Y., Li, X., and Qu, X. (2020). "Field experiments on longitudinal characteristics of human driver behavior following an autonomous vehicle." *Transportation research part C: emerging technologies*, 114, 205-224.