European Transport \ Trasporti Europei (2026) Issue 106, Paper n° 10, ISSN 1825-3997

Analyzing Rear-End Collisions: A Comparative Study of ADS and ADAS Involvement

Joydeep Banik[1] , Md Sifat Bin Siraj[2*], Tiziana Campisi[3]

[1] Department of Civil Engineering, Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh. Email: joydeepbanik1999@gmail.com ORCID: https://orcid.org/0009-0008-8459-7014.
[2] School of Transportation and Logistics, Southwest Jiaotong University, Sichuan, China. Email: sifatbinsiraj@gmail.com. ORCID: https://orcid.org/0000-0002-7734-6590.
[3]Department of Engineering and Architecture, University of Enna Kore, Enna, Italy. Email: tiziana.campisi@unikore.it ORCID: https://orcid.org/0000-0003-4251-4838

Abstract

There is a growing trend towards autonomous vehicles (AV). Such vehicles have demonstrated the potential to reduce the traffic crashes related to human errors. In recent years, the number of AV crashes has increased. Most research related to these crashes indicates that AVs are prone to being rear-ended. However, there has been little analysis of how these incidents occur or what factors contribute to them. A lack of detailed datasets on AVs further complicates the characterization of these crash dynamics, especially in a mixed-traffic environment. This study bridges these gaps and tackles the problem of rear-end collision in AVs by using the NHTSA crash data, which includes 2184 crashes between ADAS and ADS vehicles. To give a comprehensive overview, three distinct models were utilized, focusing on ADS and ADAS-controlled vehicles separately. The analysis shows that Pre-Crash Movements of both vehicles, Speed Gap Ratio and Roadway Description are the key predictors of rear-end collisions in both systems. ADS vehicles showed greater sensitivity to roadway surface, missing road markings and larger vehicles, while ADAS systems were more influenced by the presence of other entities and traffic incidents.

Keywords: Rear-end crash, Autonomous Vehicles, ADAS, ADS, GLMM

---

1. Introduction

Even though the driver is considered one of the "best sensors in the car" and may prevent many crashes, driver error remains responsible for many road crashes. The National Motor Vehicle Crash Causation Survey analyzed a sample of U.S. crashes from 2005 to 2007, revealing that driver error was the leading factor in 94% of accidents (Singh, 2018). Each year, approximately 1.3 million people die in road accidents, and another 20 to 30 million are injured (WHO, 2018). Autonomous vehicles (AVs) have the potential to reduce road crashes due to their improved efficiency, decreased human errors and environmental benefits (Chen et al., 2022; Novat et al., 2023). Advanced Driver Assistance Systems (ADAS) and Automated Driving Systems (ADS) have gained more attention.

AVs operating at level 2 use ADAS for steering and braking, whereas levels 3 to 5 employ ADS for enhanced automation. At present, ADAS is the most widely

---

[*] Corresponding author: **Md Sifat Bin Siraj** (sifatbinsiraj@gmail.com)

implemented technology in intelligent traffic systems (Nandavar et al., 2023). Recently, companies like Tesla and Waymo have thoroughly tested and introduced Level 2 and Level 4 autonomous vehicles, significantly boosting the usage of AVs in the automotive sector. Tesla's Level 2 automation software has reached over 400,000 users in North America, accumulating 100 million miles on non-highways in 2022 (Tesla, 2022). Similarly, Waymo's Level 4 AVs have covered more than 2 million miles with approximately 700 vehicles as of 2021 (California DMV, 2021). Although ADAS and ADS aim to enhance traffic safety by reducing accidents caused by human error, crashes still happen. This raises concerns about the safety of these AVs. Between 2009 and 2015, Waymo's crash rate (incidents per million miles driven) was lower than that of human-operated vehicles (Teoh & Kidd, 2017). In contrast, another study shows higher crash rates for AVs (Favarò et al., 2017). The ADS-controlled vehicle currently being tested on roads generally shows a worse safety record compared to human drivers (Goodall, 2021; Schoettle & Sivak, 2015). Nevertheless, some findings contradict this indicating that AVs have a lower crash rate than human-driven vehicles (Dixit et al., 2016; Favarò et al., 2017).

A key limitation regarding these safety issues is the insufficient datasets that have detailed crash characteristics of ADS and ADAS-controlled vehicles. The publicly available dataset from the California Department of Motor Vehicles (DMV) and the National Highway Traffic Safety Administration (NHTSA) has paved the way for conducting statistical analysis on the crash characteristics of AVs (Boggs, Arvin, et al., 2020; *California DMV*, 2021; NHTSA, 2023). Table 1 shows the summary of previous studies. Research indicates that AVs are particularly prone to rear-end collisions, especially since most incidents happen at 10 miles per hour (Boggs, Wali, et al., 2020; Favarò et al., 2019). One study utilized decision tree and association rule methods on CA DMV crash datasets and found that rear-end crashes were the most common, predominantly occurring at intersections engaged in autonomous mode (Ashraf et al., 2021). Using classification tree (Dadvar & Ahmed, 2021), Bayesian networks (Kutela et al., 2022), studies found that road surface condition, road type, incident time, crash location, and the speed limit at the site were the factors that impacted the rate and outcomes of ADS-involved crashes and mostly encountered collision was rear-end collisions (Boggs, Wali, et al., 2020). Compared to ADS-controlled vehicles, ADAS-controlled vehicles may demonstrate different behaviors since humans are involved in the control loop. There has been relatively less research on the factors influencing crashes for ADAS-controlled systems compared to ADS vehicles. Ding used a multinomial logit model and found that most ADAS crashes occurred on highways, and ADS occurred in urban places. Weather conditions, types of drivers and levels of automation were key factors that contributed to increasing crash severity (Ding et al., 2024; Dong et al., 2024).

Most of the previous literature on AVs demonstrate that these vehicles are prone to rear end crashes. However, existing studies often rely on limited datasets (or small sample sizes) and primarily examine determinants of overall crash occurrence, rather than focusing specifically on rear-end collisions. Additionally, prior research regarding AVs was fragmented and did not include both ADS and ADAS, which operated at a single level of automation, or those that included both, their focus was on general crash findings. Most of those studies have not explicitly identified the root causes and the factors influencing rear-end crashes involving ADS and ADAS-controlled vehicles in mixed traffic environments. Therefore, to bridge the gaps in the literature, research must be conducted thoroughly with more data regarding ADAS and ADS-controlled vehicle crashes to identify the factors that specifically influence rear-end collisions. This should include an overall comparison of how these differ between ADAS and ADS-controlled vehicles, which factors may lead to rear end crashes, as well as the overall crash patterns of both types of vehicles. Keeping that in mind, the objectives of this study are:

1. To identify the factors that cause rear-end collisions in ADS and ADAS-controlled vehicles in mixed traffic environments and compare the overall crash scenario
2. To explore the crash characteristics of these incidents and understand the contributing factors affecting such vehicle crashes.

Table 1: Previous Literature Summary

Literature	Dataset	Number of Data and Vehicle Type	Method
(Esenturk et al., 2021)	UK STATS19 Accident Database (2016-2018)	389,238 Conventional Road Crashes	Logistic Regression
(Xu et al., 2019)	CA DMV Crash Database (January 2015- June 2018)	72 Connected and Autonomous Vehicle (CAV) Crashes	Bootstrap-based Binary Logistic Regression
(Fu et al., 2025)	CA DMV and NHTSA Crash Database (January2015-February2022)	680 ADAS & ADS Crashes	Mixed Multinomial Logit Model
(Novat et al., 2023)	CA DMV and Transportation Injury Mapping System (TIMS) (2017-2020)	127 AV and 865 Conventional Crashes	Bayesian Network
((Yan et al., 2024)	NHTSA Crash Database (July 2021- November 2022)	202 ADAS & ADS Crashes	Random Forest, Logistic Regression
(Ashraf et al., 2021)	CA DMV Crash Database ((January 2016- February 2020)	198 AV Crashes	Decision Tree and Association Rule
(Huang et al., 2024)	NHTSA Crash Database (June 2021-May 2022)	214 ADAS & ADS Crashes	Binomial Logistic Regression
(Ding et al., 2024)	CA DMV (2014-March 2023), NHTSA, News Report	1280 ADAS & ADS Crashes	Random Parameter Logit Model
(Dong et al., 2024)	Autonomous Vehicle Operation Incident Dataset (AVOID) (2014- March 2023)	1280 ADAS & ADS Crashes	Two-mode Social Network and N-K Model
(Liu et al., 2021)	CA DMV (2014-2020) NHTSA(Conventional database)	122 AV and 2084 Conventional Crashes	NHTSA's Pre-crash Scenario Typology and Statistical Comparison
(Lee et al., 2023)	Autonomous Crash Report in CA (2019-2021)	260 AV Crashes	Path analysis framework and frequentist and Bayesian method

2. Methodology

2.1 Description of Datasets

Since July 2021, automobile manufacturers or operators must report crash incidents related to Society of Automotive Engineers (SAE) level-2 ADAS and SAE level (3-5) ADS systems as mandated by NHTSA (NHTSA, 2021). As of July 1, 2024, a total of 2,810 crashes involving ADAS and 1,305 crashes involving ADS have been reported by entities in the NHTSA dataset. NHTSA released the summarized crash report data in the format of .csv files, with 122 variables containing all information covered in the incident report form. The crash report details the subject vehicle (SV), including make, model, year, mileage, and incident information like date, time, location, city, roadway type, surface condition, speed limit, weather, and lighting. It describes the crash scene, the crash partner's (CP) role, the SV's movements and speed, contact areas, and injury severity.

2.2 Data Preprocessing

The modeling framework followed in this study is shown in Figure 1. The original dataset included various versions of a single crash, alongside incidents where necessary information was absent, such as null entries. Additionally, some crashes were withheld for confidentiality or business reasons and were labelled as 'Narrative CBI'. Moreover, certain narratives lacked adequate details about the crash progression.

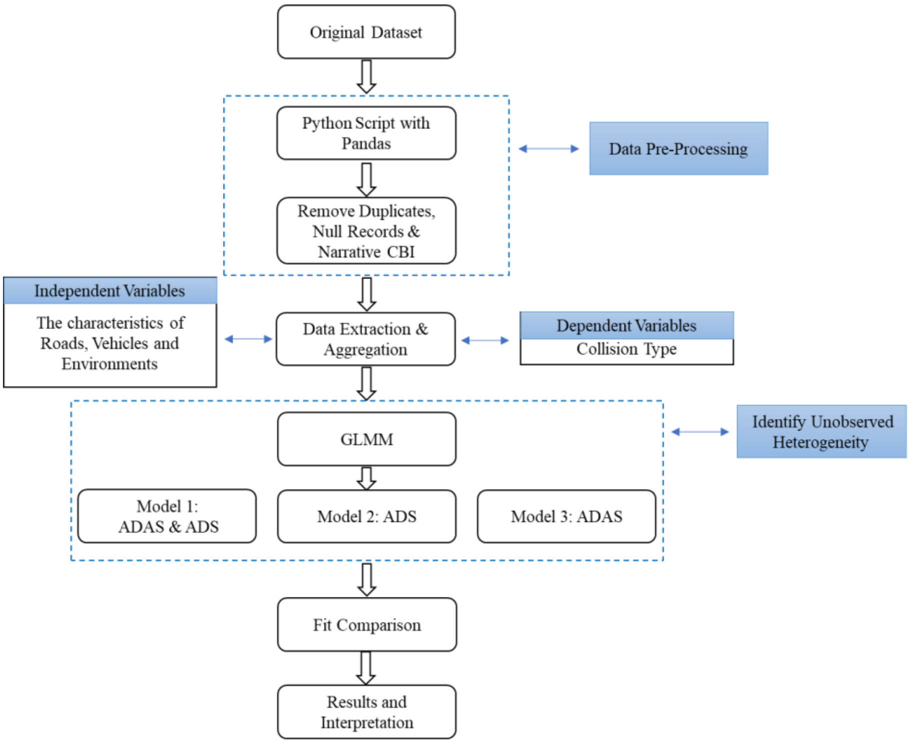

Figure 1: Study Work Flow

To remove duplicates, null entries, crashes with confidential information, and those that lack sufficient details about the contact area of CP and SV, a Python script based on the Pandas library was used. Following this pre-processing, a greater amount of data was filtered from both ADAS and ADS-controlled vehicles. Out of 2,810 ADAS-involved crashes, 1,209 and out of 1,305 ADS-involved crashes, 975, total 2184 ADAS and ADS involved crashes were kept for further analysis. Not all 122 variables are connected to crash patterns or significant for identifying the influential factors behind specific crash patterns. Therefore, the most important variables were kept for further analysis. Followed by data pre-processing, the framework follows data-extraction, data-aggregation, and data modelling, which includes three models: one where both ADAS and ADS data have been utilized, a second where ADS data has been used and a third where ADAS data has been employed. After the models fit comparison, the relevant findings regarding rear-end crashes have been presented.

2.2 Feature Extraction & Aggregation

The NHTSA's raw dataset included 122 variables for each crash report. From the original dataset (122 variables per crash report), 14 candidate explanatory variables were extracted that (i) represent pre-crash driving context and vehicle-roadway interactions relevant to rear-end collisions and (ii) have been identified as significant predictors of crash occurrence/patterns in prior studies (Dadvar & Ahmed, 2021). From the 14 variables that were considered for the analysis, 9 categorical variables were aggregated into a more meaningful form (Huang et al., 2024). These variables are described here.

Collision Type (CT) was labelled into two types. Rear-end vs Others. Vehicle Type (VT) – ADAS & ADS (Huang et al., 2024). Incident Time (IT) was divided into three categories: Morning-Noon (4:00 – 12:00), Noon-Night (12:00 – 20:00) and Night (20:00 – 4:00) (NSC, 2023). Roadway Type (RT) of five categories.

Table 2 : Aggregated Categorical Variables

ADAS- Equipped Vehicle Crashes Total: 1209	ADS- Equipped Vehicle Crashes Total: 975
After Aggregation (frequency, percentage)	
Roadway Type (RT)	
Highway/Freeway (n=737, 60.96%)	Highway/Freeway (n=62, 6.36%)
Intersection (n=143, 11.8%)	Intersection (n=435, 44.62%)
Street (n=174, 14.39%)	Street (n=431, 44.21%)
Parking Lot (n=3, 0.24%)	Parking Lot (n=44, 4.51%)
Others (n=212, 17.53%)	Others (n=3, 0.31%)
Roadway Surface (RS)	
Dry (n=828, 68.48%)	Dry (n=915, 93.84%)
Wet (n=193, 15.96%)	Wet (n=55, 5.64%)
Others (n=188, 15.55%)	Others (n=5, 0.51%)
Lighting (LT)	
Daylight (n=574, 47.47%)	Daylight (n=549, 56.31%)
Dark - Lighted (n=256, 21.17%)	Dark - Lighted (n=375, 38.46%)
Dark - Not Lighted (n=135, 11.16%)	Dark - Not Lighted (n=19, 1.95%)
Others (n=244, 20.18%)	Others (n=32, 3.28%)
Weather (WT)	
Clear (n=646, 53.43%)	Clear (n=802, 82.26%)
Cloud (n=187, 15.47%)	Cloud (n=117, 12.00%)
Rain (n=176, 14.56%)	Rain (n=40, 4.10%)
Adverse (n=12, 0.99%)	Adverse (n=1, 0.10%)
Others (n=188, 15.55%)	Others (n=15, 1.54%)
Crash With (CW)	
Passenger Car (n=214, 17.71%)	Passenger Car (n=403, 41.33%)
SUV (n=132, 10.92%)	SUV (n=134, 13.74%)
Larger Vehicles (n=138, 11.41%)	Larger Vehicles (n=215, 22.05%)
Fixed Objects (n=318, 26.30%)	Fixed Objects (n=23, 2.35%)
Others (n=408, 33.75%)	Others (n=163, 16.72%)
Pre-crash Movement of Crash Partner (PMCP)	
Forward Motion (n=161, 13.32%)	Forward Motion (n=474, 48.61%)
Lane Change Maneuvers (n=88, 7.28%)	Lane Change Maneuvers (n=100, 10.26%)
Turn Maneuvers (n=66, 5.46%)	Turn Maneuvers (n=96, 9.85%)
Stopped/Slow (n=68, 5.62%)	Stopped/Slow (n=163, 16.72%)
Non-Motorist (n=4, .33%)	Non-Motorist (n=8, 0.82%)
Others (n=822, 67.99%)	Others (n=134, 13.74%)
Pre-crash Movement of Subject-vehicle (PMSV)	
Forward Motion (n=914, 75.59%)	Forward Motion (n=398, 40.82%)
Lane Change Maneuvers (n=78, 6.45%)	Lane Change Maneuvers (n=37, 3.80%)
Turn Maneuvers (n=23, 1.90%)	Turn Maneuvers (n=84, 8.62%)
Stopped/Slow (n=12, 0.99%)	Stopped/Slow (n=424, 43.49%)
Others (n=182, 15.06%)	Others (n=32, 3.28%)

Pre-Crash Movement of CP (PMCP) & SV (PMSV) are defined as the movement of CP & SV before the crash occurred. The rest of the features are Roadway Surface (RS), Lighting (LT), Weather (WT), Injury Severity (IS), Roadway Description (RD), Property Damage (PD), and Crash With (CW). Table *2* displays the number of crashes and their corresponding percentage in brackets for ADAS and ADS.

To measure the impact of speed on crashes involving ADS or ADAS, a new variable called the Speed Gap Ratio (SGR) was created, as shown in sotto.

$$SGR = \frac{PSL - SVPCS}{PSL} \qquad \text{(Equation 1)}$$

where PSL represents the posted speed limit (in mph) on the roadway where the incident took place; SVPCS, as outlined in the NHTSA dataset, refers to the speed (in mph) of the SV at the moment of the incident. SGR is defined as a normalized deviation of the subject vehicle’s speed from the posted speed limit, using PSL as a consistent roadway-specific benchmark for the speed environment (Huang et al., 2024). Dividing by PSL produces a dimensionless measure that enables comparison across crashes occurring on roads with different speed limits. In this formulation, SGR > 0 indicates the subject vehicle was traveling below the posted speed limit, SGR = 0 indicates travel at the posted speed limit, and SGR < 0 indicates travel above the posted speed limit, larger absolute values indicate greater proportional deviation from PSL.

2.3 Generalized Linear Mixed Model (GLMM)

The generalized linear mixed model (GLMM) is now widely recognized in the literature for modeling the uncertainty linked to performance judgments derived from performance indicators. (Goldstein & Spiegelhalter, 1996; Morris & Christiansen, 1996). GLMMs include ‘random effects’ alongside the fixed parameters in traditional generalized linear models. The model recognizes a variance component, representing a source of uncertainty related to feature-specific performance effects, potentially arising from unmeasured factors like varying exposure, different recording practices, data quality and more (Bailey & Hewson, 2004). GLMM can effectively handle unobserved heterogeneity among observations by assuming a continuous distribution for random parameters (Anastasopoulos & Mannering, 2009).

Three models have been utilized for the analysis. Model 1 demonstrates the likelihood of rear-end collisions involving vehicles controlled by ADAS and ADS, highlighting the risk comparison between these two types of autonomous vehicles. Model 2 focuses on examining the factors that contribute to rear-end collisions in ADAS, while Model 3 focuses on ADS. In every model, the dependent variable was CT (Rear-end vs. Others). The other variables were independent, including two-way interactions among independent variables such as SGR and VT, PMCP and VT and SGW and WT for Model 1. In Models 2 and 3, the interaction terms involve independent variables except for VT, such as SGR and WT and SGR and RT. Independent variables are: VT, IT, RT, RS, LT, WT, CW, IS, RD, PD, PMCP and PMSV.

$$ln\left(\frac{p(CT=Rear-End)}{p(CT=Others)}\right) = \beta_0 + \beta_1 * VT + \beta_2 * IT + \beta_3 * RT + \beta_4 * RS + \beta_5 * LT + \beta_6 * WT + \beta_7 * CW + \beta_8 * IS + \beta_9 * RD + \beta_{10} * PD + \beta_{11} * SGR + \beta_{12} * PMCP + \beta_{13} * PMSV + \sum \beta_{\frac{VT}{x}} * \left(\frac{VT}{X}\right) + \sum \beta_{\frac{M}{N}} * \left(\frac{M}{N}\right) \qquad \text{(Equation 2)}$$

Mode 1 follows the sopra **Equation 2**,

where $X \in \{IT, RT, RS, LT, WT, CW, IS, RD, PD, SGR, PMCP, PMSV\}$,

(VT/X) denotes the two-way interaction between VT and other independent variables such as PMCP x VT, SGR x VT,

(M/N) denotes the interaction effect between other independent variables excluding VT, such as SGR x WT, SGR x RT.

The equations for Model 2 (ADAS-controlled) and Model 3 (ADS-controlled) are presented **sotto**:

$$ln\left(\frac{p(CT=Rear-End)}{p(CT=Others)}\right) = \beta_0 + \beta_1 * IT + \beta_2 * RT + \beta_3 * RS + \beta_4 * LT + \beta_5 * WT + \beta_6 * CW + \beta_7 * IS + \beta_8 * RD + \beta_9 * PD + \beta_{10} * SGR + \beta_{11} * PMCP + \beta_{12} * PMSV + \sum \beta_{\frac{M}{N}} * (\frac{M}{N}) \qquad \text{(Equation 3)}$$

In the equation **sopra,**

$$\{M, N\} \in \{IT, RT, RS, LT, WT, CW, IS, RD, PD, SGR, PMCP, PMSV\}$$

which denotes the two-way interaction effects, such as SGR x WT, SGR x RT, IT x WT etc.

These three models were developed using GLMM in SPSS software. Model fit was assessed using information criteria based on the -2log likelihood, with both Akaike Corrected and Bayesian information criteria reported. As noted in the output, models with smaller information criterion values fit better, confirming the suitability of our final model specification.

3. Results

3.1 Crash Distribution

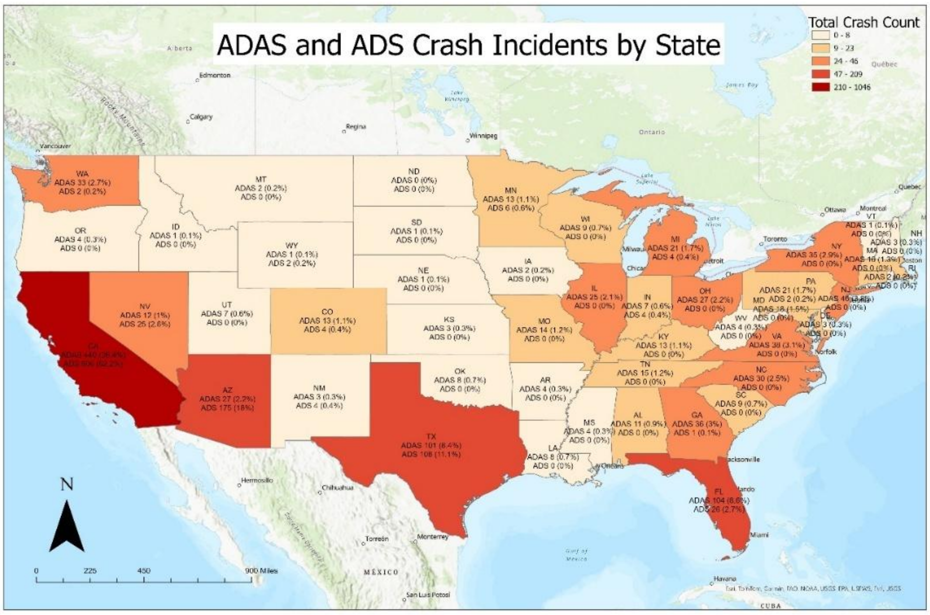

Figure 2 : Crash Distribution by States

Figure 2 presents the state-wise distribution of ADAS and ADS crashes across the USA. The map was produced in ArcGIS Pro using a graduated color (choropleth) symbology of the state 'total crash count' (ADAS + ADS) with five classes from light beige (0–8 crashes) to dark red (210–1046 crashes). The class intervals were generated automatically using the default Natural Breaks (Jenks) classification in ArcGIS Pro meaning the breakpoints were computed from the observed distribution of crash totals to group similar values and separate dissimilar values, which is appropriate here because the crash totals are strongly right-skewed (many states with low counts and a few states with very high counts). Each state label shows the ADAS and ADS crash counts separately and the percentages in parentheses indicate that state's share of the national ADAS total and the national ADS total, respectively. High-count states include California (Grand Total = 1046), Texas (Grand Total = 209), Arizona (Grand Total = 202), and Florida (Grand Total = 130), while moderate totals include New Jersey (46), Virginia (38), Nevada (37), Georgia (37), Washington (35), and New York (35). In contrast, very low totals appear in states such as Hawaii (0) and North Dakota (0), as well as states with only 1–3 crashes (e.g., Idaho, Nebraska, Vermont, and Wyoming), highlighting the geographic concentration of reported ADAS and ADS-related crashes across the country.

3.2 Statistical Analysis of Crash Factors

Table 3 below shows the collision type predictors for three models that have been analysed in this study. There are several factors that greatly contribute to the occurrence of collisions and there are differences between the two types of systems. For RS (Roadway Surface), a significant positive relationship is found for ADAS ($\beta = 3.034$, $p = 0.048$) indicating that certain roadway surfaces increase the odds of colliding, whereas the relationship for ADS is not significant ($p = 0.169$). Likewise, CW (Crash with) exhibits a pronounced and consistent positive effect in both ADS and ADAS models, with coefficients of 9.914 ($p < 0.001$) and 11.351 ($p < 0.001$), respectively, demonstrating that other entities being present in crashes significantly heighten the risk of collision. Another powerful predictor of collisions is the Pre-Crash Movements of the Crash Partner (PMCP), with coefficients of 25.549 ($p < 0.001$) for ADAS and 19.356 ($p < 0.001$) for ADS, emphasizing that the represented behaviors of the crash partner before the crash are the most important predictor of such events in either system.

Table 3 : F-Statistics for Three types of analysis

Variables	**(ADS & ADAS)**		**ADAS**		**ADS**	
	F-Value	**P-Value**	**F-Value**	**P-Value**	**F-Value**	**P-Value**
Corrected Model	12.147	0.000 ***	3.92	0.000 ***	6.032	0.000 ***
RS	3.034	.048 **	1.78	0.169	5.021	0.007 ***
CW	9.914	0.000 ***	11.351	0.000 ***	3.261	0.011 **
PMCP	25.549	0.000 ***	4.583	0.000 ***	19.356	0.000 ***
PMSV	4.866	0.001 ***	2.114	0.077 *	4.921	0.001 ***
IS	5.689	0.000 ***	1.262	0.278	5.439	0.000 ***
RD	3.85	0.004 ***	3.755	0.005 ***	2.859	0.036 **
PD	3.614	0.027 **	4.713	0.009 ***	1.084	0.339
SGR	15.638	0.000 ***	3.911	0.048 **	7.815	0.005 ***
SGR_RT	1.413	0.235	0.011	0.915	2.792	0.095 *
SGR_VT	15.701	0.000 ***	-	-	-	-
PMCP_VT	2.687	0.020 **	-	-	-	-
VT_IS	3.485	0.008 ***	-	-	-	-
SGR_WT	1.308	0..253	3.936	0.047 **	0.981	0.322
PMSV_VT	1.899	0.108	-	-	-	-
VT	0.044	0.834	-	-	-	-
WT	1.802	0.126	1.937	0.102	2.884	0.022 **
RT	1.384	0.246	7.61	0.000 ***	1.106	0.346
LT	0.59	0.622	0.739	0.529	1.415	0.237
Akaike Corrected	9202.728		6058.638		4761.509	
Bayesian Information Criteria	9208.386		6063.697		4766.343	
-2 log likelihood	9200.726		6056.635		4759.505	
*** p<.01, ** p<.05, * p<.1						

The Pre-Crash Movements of the Subject Vehicle (PMSV), affecting collision outcomes, are also registered as significant for the both systems, leading to a positive coefficient for ADS & ADAS model 4.866 ($p =0.001$), for ADAS 2.114 ($p < 0.077$) and for ADS 4.921 ($p = 0.001$). It indicates that maneuvers by a vehicle leading up to the crash are among the strongest predictors. RD (Roadway Description) was also a significant contributor to both models having a positive effect on collision risk (ADS & ADAS $\beta = 3.850$, $p = 0.004$, ADAS $\beta = 3.755$, $p = 0.005$; ADS $\beta = 2.859$, $p = 0.036$)

additionally reinforcing the importance of the roadway conditions in collision risk. SGR_VT (Speed Gap Ratio x Vehicle Type) was found highly statistically significant ($\beta = 15.701$, $p < .001$), suggesting a substantial increase in the likelihood of rear-end crashes as SGR changes depending on the vehicle type. The β and p-value of SGR for ADAS were 3.911 and .048, while for ADS, they were 7.815 and .005. This shows that changes in SGR significantly increase the likelihood of rear-end crashes in ADS-controlled vehicles. The VT_IS (Vehicle Type x Injury Severity) interaction term was also found to be significant, indicating that the particular vehicle type being involved in rear-end crashes differs significantly across injury severity levels. For ADS, 'IS' was determined to be $\beta = 5.439$ and $p < .001$, indicating that the probability of a rear-end collision involving an ADS-controlled vehicle varies significantly across different levels of injury severity. PD (Property Damage) was found to be statistically significant (for ADAS, $\beta = 4.713$, $p = .009$), suggesting that rear-end crashes involving ADAS differ in terms of property damage. RT (Roadway Type) is significant in ADS ($\beta = 7.610$, $p < 0.001$) but it does not affect ADAS. The interaction term SGR_WT (Speed Gap Ratio x Weather) exhibited strong significance for the ADAS, with $\beta = 3.936$ and $p = 0.047$. This indicates that the speed deviations heightening the chances of rear-end collisions vary significantly under different weather conditions. WT (Weather) was found to have effects on the probability of ADS vehicles ($\beta = 2.884$, $p = .022$) facing rear-end crashes. This capture of potential complex inter-relationships between roadway conditions, vehicle behaviors and environmental factors impacting crashes in ADS and ADAS systems is just one set of findings from this research analysis.

Table 4 below displays the fixed effects parameter estimates for Model 1, which examines the predictors of collision types in the context of both Automated Driving Systems (ADS) and Advanced Driver Assistance Systems (ADAS).

Table 4 : Parameter Estimates for Model 1 (ADS & ADAS)

Variables	**β**	**Std. Error**	**t**	**P-Value**	**95% Confidence Interval**		**Exp. (β)**	**95% Confidence Interval for Exp. (Coefficient)**	
					Lower	**Upper**		**Lower**	**Upper**
Intercept	-2.41	2.184	-1.1	0.27	-6.694	1.873	0.09	0.001	6.508
RS (Others)	1.205	0.49	2.461	0.014 **	0.245	2.166	3.338	1.277	8.725
WT (Rain)	1.692	0.788	2.146	0.032 **	0.146	3.237	5.428	1.157	25.465
CW (Others)	1.023	0.196	5.213	.000 ***	0.638	1.408	2.781	1.893	4.086
CW (Larger Vehicles)	0.808	0.255	3.17	.002 ***	0.308	1.308	2.243	1.361	3.698
PMCP (Forward Motion)	-3.01	0.902	-3.34	.001 ***	-4.778	-1.24	0.049	0.008	0.289
PMCP (Turn Maneuvers)	-2.51	0.931	-2.69	.007 ***	-4.334	-0.681	0.081	0.013	0.506
PMSV (Forward Motion)	0.757	0.216	3.499	.000 ***	0.333	1.181	2.131	1.395	3.256
PMSV (Lane Change Maneuvers)	1.298	0.442	2.936	.003 ***	0.431	2.166	3.663	1.539	8.721
IS (Minor)	-1.08	0.282	-3.84	.000 ***	-1.632	-0.528	0.34	0.196	0.59
IS (Serious)	-1.84	0.861	-2.13	.033 **	-3.528	-0.149	0.159	0.029	0.862
RD (Traffic Incident)	1.786	0.527	3.39	.001 ***	0.753	2.819	5.963	2.123	16.754
SGR	-2.24	0.568	-3.96	.000 ***	-3.357	-1.131	0.106	0.035	0.323
SGR_VT	1.838	0.464	3.962	.000 ***	0.928	2.748	6.285	2.53	15.609
PMSV_VT (Others)	2.024	0.865	2.34	.019 **	0.328	3.72	7.568	1.388	41.272

VT_IS (Minor)	1.619	0.647	2.503	.012	**	0.35	2.887	5.047	1.419	17.943

*** p<.01, ** p<.05, * p<.1

This model classifies the kind of collisions, with a dependent variable for which some key variables play a significant role. RS (Others- snow / Slush and unknown conditions) has a positive and significant coefficient (β = 1.205, p = 0.014). Thus, there is 3.338 times higher Exp (β) for collisions occurring on such surfaces. Likewise, WT (Rain), signaling rainy weather, is significantly predictive of collision risk (β = 1.692, p = 0.032), with an Exp (β) of 5.428, which underscores the high likelihood of the collision event occurring under rainy weather.

Collisions in terms of CW (Others- animal, pedestrian, cyclist, etc.) significantly increased the likelihood of collisions (β = 1.023, p < 0.001). Also, CW (Larger Vehicles), which shows a positive association (β = 0.808, p = 0.002), indicates that the larger vehicles are engaged in rear-end collisions, with an Exp (β) of 2.243. The negative coefficients for PMCP (Forward Motion) and PMCP (Turn Maneuvers), which are -3.009 and -2.508, respectively, indicate that the probability of collision is less likely in this dataset when certain driving parameters, such as forward motion and turning, are present in the pre-crash state. Whereas the PMSV (Forward Motion) and PMSV (Lane Change Maneuvers) variables demonstrate positive coefficients, suggesting that they increase collision risk (β = 0.757 and β = 1.298, p < 0.01), with Exp (β) of 2.131 and 3.663, respectively.

Regarding injury severity, IS (Minor) and IS (Serious) are inversely associated with collision likelihood, suggesting that minor or serious injuries may be related to a decrease in rear-end collision probability, with coefficients of -1.080 and -1.838, correspondingly. In terms of RD (Roadway Description), ‘Traffic Incident’ significantly increases the likelihood of collision (β = 1.786, p = 0.001) with an Exp (β) of 5.963, indicating that rear-end collisions mostly occur on roads with past traffic incident occurrences. The SGR exhibits a significant negative correlation which denotes that decreased SGR significantly increases the likelihood of rear-end crashes. The lower gap ratio indicates that the speed of the SV is close to or exceeds the speed limit, which refers to the high-speed situation, ultimately increasing the probability of a rear-end crash. The SGR_VT, PMSV_VT, VT_IS interaction terms exhibit significant positive correlations with collision occurrences, which suggests that the combined effect of two variables (i.e, SGR with VT) increases the likelihood of rear-end crashes.

Table 5 below shows how various predictors influence the probability of rear-end collisions in ADS-controlled vehicles. The probability is associated with important features, such as road conditions, vehicle maneuvers, and environmental factors. For RS (Dry), meaning dry road conditions, the coefficient is highly positive (β = 2.697, p = 0.006) suggesting that dry roads are associated with a higher risk of collisions, with an Exp (β) of 14.832. This means that under dry road conditions, collisions can actually increase due to higher speeds or vehicle performance. In contrast, LT (Dark Not Lighted, reflecting low light levels on-road) has a slightly negative coefficient (β = -1.302, p = 0.056), denoting that lack of lighting may explain a lower probability of some collision types. The potential rationale may be that the vehicle’s autonomous system operates effectively under such conditions due to the heightened awareness of increased crash risk in dark environments. The model further indicates each pre-crash movement (Forward Motion, β= 0.902 Lane Change Maneuvers, β= 1.328) as having a strong positive relationship to collision occurrence, confirming that forward motion and lane changes are key behaviors in the collision prediction space. Alternatively, IS (Minor) was associated with minor injuries and had a negative coefficient of -1.291 (p < 0.001), indicated that minor injury types were negatively associated with the probability of rear-end collisions.

Table 5 : Parameters Estimates for Model 2 (ADS)

Variables	**β**	**Std. Error**	**t**	**P-Value**	**95% Confidence Interval**		**Exp. (β)**	**95% Confidence Interval for Exp. (Coefficient)**	
					Lower	**Upper**		**Lower**	**Upper**
Intercept (Collision Type)	1.661	3..1158	0	0.57	3.948	3.626	0.001	0	.
RS (Dry)	2.697	0.9703	2.779	.006***	0.793	4.601	14.832	2.209	99.57
LT (Dark Not Lighted)	-1.3	0.6802	-1.92	.056 **	-2.637	0.032	0.272	0.072	1.033
PMSV (Forward Motion)	0.902	0.2377	3.793	.000***	0.435	1.368	2.464	1.545	3.92
PMSV (Lane Change Maneuvers)	1.328	0.5121	2.594	.010 **	0.323	2.333	3.774	1.382	10.31
IS (Minor)	-1.29	0.3536	-3.65	.000***	-1.985	-0.597	0.275	0.137	0.55
RD (Missing/ Degraded Markings)	3.26	1.1376	2.865	.004***	1.027	5.492	26.037	2.792	242.78
SGR	-4.4	1.5736	-2.8	.005***	-7.487	-1.311	0.012	0.001	0.27
SGR_RT	0.451	0.2698	1.671	.095 *	-0.079	0.98	1.57	0.924	2.665

**** p<.01, ** p<.05, * p<.1*

The term RD (Missing/Degraded Markings) is also significantly positive (β=3.260; p=0.004), indicating that missing or degraded road markings increase the odds of a crash, which highlights the importance of having clearly marked and well-maintained road markings to prevent accidents. Moreover, the fact that SGR (Speed Gap Ratio) (β = -4.399, p = 0.005) has a significantly negative coefficient further reinforces the inherent association between increased speed gap ratio of cars and reduced collisions. When SGR decreases, it indicates that the vehicle was traveling at a significantly lower speed than the posted limit. This suggests that the higher the speed, the lower the ratio, which presents the most vulnerable conditions for rear-end crashes. Going at a fast pace can increase the probability of a rear-end crash. The analysis reveals the intricate interplay of vehicle behavior, temperature variance, and road conditions in predicting collision probability.

Table 6 below shows factors that influence rear-end crashes in the context of ADAS-controlled vehicles. Several roadway and vehicle-related variables significantly affect the likelihood of rear-end crashes. The RT, 'Others' which includes rural road, traffic circle and unknown (β = -0.936, p.001), shows the decreased collision rate for ADAS. On the contrary roadway surface, 'Others' increased the risk of rear-end collisions in ADAS with Exp. (β) = 3.222). Moreover, the SGR (Speed Gap Ratio) has a considerable negative impact on collisions (β = -1.100, p = 0.048), revealing that the higher the ratio, the less likely the crash will occur. There are other behavioral aspects that have a considerable effect on the outcome of a collision. CW (Others) refers to crashes with animals, motorcycles, pedestrians, unknowns etc. has a strong positive association (β = 1.640, p < 0.001) with an increased odds of 5.155, meaning such things influence vehicles to be rear-ended. Significant positive coefficients also appear in PMSV (Forward Motion) (β=

1.995, p= 0.015) and PMSV (Lane Change Maneuvers) (β= 2.241, p= 0.010), indicating that these pre-crash driving maneuvers are strongly correlated to a higher risk of collision.

Table 6 : Fixed Parameter Estimates for Model 3 (ADAS)

Variables	**β**	**Std. Error**	**t**	**P-Value**	**95% Confidence Interval**		**Exp. (β)**	**95% Confidence Interval for Exp. (Coefficient)**	
					Lower	**Upper**		**Lower**	**Upper**
Intercept	-1.44	3.6907	-0.39	0.696	-8.684	5.798	0.236	0	329.67
RT (Others)	-0.94	0.2839	-3.3	.001 ***	-1.493	-0.379	0.392	0.225	0.685
RS (Others)	1.17	0.676	1.731	.084 **	-0.156	2.496	3.222	0.855	12.138
CW (Others)	1.64	0.2563	6.399	.000 ***	1.137	2.143	5.155	3.118	8.523
PMSV (Forward Motion)	1.995	0.8213	2.429	.015 **	0.384	3.607	7.354	1.468	36.844
PMSV (Lane Change Maneuvers)	2.241	0.8723	2.569	.010 **	0.529	3.952	9.402	1.698	52.061
RD (Traffic Incident)	2.273	0.772	2.945	.003 ***	0.759	3.788	9.71	2.135	44.158
SGR	-1.1	0.5561	-1.98	.048 **	-2.191	-0.009	0.333	0.112	0.991
SGR_WT	0.481	0.2426	1.984	.047 **	0.005	0.957	1.618	1.005	2.604

.*** *p<.01, ** p<.05, * p<.1*

RD (Traffic Incident) is another important variable with a coefficient of 2.273 (p = 0.003), indicating that roads with past traffic incidents represent a significant risk factor for collision occurrence with an odds ratio of 9.710. SGR_WT (Speed Gap Ratio_Weather), finally, has a substantial and positive effect (β = 0.481, p = 0.047), meaning the combination of speed gaps and weather conditions increases the probability of a collision. Across the three model specifications, several predictors repeatedly emerged as robust determinants of rear-end collision outcomes. Pre-crash movements of the crash partner and the subject vehicle (PMCP and PMSV) and the Speed Gap Ratio (SGR) were consistently significant, indicating that rear-end crashes in mixed traffic are strongly linked to interaction dynamics and the prevailing speed environment. Roadway description variables particularly traffic-incident-related contexts and crash-partner involvement (CW) also showed strong associations, while system-specific effects suggest ADS outcomes are more sensitive to infrastructure conditions such as missing/degraded markings, whereas ADAS outcomes exhibit stronger dependence on contextual and weather-related interactions.

4. Conclusion

In this research, 2184 rear-end crash data points were analyzed, making it the most extensive dataset utilized to date for ADS and ADAS crash analysis.

Road surface conditions substantially impact collision likelihood, with differing effects between systems. For ADAS, certain roadway surfaces (especially "Others" category including snow/slush) increase collision odds while for ADS, dry road conditions exhibited a more potent effect. Rainy conditions significantly increased collision probabilities, particularly affecting ADS systems. Roads with traffic incidents or missing/degraded markings showed strong associations with collisions across both systems.

The movements of both the subject vehicle (PMSV) and crash partner (PMCP) emerged as the most significant predictors of collisions. Forward motion and lane change maneuvers of SV significantly increased collision risk for both systems. Speed Gap Ratio (SGR) was a significant predictor in the combined and system-specific models, and its interactions (SGR×VT and SGR×WT) indicate that the effect of speed context differs by system type and weather conditions. Crash partner type (CW) was significant, showing that interactions with larger vehicles and 'other' entities (e.g., pedestrians/animals) are associated with higher rear-end collision likelihood in the models.

These findings suggest distinct priorities for the two systems. ADS outcomes appear more sensitive to roadway infrastructure quality and roadway surface conditions, implying that maintenance of lane markings and consistent roadway guidance is critical where higher automation is expected. In contrast, ADAS outcomes point to a need for operational deployment policies that manage mixed-traffic complexity and condition-dependent risk, such as restricting or warning against use in adverse weather and high-conflict environments, strengthening driver monitoring/engagement requirements, and applying geofenced or context-based activation (or reduced functionality) when traffic incidents or complex roadside entities are present.

This study used only the NHTSA crash dataset (July 2021–July 2024) and many records had missing or incomplete fields that were removed during preprocessing, reducing the final sample size. Using additional sources (e.g., CA DMV) and more complete reporting would likely improve the robustness and accuracy of the findings.

Reference

Anastasopoulos, P. Ch., & Mannering, F. L. (2009). A note on modeling vehicle accident frequencies with random-parameters count models. *Accident Analysis & Prevention*, *41*(1), 153–159. https://doi.org/10.1016/j.aap.2008.10.005

Ashraf, M. T., Dey, K., Mishra, S., & Rahman, M. T. (2021). Extracting Rules from Autonomous-Vehicle-Involved Crashes by Applying Decision Tree and Association Rule Methods. *Transportation Research Record*, *2675*(11), 522–533. https://doi.org/10.1177/03611981211018461

Bailey, T. C., & Hewson, P. J. (2004). Simultaneous Modelling of Multiple Traffic Safety Performance Indicators by Using a Multivariate Generalized Linear Mixed Model. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, *167*(3), 501–517.

Boggs, A. M., Arvin, R., & Khattak, A. J. (2020). Exploring the who, what, when, where, and why of automated vehicle disengagements. *Accident Analysis & Prevention*, *136*, 105406. https://doi.org/10.1016/j.aap.2019.105406

Boggs, A. M., Wali, B., & Khattak, A. J. (2020). Exploratory analysis of automated vehicle crashes in California: A text analytics & hierarchical Bayesian heterogeneity-based approach. *Accident Analysis & Prevention*, *135*, 105354. https://doi.org/10.1016/j.aap.2019.105354

California DMV. (2021). California DMV. https://www.dmv.ca.gov/portal/

Chen, F., Lu, G., Tan, H., Liu, M., & Wan, H. (2022). Effects of assignments of dedicated automated vehicle lanes and inter-vehicle distances of automated vehicle platoons on car-following performance of nearby manual vehicle drivers. *Accident Analysis & Prevention*, *177*, 106826. https://doi.org/10.1016/j.aap.2022.106826

Dadvar, S., & Ahmed, M. M. (2021). *California Autonomous Vehicle Crashes: Explanatory Data Analysis and Classification Tree* (No. TRBAM-21-04068). Article TRBAM-21-04068. Transportation Research Board 100th Annual MeetingTransportation Research BoardTransportation Research Board. https://trid.trb.org/View/1759180

Ding, S., Abdel-Aty, M., Barbour, N., Wang, D., Wang, Z., & Zheng, O. (2024). Exploratory analysis of injury severity under different levels of driving automation (SAE Levels 2 and 4) using multi-source data. *Accident Analysis & Prevention*, *206*, 107692. https://doi.org/10.1016/j.aap.2024.107692

Dixit, V. V., Chand, S., & Nair, D. J. (2016). Autonomous Vehicles: Disengagements, Accidents and Reaction Times. *PLOS ONE*, *11*(12), e0168054. https://doi.org/10.1371/journal.pone.0168054

Dong, S., Cui ,Zhiwei, Sun ,Daniel (Jian), Ye ,Liming, & and Ding, S. (2024). How do environmental and road factors impact automated vehicle crashes? An evidence from ADAS and ADS. *Journal of Transportation Safety & Security*, *0*(0), 1–29. https://doi.org/10.1080/19439962.2024.2415957

Esenturk, E., Khastgir, S., Wallace, A., & Jennings, P. (2021). Analyzing Real-world Accidents for Test Scenario Generation for Automated Vehicles. *2021 IEEE Intelligent Vehicles Symposium (IV)*, 288–295. https://doi.org/10.1109/IV48863.2021.9576007

Favarò, F. M., Eurich ,Sky O., & and Rizvi, S. S. (2019). “Human” Problems in Semi-Autonomous Vehicles: Understanding Drivers’ Reactions to Off-Nominal Scenarios. *International Journal of Human–Computer Interaction*, *35*(11), 956–971. https://doi.org/10.1080/10447318.2018.1561784

Favarò, F. M., Nader, N., Eurich, S. O., Tripp, M., & Varadaraju, N. (2017). Examining accident reports involving autonomous vehicles in California. *PLOS ONE*, *12*(9), e0184952. https://doi.org/10.1371/journal.pone.0184952

Fu, H., Ye, S., Fu, X., Chen, T., & Zhao, J. (2025). New insights into factors affecting the severity of autonomous vehicle crashes from two sources of AV incident records. *Travel Behaviour and Society*, *38*, 100934. https://doi.org/10.1016/j.tbs.2024.100934

Goldstein, H., & Spiegelhalter, D. J. (1996). League Tables and Their Limitations: Statistical Issues in Comparisons of Institutional Performance. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *159*(3), 385–409. https://doi.org/10.2307/2983325

Goodall, N. J. (2021). Comparison of automated vehicle struck-from-behind crash rates with national rates using naturalistic data. *Accident; Analysis and Prevention*, *154*, 106056. https://doi.org/10.1016/j.aap.2021.106056

Huang, C., Wen, X., & He, D. (2024). Characteristics of Rear-End Collisions: A Comparison between Automated Driving System-Involved Crashes and Advanced Driving Assistance System-Involved Crashes. *Transportation Research Record*, *2678*(7), 771–782. https://doi.org/10.1177/03611981231209319

Kutela, B., Avelar, R. E., & Bansal, P. (2022). Modeling Automated Vehicle Crashes with a Focus on Vehicle At-Fault, Collision Type, and Injury Outcome. *Journal of Transportation Engineering, Part A: Systems*, *148*(6), 04022024. https://doi.org/10.1061/JTEPBS.0000680

Lee, S., Arvin, R., & Khattak, A. J. (2023). Advancing investigation of automated vehicle crashes using text analytics of crash narratives and Bayesian analysis. *Accident Analysis & Prevention*, *181*, 106932. https://doi.org/10.1016/j.aap.2022.106932

Liu, Q., Wang, X., Wu, X., Glaser, Y., & He, L. (2021). Crash comparison of autonomous and conventional vehicles using pre-crash scenario typology. *Accident Analysis & Prevention*, *159*, 106281. https://doi.org/10.1016/j.aap.2021.106281

Morris, C. N., & Christiansen, C. L. (1996). Hierarchical Models for Ranking and for Identifying Extremes, with Applications[1]. In J. M. Bernardo, J. O. Berger, A. P. Dawid, & A. F. M. Smith (Eds.), *Bayesian Statistics 5: Proceedings of the Fifth Valencia International Meeting* (p. 0). Oxford University Press. https://doi.org/10.1093/oso/9780198523567.003.0015

Nandavar, S., Kaye, S.-A., Senserrick, T., & Oviedo-Trespalacios, O. (2023). Exploring the factors influencing acquisition and learning experiences of cars fitted with advanced driver assistance systems (ADAS). *Transportation Research Part F: Traffic Psychology and Behaviour, 94,* 341–352. https://doi.org/10.1016/j.trf.2023.02.006

NHTSA. (2023). *Crash Data Systems | NHTSA*. Crash Data Systems. https://www.nhtsa.gov/data/crash-data-systems

Novat, N., Kidando, E., Kutela, B., & Kitali, A. E. (2023). A comparative study of collision types between automated and conventional vehicles using Bayesian probabilistic inferences. *Journal of Safety Research, 84,* 251–260. https://doi.org/10.1016/j.jsr.2022.11.001

NSC. (2023). Car Crashes by Time of Day and Day of Week. *Injury Facts*. https://injuryfacts.nsc.org/motor-vehicle/overview/crashes-by-time-of-day-and-day-of-week/

Schoettle, B., & Sivak, M. (2015). *A PRELIMINARY ANALYSIS OF REAL-WORLD CRASHES INVOLVING SELF-DRIVING VEHICLES*. https://www.semanticscholar.org/paper/A-PRELIMINARY-ANALYSIS-OF-REAL-WORLD-CRASHES-Schoettle-Sivak/0c0d9cf05025a885c0c55c99fec862860410db3c

Singh, S. (2018). Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey. *Traffic Safety Facts - Crash Stats*, Article DOT HS 812 506. https://trid.trb.org/View/1507603

Teoh, E. R., & Kidd, D. G. (2017). Rage against the machine? Google’s self-driving cars versus human drivers. *Journal of Safety Research, 63,* 57–60. https://doi.org/10.1016/j.jsr.2017.08.008

Tesla. (2022). *Tesla Quarterly Update*. https://krungthai.com/Download/generalcontent/MediaFile_28748TeslaIncFinancialresultsforQ4andFY2022Shareholderdeck.pdf

WHO. (2018). *Global status report on road safety 2018*. https://www.who.int/publications/i/item/9789241565684

Wu, J., Flannagan, C., Sander, U., & Bärgman, J. (2024). *Model-based generation of representative rear-end crash scenarios across the full severity range using pre-crash data* (No. arXiv:2406.15538). arXiv. https://doi.org/10.48550/arXiv.2406.15538

Xu, C., Ding, Z., Wang, C., & Li, Z. (2019). Statistical analysis of the patterns and characteristics of connected and autonomous vehicle involved crashes. *Journal of Safety Research, 71,* 41–47. https://doi.org/10.1016/j.jsr.2019.09.001

Yan, S., Huang ,Chunxi, & and He, D. (2024). A comparison of patterns and contributing factors of ADAS and ADS involved crashes. *Journal of Transportation Safety & Security, 16*(9), 1061–1088. https://doi.org/10.1080/19439962.2023.2284175

Acknowledgments

The authors sincerely acknowledge the National Highway Traffic Safety Administration (NHTSA) for their dedicated efforts in collecting and releasing the raw data utilized in this study.