

A study on in-depth accident investigation of intelligent connected vehicles based on reverse scenario construction

Received: 14 May 2025

Accepted: 12 November 2025

Published online: 21 November 2025

Cite this article as: Zhang S., Li X., Zha Z. *et al.* A study on in-depth accident investigation of intelligent connected vehicles based on reverse scenario construction. *Sci Rep* (2025). <https://doi.org/10.1038/s41598-025-28747-3>

Shibo Zhang, Xin Li, Zhiyong Zha, Huiwen Xiong, Chengcheng Wu, Pingfei Li & Zhenyu Kang

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

A Study on In-depth Accident Investigation of Intelligent Connected Vehicles Based on Reverse Scenario Construction

Shibo Zhang^{*1,2,3}, Xin Li¹□Zhiyong Zha⁴, Huiwen Xiong¹,
Chengcheng Wu¹, Pingfei Li^{1,2}, Zhenyu Kang¹

1. School of Automobile and Transportation, Xihua University, Chengdu, 610039 China;

2. Sichuan Xihua Jiaotong Forensic Centre, Xihua University, Chengdu, 610039 China;

3. Provincial Engineering Research Centre for New Energy Vehicle Intelligent Control and Simulation Test Technology of Sichuan, Chengdu, 610039 China;

4. China Merchants Testing Vehicle Technology Research Institute Co., Ltd, Chongqing 401122, China

Correspondence to:

Shibo Zhang

School of Automobile and Transportation, Xihua University, Chengdu 610039, China;

E-mail: zhshbox@163.com

Phone: +86-028-87720534

A Study on In-depth Accident Investigation of Intelligent Connected Vehicles Based on Reverse Scenario Construction

Abstract □ To address the limitations of current in-depth investigation systems for intelligent connected vehicle accidents, this paper analyzes the complexity and characteristics of modern self-driving crashes using the 2021 Annual Report on Intelligent Connected Vehicles published by the California Department of Motor Vehicles. A closed-loop framework is proposed, integrating human-vehicle-road collaborative data collection, reverse scenario reconstruction, and accident causation analysis. At the technical level, high-precision, time-sequenced motion parameters are extracted from vehicle log data. A two-wheel vehicle dynamics model is established, and a 20-second pre-crash spatiotemporal trajectory reconstruction is developed using MATLAB. OpenStreetMap (OSM) data is integrated and converted to the Open DRIVE format via SUMO tools, enabling accurate reconstruction of road geometry. To handle heterogeneous multimodal log data, Python-based algorithms are employed for data cleaning and feature extraction, capturing lateral vehicle displacement and surrounding traffic dynamics. Finally, a parameterized model of the pre-collision dynamic scene is built using the Open SCENARIO standard, achieving high-fidelity accident reproduction. This method offers scientific support for liability determination, safety enhancement of autonomous systems, and the development of standardized evaluation frameworks.

Key words □ Intelligent Connected Vehicles □ In-Depth Accident Investigation □ Reverse Scenario Construction □ Driving Log Data □ Open DRIVE □ Open SCENARIO

1 Introduction

With the advancement of intelligent manufacturing, China has prioritized transportation development in its 14th Five-Year Plan, accelerating progress in autonomous driving technologies. As a pivotal transitional phase in the evolution of intelligent connected vehicles (ICVs), this sector has become a key indicator of national manufacturing strength and has witnessed rapid growth. According to the China Internet Development Report (2023), over 8 million new L2-level vehicles equipped with advanced driver-assistance systems were sold in China in 2022 alone¹.

As autonomous and human-driven vehicles are expected to coexist in the

foreseeable future, new challenges emerge in traffic accident investigations. ICVs generate extensive pre-collision data, offering precise and reliable information to support accident analysis. This data enables a comprehensive understanding of causal factors and facilitates a more accurate attribution of responsibility among ICV systems, human drivers, and manufacturers. Consequently, building upon traditional investigation approaches, researchers have explored various aspects of ICV-related accidents—including data acquisition, cause analysis, and accident reconstruction—leading to a growing body of valuable insights.

Beck et al.² proposed a data pipeline that collects and processes raw data from all onboard autonomous vehicle sensors, enabling crash investigators to visualize accident scenarios without prior knowledge of vehicle perception systems. Gong et al.³ conducted a statistical analysis of 418 major traffic accidents in China (2016–2019), identifying accident distribution patterns using hierarchical clustering. Wang et al.⁴ applied complex network topology to categorize real-world accident causes into time-related, environmental, and traffic management factors, finding traffic management to be the most influential. Bauder et al.⁵ developed a reconstruction method based on the CISS database and PCM methodology, calculating the vehicle center of gravity from deformation data to enhance the realism of crash scenario simulations. Alemazkoor et al.⁶ proposed an efficient data collection scheme with a collection rate as low as 0.05, depending on the required accuracy. Yoo et al.⁷ have developed a platform that extends the range of big data collected from connected cars from OBU (on-board unit), and CAN to cameras, lidar and GPS sensors, which can be used to analyze the range of roads.

In terms of accident causation analysis, Yuan et al.⁸ examined the underlying causes of ICV accidents. Addressing the challenges in ICV accident investigation, analysis, and liability attribution, they summarized key aspects from the perspectives of technical defects, information security, and passive safety. The study clarified the responsible parties in ICV-related accidents and their corresponding legal liabilities. Liu et al.⁹, based on in-depth crash investigation research, identified the main causes of autonomous vehicle collisions in emergency scenarios. These include the perception-reaction time gap between ICVs and human drivers, the failure of autonomous systems to interpret the lane-changing intentions of other vehicles accurately. Cummings et al.¹⁰ examined the recent accident where a pedestrian was dragged by a self-driving car, major themes include a lack of consideration for remote operations issues, significant post-collision assessment deficiencies, and a lack of periodic design and test protocol reviews. Wang et al.¹¹ employed a complex network topology approach to

construct a causality network for traffic accidents. Real-world accident data were categorized into temporal, environmental, and traffic management factors, among which traffic management was found to have the most significant impact on accident occurrence.

In the research of ICV accidents, the reference and integration of multi-dimensional methods are of great significance to improve the rigor of analysis. Silmi Juman ZAM et al.¹² developed a new $O(N^2)$ polynomial algorithm. The data-driven modeling method can provide support for the analysis of accident-related vehicle running state, environmental impact, and other dimensions. G M Hasan Shahariar et al.¹³ built an emission and fuel consumption prediction model by combining various machine learning algorithms, which can provide a method reference for accurate analysis of vehicle operating parameters in accident scenarios. Liu et al.¹⁴ proposed an enhanced SLAM framework that addresses operational efficiency challenges in GPS-unavailable environments, providing technical insights for resolving critical issues in ICV accident scene reconstruction, including complex environmental localization failures, missing scene details, and multi-source data interference.

In China, accident scene measurements are still primarily conducted manually, and scene diagrams are usually drawn by hand. In earlier studies, investigators estimated vehicle speeds at the time of a crash by examining scattered debris and tire marks based on the types of vehicles involved. However, with the mandatory adoption of Anti-lock Braking Systems (ABS), visible braking marks are often shorter than actual braking distances, which increases the error in speed calculations and complicates accident reconstruction. With the advancement of computer simulation, digital technologies, and video analysis, digital methods are gradually replacing traditional approaches in processing and analyzing road traffic accident scenes. In terms of vehicle speed reconstruction, when accurate modeling of the involved parties is challenging, Shang Gao et al.¹⁵ proposed a rigid-flexible coupled human model (HM) along with a Facet vehicle model to analyze both vehicle and human injuries, showing a high degree of consistency. For frontal collision scenarios, Han¹⁶ utilized crash test data from the U.S. National Highway Traffic Safety Administration (NHTSA) for curve fitting and analyzed video data from offset collision tests conducted by the Insurance Institute for Highway Safety (IIHS), ultimately developing a new formula for estimating collision speed. Esenturk et al.¹⁷ analyzed the UK STATS19 accident database to identify historical crash patterns and used these insights to systematically generate test scenarios for ICVs, contributing to safety improvements. Zhao et al.¹⁸ introduced a new method for simulating vehicle crashes in numerical environments, based

on the premise that most collisions are associated with specific driver characteristics. This approach supports future testing and validation of ICVs, particularly in evaluating their response to accident scenarios.

In summary, with the increasing prevalence of ICVs, researchers have extended traditional accident investigation methods—especially those based on driving record data—toward the study of ICV-specific crashes. Building on relevant reports from California, this paper proposes a closed-loop accident investigation framework that incorporates reverse scenario construction. By leveraging driving log data, vehicle dynamics modeling, and multi-source data processing techniques, the pre-collision dynamic scenario can be accurately reconstructed, offering scientific support for causation analysis and the development of industry standards.

2 Materials

According to statistics from the California Department of Motor Vehicles (DMV), 26 companies submitted autonomous vehicle disengagement reports in 2021. However, only 7 of these companies reported cases of forced disengagement in autonomous driving mode, involving 439 incidents, accounting for 16.4% of all disengagement events, representing a 4.4% decrease compared to 2020. As shown in Table 1, the leading causes of forced disengagements were software-related issues, including perception failures, decision-making errors, trajectory prediction problems, and path planning malfunctions. These accounted for 181 cases, representing 41.2% of the total. In contrast, hardware-related failures, such as component malfunctions, data recorder issues, and excessive braking, were less frequent, totaling 16 cases.

Table.1 In 2021, ICVs were forced out of takeover cases

collection	reasons	quantities
Perception issues	Detection of danger ahead leads to disengagement	10
	Sensing component issues leads to disengagement	2
	Perceiving differences leads to disengagement	2
	The software detects a target in the path causing disengagement	61
Modeling Algorithm Problems	Environmental modeling issues lead to disengagement	1
	Incorrect forecasting leads to disengagement	1
	Movement/behavioral planning differences lead to disengagement	1
	The inability to generate a valid trajectory within a period leads to disengagement	47
Controller issues	Motion control system problems lead to disengagement	2
	Unsuitable maneuvers lead to system disengagement	8
	Motion control operation status check causes software exit and disengagement	59
Software issues	Intermediate components are offline causing disengagement	1
	Localization components issues lead to disengagement	5
	Planner component issue leads to disengagement	31
	Software differences leading to disengagement	133

Hardware issues	Software system issues causing disengagement	2
	Fusion component errors lead to disengagement	9
	Hardware differences lead to disengagement	10
	Hardware diagnostics cause the system to exit and disengage	7
	Hardware diagnostics found problems with the hardware causing disengagement	30
	Disengagement due to hardware differences or system malfunction	1
else	Data logger problems lead to disengagement	11
	Braking too hard leads to disengagement	5

Among the 95 ICV accidents reported, excluding those that occurred under manual control by safety drivers, only 39 cases were caused while the vehicle was operating in autonomous mode, accounting for 41% of the total number of accidents throughout the year. As shown in Table 2, the most frequent accident type was rear-end collisions in which the ICV was hit from behind, with 22 cases, representing 56.4% of the autonomous mode accidents. There were only two incidents in which the AV was identified as the primary cause of the accident: one involved side-scraping the center divider during a left lane change, and the other involved reversing into a stationary vehicle. Both accidents were classified as minor with no reported injuries.

Table.2 ICV driving accident type distribution of DMV

Types of accident	Case count	Illustration (orange for ICV, blue for others)
ICV stops (rear-ended by straight /turning vehicle)	17	
ICV movement: low speed/braking (rear-ended by straight /turning vehicle)	5	
ICV scratched (opposite/illegal left-turn vehicles, riders/skateboards)	4	
ICV actively causing accidents (left lane change/strip contact, reverse into stationary vehicles)	2	
ICV hit (turning, other vehicles reversing, cyclists break mirror)	11	

Based on the analysis of ICV disengagement reports and accident characteristics, vehicles equipped with autonomous driving assistance systems exhibit more complex and diverse causes of accidents compared to traditional vehicles. Moreover, most accidents occur passively, with relatively few incidents directly initiated by the system. Although fully autonomous driving represents the ultimate stage of intelligent connected vehicle (ICV) development¹⁹, it has not yet been realized. In the meantime, the market share and penetration rate of ICVs below Level 3 are increasing steadily over successive years. However, reasonable and legally sound

methods for accident investigation and analysis remain underdeveloped. The *Regulations on In-Depth Investigation of Road Traffic Accidents (Trial)*²⁰ clearly state that in-depth vehicle accident investigations aim to analyze single or multiple cases in detail to uncover root causes, eliminate potential hazards, and promote the safe and orderly development of the automotive industry. Therefore, a systematic in-depth investigation mechanism for ICV accidents is of significant practical importance.

In summary, the causes of ICV accidents are diverse. Therefore, the analysis of ICV accident causation must be more comprehensive. The advanced equipment installed on ICVs provides strong support for accident reconstruction, greatly aiding the reverse reconstruction of accidents. During the investigation process, it is essential to fully utilize the multi-source information collected, including driver behavior, vehicle status, and road environment data, to achieve an accurate reconstruction of the accident. The animated reconstruction of the vehicle accident scene can visually represent the vehicle's status prior to the accident, effectively restoring the motion states of the vehicle and other involved traffic participants. It provides valuable assistance in determining accident responsibility and ultimately offers scientific support for improving ICV safety strategies.

3 Method

This study utilizes Open DRIVE and Open SCENARIO as research tools to complete the reverse reconstruction of ICV accidents.

Open DRIVE is a high-precision map format designed for representing static road networks, including elements such as lanes, traffic signs, and intersections²¹. Initially developed in 2006 by VIRESS and Daimler, it was later maintained by ASAM from 2018 onwards. Open DRIVE supports autonomous driving simulations by accurately describing road layouts and providing a hierarchical tree structure for efficient map storage and scene construction.

Open SCENARIO is a standard used to describe dynamic traffic scenarios in autonomous driving simulations²². Released by ASAM in 2020, it defines the motion behaviors, positions, and attributes of dynamic traffic elements. It includes components like Entities (scene elements) and Storyboard (dynamic behaviors), offering a standardized framework for modeling traffic interactions and simulating autonomous driving scenarios.

3.1 Generation of Accident Vehicle Trajectories

By utilizing the vehicle driving log data's time series, along with the steering angle and vehicle speed, it is possible to generate the vehicle's motion trajectory at different time points. To simulate the

motion of intelligent vehicles, it is essential to consider their kinematic characteristics. Simplifying the vehicle as a bicycle model significantly reduces the complexity of the model. The vehicle kinematic model studies the vehicle's motion law from a geometric perspective, including the changes in the vehicle's spatial posture, speed, and other factors over time. Figure1 illustrates the simplified vehicle model.

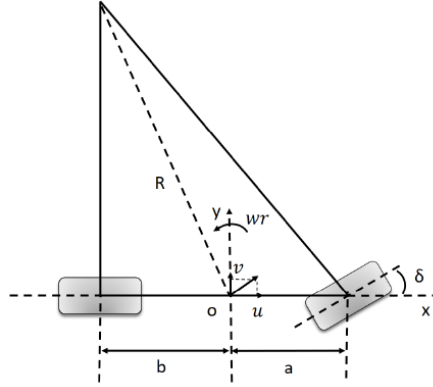


Figure.1 Simplified vehicle model

The steady-state yaw rate gain of the vehicle based on the model²³ is:

$$\left(\frac{w_r}{d}\right)_s = \frac{u/L}{1 + \frac{m}{L^2} \left(\frac{a}{k_2} - \frac{b}{k_1}\right) u^2} = \frac{u/L}{1 + Ku^2} \quad (1)$$

$$K = \frac{m}{L^2} \left(\frac{a}{k_2} - \frac{b}{k_1}\right) \quad (2)$$

In the above formula, $\left(\frac{w_r}{d}\right)_s$ refers to the steering sensitivity, K refers to the stability factor, k refers to the cornering stiffness. a and b refers to the distance from the vehicle's center of gravity to the front and rear axles. u refers to the velocity component of the vehicle along the x-axis in its own coordinate system. d refers to the steering angle of the front wheels.

When the front wheel steering angle is constant and lateral acceleration is neglected, the turning radius is R_0 . However, when the vehicle speed is constant, the turning radius with lateral acceleration is R , R/R_0 can be used to characterize the steady-state response of the vehicle:

$$R = \frac{u}{w_r} = \frac{(1 + Ku^2)L}{d} = (1 + Ku^2)R_0 \quad (3)$$

$$R_0 = L/d \quad (4)$$

If the vehicle's current position at T_0 is (X_0, Y_0) , the yaw angle is α_0 . After t , the vehicle's position changes to (X_t, Y_t) the heading angle becomes

q_t As shown in Figure 2, the change in the vehicle's heading angle q is:

$$q = w_r \times t \quad (5)$$

$$q_t = q_0 + q \quad (6)$$

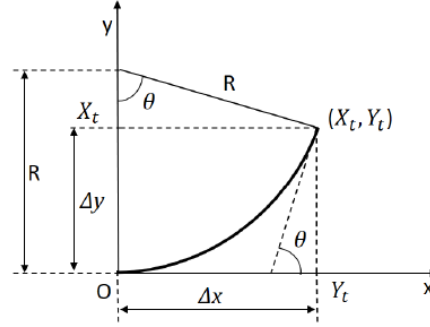


Figure.2 Schematic diagram of vehicle position after time

After t , the vehicle has traveled distances along the x and y axes in Figure 2, represented as Δx and Δy :

$$\Delta x = R \sin(q) = \frac{(1 + Ku^2)L}{d} \sin\left(\frac{u/L}{1 + Ku^2} t\right) \quad (7)$$

$$\Delta y = R - R(\cos(q)) = R(1 - \cos(q)) = \frac{(1 + Ku^2)L}{d} \left(1 - \cos\left(\frac{u/L}{1 + Ku^2} t\right)\right) \quad (8)$$

According to the derived formulas, given the initial speed and steering wheel angle of the vehicle, the vehicle's position and steering angle can be obtained after a certain period of time.

3.2 Construction of Static Accident Scene

The elements for static scene construction include basic road features, streetlights, signs and markings, obstacles, and others. The challenge in constructing road features lies in the difficulty of obtaining the road curvature of the accident segment. By utilizing the OpenStreetMap (OSM) open-source map, the basic characteristics of the accident segment can be more accurately identified. However, obtaining the corresponding starting position of the accident segment within the 20 seconds before the collision is more challenging. In this study, the position of the vehicle 20 seconds before the collision and at time zero was determined through a combined comparison of Baidu Map street view and dashcam video. The technical approach is illustrated in Figure 3.



Figure.3 Static scene construction technique route

In Step 6 of Figure 3, since the OSM (OpenStreetMap) open-source map format differs from the Open DRIVE format, a conversion tool is required to transform the open-source map format into the Open DRIVE standard map format for subsequent road reconstruction. The tool conversion program from the SUMO traffic simulation software serves as a bridge, effectively converting the open-source map format into the Open DRIVE format.

3.3 Construction of Dynamic Accident Scene

To fully restore the accident scene, in addition to reconstructing the static scene, it is essential to reconstruct the status of the main vehicle at various time points. This includes the main vehicle's speed, steering angle, and the speeds of vehicles in adjacent lanes detected by the main vehicle's sensors. The dynamic aspects of the scene are described using Open SCENARIO. Based on the information recorded in the driving log data, the accident can theoretically be recreated.

Given the vast and detailed information contained in the driving log data, and considering factors such as reconstruction accuracy and software limitations, Python programming is used to extract key data, including the time sequence, main vehicle's steering angle, main vehicle's speed, and the longitudinal distances of vehicles in the left, front, and right lanes relative to the main vehicle. This processed information is organized into an Excel sheet.

The steering wheel angle recorded in the driving log data is converted into the wheel's deflection angle per unit time using the vehicle's

transmission ratio (10:1). Using formula (9), the vehicle's lateral displacement per unit of time can be calculated, where the steering wheel angle is one-tenth of the actual steering wheel angle.

$$v_x = v \sin \alpha \quad (9)$$

The longitudinal distance of the vehicles in the left, front, and right lanes detected by the main vehicle's sensors relative to itself in the Excel information table needs to be calculated using the velocity-time-distance relationship to determine the speed of other vehicles at different moments. The formula is as follows:

$$V = \frac{v_1 + v_1'}{2} - \frac{d_1 - d_2}{Dt} \quad (10)$$

V Represents the current speed of the vehicle in front of the left lane, v_1 represents the current speed of the main vehicle, v_1' represents the speed of the main vehicle in the next second, d_1 represents the current longitudinal distance between the main vehicle and the vehicle ahead in the left lane, d_2 represents the longitudinal distance between the main vehicle and the vehicle ahead in the left lane in the next second.

Using the method outlined above, the lateral displacement of the ego vehicle and the speed of other vehicles at each time step can be derived. Based on the movement information of all involved vehicles and the requirements of the Open SCENARIO standard, it is clear that the static scene information references the static scene data. The updated information primarily involves motion data in the *Entities* and *Storyboard* components. In the storyboard, the key updates include the ego vehicle's speed, lateral displacement at different time steps, and the speed of the front and left-side vehicles at the same time intervals.

4 Result

4.1 Accident Case Introduction

In June 2021, Mr. Chen was driving a branded intelligent connected vehicle on Hongmei South Road in Shanghai under clear weather conditions. At the time of the collision, the vehicle was traveling at approximately 60 km/h. A sudden brake by the vehicle ahead led to a three-car rear-end collision, resulting in injuries to individuals in the front vehicle. The available accident data includes dashcam footage, on-site photographs, audio recordings, vehicle EDR (Event Data Recorder) data, and driving log data. These data sources enable an effective reconstruction of the vehicle's state before the accident using the reverse scene construction method.

Comparative analysis with the EDR data and dashcam footage confirms the feasibility of the reconstruction approach.

4.2 Vehicle Trajectory Analysis

Based on the derived formulas (1) - (8), and knowing the initial position, speed, and steering angle of the vehicle, after a certain period, the position and steering angle of the accident vehicle can be determined. Table 3 presents a subset of the raw data, including time sequences, speed, and steering angles, which are used in the calculation process.

By applying these formulas, we can track the vehicle's position and orientation over time, providing essential data for reconstructing the accident scene and analyzing the dynamics of the vehicle's movement before the collision.

Table. 3 Track raw data sheet

datetime	DI_vehicleSpeed	SCCM_steeringAngle
44368.76734	66.16	1.8
44368.76734		2.3
44368.76735	66.32	
44368.76736		2.4
44368.76736	66.72	
44368.76737		2
...	...	...

Given the known wheelbase of the accident vehicle and the vehicle trajectory generation formula mentioned above, the motion trajectory of the main vehicle for the 20 seconds before the collision can be implemented in MATLAB. By using the vehicle's driving log data recorded at different time intervals, along with the speed and steering angle, a MATLAB program can be written to obtain the vehicle's motion trajectory over the 20 seconds.

Figure 4 presents the actual vehicle trajectory on the map, while Figure 5 shows the vehicle motion trajectory reconstructed using the proposed methodology (represented by the middle blue and green line segments). Comparison of these two trajectories reveals a close alignment with both each other and the factual evidence, confirming the accuracy of the reconstruction method.



Figure.4 trajectory generation on the Map

4.3 Static Scene Analysis

After converting the open-source map of the accident location into an Open DRIVE format, the road sections became numerous, and the scene appeared somewhat chaotic. Additionally, it contained non-internationally standardized recording formats, necessitating the extraction of road curvature features from the accident section. Open DRIVE file information is recorded in XML format, and the Open DRIVE map viewer allows users to query the ID of each road. By referencing these IDs, users can return to VS Code to access the corresponding road information.

According to the Open DRIVE viewer, the reconstructed road section is composed of ID425 and ID426. By measuring the distance, it was found that the distance from the end of ID425 to the collision point at the 20-second mark was 209 meters. Therefore, to fully cover the accident section, the road features of the ID425 section must extend beyond 209 meters. Figure 5 illustrates the key point information of the road section.

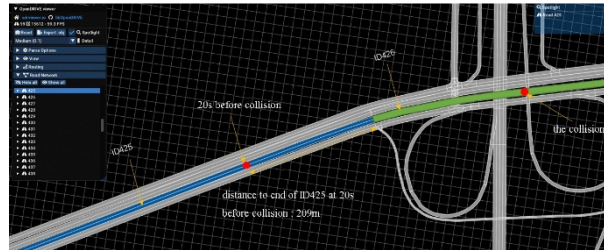


Figure.5 Key point information of section

The specific details of the accident road section were redefined according to the Open DRIVE standards. Firstly, the spacing of the white dashed lane markings on the roadway was set such that the dashed segment length is 4 meters, with a blank segment length of 6 meters. Additionally, the streetlight information was referenced from Table 4, while relevant road information was based on the data in Table 5. A portion of the Python code is written the update the road's total length, the representation of different curvature segments, road elevation, lane width, the central lane marking, and streetlight placement. These detailed parameters and standard requirements ensure that the reconstructed road model closely matches the actual scene where the accident occurred, providing precise road data support for subsequent dynamic scene analysis.

Table.4 Preparation of key information on street lights on accident roadways

streetlight information	streetlight coverage length(m)	streetlight spacing(m)	streetlight height(m)
	678.69041303	30	7.5

Table.5 Preparation of key information on accident roadways

Key Road Segment Information for Accident Scenario Reconstruction

Road ID	Segment Length (m)	Straight/Curved	Inertial Heading Angle (rad)	x Coordinate (m)	Y Coordinate (m)
ID425	143.83228235	Straight	- 1.05260427	399.22134363	2155.02573569
	27.77996235	Curved	- 1.05260427	470.46296266	2030.07631727
	92.18122496	Straight	- 1.05630399	484.17799490	2005.91801331
	27.77957597	Curved	- 1.05630399	529.53972788	1925.67032540
	9.59509299	Curved	- 1.06361649	543.06889362	1901.40792762
	27.77718139	Curved	- 1.07909163	547.69301562	1893.00066688
	23.45364739	Straight	- 1.09794506	560.71288465	1868.46420947
ID426	15.80673042	Straight	- 1.13786059	571.55503741	1847.25374010
	16.79781555	Curved	- 1.13786059	578.18655334	1832.90537005
	53.08653011	Straight	- 1.22957602	584.73377516	1817.44020663
	27.76379978	Curved	- 1.22957602	602.49850792	1767.41427386
	49.94894980	Straight	1.30631831	610.77754195	1740.91804200
	27.77982458	Curved	1.30631831	623.83447081	1692.70586310
	107.32783864	Straight	- 1.29833213	631.20326887	1665.92122366
	27.77995675	Curved	- 1.29833213	660.08578554	1562.55263472

The static reconstruction elements include the key features of the accident site, such as the road segment and streetlights. Following the method outlined in the previous section for static scene construction, critical road information is extracted and used to write Open DRIVE code for generating static data. Figure 6 presents the reconstructed road, which provides the road model for the dynamic scene construction. The total length of the reconstructed road is 678 meters. By using map measurements, the location of the reconstructed road on the map is determined.

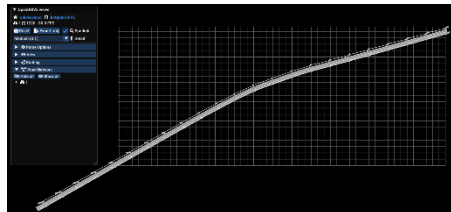


Figure.6 Reconstruction of the accident section of the road

4.4 Dynamic Scene and Accident Causation Analysis

Given the wheelbase of the accident vehicle, the steering model of the vehicle is generated according to the trajectory generation formula (9) as

shown in Figure 7. Using formula (10), the vehicle's positional changes at each time step are obtained, which represent the dynamic changes of the scene elements within one second, as illustrated in Figure 8.

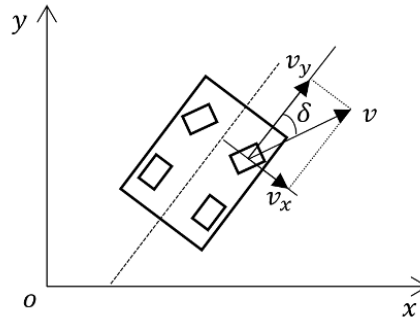


Figure.7 Vehicle steering model

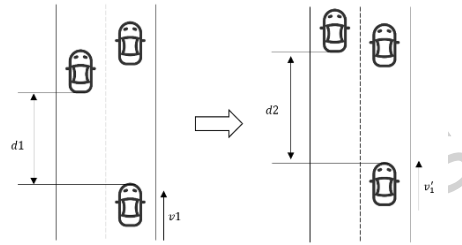


Figure.8 Vehicle position change in unit timing

Based on the proposed method and the reconstructed data for the 20 seconds preceding the collision, the lateral displacement of the primary vehicle and the velocities of other traffic participants at each time step are extracted. This motion data is then used to update the storyboard, ensuring that the trajectories of all entities comply with the dynamic characteristics of the accident participants and align with the formatting and structural requirements of the Open SCENARIO standard. Comparison of the simulated speed and the logging data speed of the main vehicle is shown in Figure 9.

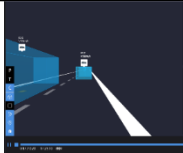


Figure.9 Main vehicle speed comparison

A comprehensive dynamic accident scene includes the motion characteristics of the main vehicle and other vehicles leading up to the collision. Using the dynamic scene construction method outlined in the

motion animations of all vehicles before the collision was generated. At this point, the main vehicle is traveling at 66 km/h, the vehicle in front on the left is moving at 73 km/h, and the vehicle in front on the right is traveling at 65 km/h. This pattern holds for both the 3-second and 0-second moments before the collision. Table 6 shows the pre-crash Animation before the collision at the three pre-crash time points compared to the Car log video before the collision.

Table.6 Comparison of pre-crash main vehicle crash animation and tachograph video

pre-collision time	Animation before the collision	Car log video before the collision
20s		
3s		
0s		

The comparative analysis demonstrates that the reconstructed animations of the selected rear-end collision cases closely resemble the actual accident footage. This not only validates the authenticity of the reconstruction results but also provides preliminary confirmation that the accident reconstruction framework developed in this study is applicable to such scenarios. The alignment between quantitative analysis outcomes and factual evidence further indicates the framework's practical value in determining liability for similar rear-end collisions. It should be noted that the applicability of this framework to other accident scenarios, such as side collisions or adverse weather conditions requires further validation through more diverse case studies.

4.5 Accident Causation Analysis

Through a comprehensive analysis of vehicle log data, motion trajectories, static and dynamic scenes, and surveillance footage as well, the primary cause of the accident can be identified. First, after the activation of the Automatic Emergency Braking (AEB) system at 06:25:17.94 PM, it was deactivated within a very short time, at 06:25:18.26 PM. This abnormal situation significantly compromised the vehicle's safety capabilities in the face of potential danger. Immediately following, at

06:25:18.30 PM, the accelerator pedal position suddenly increased to 100%, causing the vehicle to accelerate rapidly, further exacerbating the dangerous situation. Most crucially, before the collision, the driver never applied the brake pedal and took no braking action to avoid the crash. Furthermore, from the driving behavior perspective, the driver was consistently tailgating in the same lane. Although no lane change occurred, the following distance was too short, with the closest distance to the lead vehicle being only 11.3 meters during the 20 seconds before the collision. At the same time, the average speed was 65.8 km/h. This high-speed, close-distance tailgating greatly increased the likelihood of an accident, ultimately leading to the collision.

4.6 Error Analysis

To verify the reliability of the simulation model, a quantitative error analysis must be conducted on key parameters that characterize vehicle motion states. As core indicators of dynamic behavior, speed and steering angle directly determine the reconstruction accuracy of vehicle trajectories and collision processes in accident scenarios. Therefore, we extracted speed and steering angle data from the first 20 seconds of simulation results for comparison with real data collected by the Event Data Recorder (EDR) in the accident vehicle. The results are shown in Table 7.

Table 7. Comparison between Pre-collision Simulation Data and EDR Data

Pre-collision Time(s)	Host Vehicle Speed(km/h)	Simulated Speed(km/h)	Steering Angle [Degree]	Steering Angle Difference
20	66.905	66.161	2.3	0.5
19	68.107	66.473	2.4	0.1
18	69.164	67.962	2	-0.4
...	...	...	...	...
5	62.375	62.437	6.7	0
4	61.394	61.548	5.4	-0.3
3	61.260	59.345	5.4	0
2	62.148	60.803	4	-0.4
1	55.590	54.245	3.5	-0.5
0	47.374	47.374	3.1	-0.4

By calculating the Mean Relative Error (MRE) between simulated and actual vehicle speeds, the results show: The average relative error in speed is approximately 1.79%, with a maximum relative error of 3.67%; the average relative error in steering angle is about 9.54% (primarily affected by fluctuations in individual time points). The steering angle is directly related to the vehicle's steering geometry and kinematics. The two-wheeled (bicycle) model simplifies the steering mechanism of the four-wheeled vehicle, which directly increases the steering angle deviation. Speed is a macroscopic parameter reflecting the vehicle's overall translational motion, and it is mainly associated with the vehicle's overall

force conditions. Since the model simplification exerts a smaller impact on such parameters, the speed error is consequently lower.

Overall, the simulation data demonstrates high consistency with real-world data, validating the sub-model's capability to accurately reproduce the motion characteristics of such accident vehicles.

5. Discussion

This study proposes and validates a reverse accident scenario reconstruction model tailored to rear-end collision cases, yet it is constrained by inherent limitations.

With respect to data dependency, the framework is highly reliant on detailed, high-precision driving log data and on-site road information. In practical accident investigations, critical data may be compromised or lost, while the lack of unified data formatting across vehicle manufacturers necessitates extensive preprocessing (e.g., calibration, format conversion). Future work will develop time-series prediction-based data imputation algorithms to compensate for partially missing data²⁴ and explore industry-wide standardized interfaces compatible with mainstream manufacturers, to resolve format inconsistencies and enable compliant access to encrypted data.

Regarding model simplification, the 2-wheel bicycle model was employed for vehicle motion simulation to balance computational efficiency and real-time performance—a common simplification in preliminary trajectory reconstruction. However, this model cannot fully capture the intricate physical behaviors inherent in real-world collisions (e.g., tire slip during emergency braking, suspension dynamics in lateral impacts), resulting in potential discrepancies in reconstructed vehicle attitude (e.g., yaw angle, lateral offset) at the collision instant²⁵. In addition, the coupling relationship between the geometric parameters of the four-wheeled vehicle's steering mechanism (e.g., steering ratio, wheelbase effects) and the side-slip angle contributes to a further amplification of errors in local geometric parameters—specifically, the steering angle error²⁶. It also neglects post-collision body deformation and secondary motions. Subsequent research will integrate multi-body dynamics models to enhance physical fidelity and combine the simplified bicycle model with high-precision counterparts to balance efficiency and accuracy.

A further constraint relates to generalizability; its validity has solely been validated through a single rear-end collision case on an urban highway. Future validation will expand to a diversified case library covering side-impact, intersection, and low-visibility scenarios²⁷. The framework's modular architecture will also be refined with scenario-

adaptive parameter modules and multi-vehicle interaction logic to boost applicability across ICV accident scenarios.

6 Conclusion

In response to the challenges of in-depth investigations of ICV accidents, this study proposes a closed-loop methodology that incorporates reverse scenario construction, grounded in traditional frameworks and typical accident characteristics. Through a human-vehicle-road collaborative data collection mechanism and multi-source evidence validation, the method emphasizes reverse reconstruction as a core tool for tracing accident causation in complex scenarios.

By combining vehicle dynamics theory with multi-source data processing, the study reconstructs high-precision pre-collision scenes using a two-wheel model built from driving log data. Road geometry is standardized via the SUMO toolchain, and Python algorithms process heterogeneous data to extract dynamic parameters of traffic participants. Based on the Open SCENARIO standard, a parameterized dynamic model of the 20 seconds prior to collision is developed.

This work establishes a scene-adapted technical framework of “data-scenario simulation analysis” for the selected rear-end collision case, initially forming a quantitative analysis tool tailored to responsibility determination in such accidents. Based on the validation of this single case, the framework also provides certain technical reference for evaluating autonomous driving system safety and formulating relevant standards in similar scenarios. While it enriches existing ICV accident investigation approaches and offers valuable insights for future research and practical application, its generalizability to other accident types (e.g., side-impact collisions, intersection accidents) and complex conditions (e.g., adverse weather) remains to be further verified through diversified case studies.

Reference:

1. Xiaohui Yu. China Internet Development Report (2023) published in July 2023. <https://www.isc.org.cn/article/17333342358990848.html> (2023).
2. Beck, J., Arvin, R., Lee, S., Khattak, A. & Chakraborty, S. Automated vehicle data pipeline for accident reconstruction: New insights from LiDAR, camera, and radar data. *Accident Analysis & Prevention* **180**, 106923 (2023).

3. Gong, P., Wang, Q. & Zhu, J. Spatiotemporal Characteristics of Traffic Accidents in China, 2016-2019. *Intelligent Automation & Soft Computing* **29**, 31-42 (2021).
4. Wang, Y., Zhai, H., Cao, X. & Geng, X. Cause Analysis and Accident Classification of Road Traffic Accidents Based on Complex Networks. *Appl. Sci.-Basel* **13**, 12963 (2023).
5. Bauder, M. *et al.* Determination of accident scenarios via freely available accident databases. *Open Eng.* **12**, 453-467 (2022).
6. Alemazkour, N. & Meidani, H. Efficient Collection of Connected Vehicles Data With Precision Guarantees. *IEEE Trans. Intell. Transp. Syst.* **21**, 4637-4645 (2020).
7. Yoo, A., Shin, S., Lee, J. & Moon, C. Implementation of a Sensor Big Data Processing System for Autonomous Vehicles in the C-ITS Environment. *Appl. Sci.-Basel* **10**, 7858 (2020).
8. Yuan, Q., Peng, Y., Xu, X. & Wang, X. Key points of investigation and analysis on traffic accidents involving intelligent vehicles. *Transp. Saf. Environ.* **3**, tdab020 (2021).
9. Liu, Q., Wang, X., Wu, X., Glaser, Y. & He, L. Crash comparison of autonomous and conventional vehicles using pre-crash scenario typology. *Accident Analysis & Prevention* **159**, 106281 (2021).
10. Cummings, M. M., Wheeler, B. & Kliem, J. A Root Cause Analysis of a Self-Driving Car Dragging a Pedestrian. *Computer* **57**, 31-40 (2024).
11. Wang, Y., Zhai, H., Cao, X. & Geng, X. Cause Analysis and Accident Classification of Road Traffic Accidents Based on Complex Networks. *Applied Sciences* **13**, 12963 (2023).
12. Silmi Juman, Z. A. M. *et al.* A new algorithm for solving uncapacitated transportation problem with interval-defined demands and suppliers capacities. *J. Intell. Fuzzy Syst.* **41**, 625-637 (2021).
13. Shahariar, G. M. H. *et al.* Real-driving CO₂, NO_x and fuel

consumption estimation using machine learning approaches. *Next Energy* **1**, 100060 (2023).

14. Liu, X., Zhang, X., Yan, X., Liu, P. & Lao, H. RGB-D camera and graph neural network-based SLAM for dynamic and low-texture environments. *Sci Rep* **15**, 28214 (2025).

15. Gao, S. *et al.* A Research on Accident Reconstruction of Bus-Two-Wheeled Vehicle Based on Vehicle Damage and Human Head Injury. *Int. J. Environ. Res. Public Health* **19**, 14950 (2022).

16. Han, I. Vehicle collision analysis from estimated crush volume for accident reconstruction. *Int. J. Crashworthiness* **24**, 100–105 (2019).

17. Esenturk, E., Wallace, A. G., Khastgir, S. & Jennings, P. Identification of Traffic Accident Patterns via Cluster Analysis and Test Scenario Development for Autonomous Vehicles. *IEEE Access* **10**, 6660–6675 (2022).

18. Zhao, L., Farhi, N., Christoforou, Z. & Haddadou, N. Analysis of Driver Behavior and Intervehicular Collision: A Data-Based Traffic Modeling and Simulation Approach. *J. Adv. Transp.* **2022**, 1068311 (2022).

19. Xu, Q. *et al.* The status, challenges, and trends: an interpretation of technology roadmap of intelligent and connected vehicles in China (2020). *JICV* **5**, 1–7 (2022).

20. Xinquan Du & Feng Liu. Exploration of the cooperation mechanism of in-depth investigation and forensic appraisal of traffic accidents. *Chinese Journal of Forensic Sciences* 91–94 (2020).

21. Introduction :: OpenDRIVE®.

https://publications.pages.asam.net/standards/ASAM_OpenDRIVE/ASAM_OpenDRIVE_Specification/latest/specification/00_preface/00_introduction.html.

22. Introduction :: ASAM OpenSCENARIO.

https://publications.pages.asam.net/standards/ASAM_OpenSCENARIO/AS

AM_OpenSCENARIO_XML/latest/00_preface/01_introduction.html.

23. Yu, Z. (ed.) *Automobile Theory* (in Chinese). Tsinghua University Press, Beijing (2005).

24. Shang, Q., Tang, Y. & Yin, L. A hybrid model for missing traffic flow data imputation based on clustering and attention mechanism optimizing LSTM and AdaBoost. *Sci Rep* **14**, 26473 (2024).

25. Jin, L., Zhao, S. & Xu, N. A Survey of Vehicle Dynamics Models for Autonomous Driving. in 2024-01-2325 (Detroit, Michigan, United States, 2024). doi:10.4271/2024-01-2325.

26. Alves Torres, L., Murilo, A., Vilela Lopes, R. & Leal, V. Analysis and Virtual Validation of Vehicle Dynamics Models for Electronic Stability Control. *IEEE Latin America Transactions* **22**, 166-172 (2024).

27. ISO 21448:2022. *ISO* <https://www.iso.org/standard/77490.html>.

Author contributions statements:

S. Z provided experimental data and methodological guidance. X. L. was responsible for figure preparation and drafting the initial manuscript. Z. Z. conducted data analysis. H. X. contributed to the writing of the initial draft. C. W. provided guidance on the manuscript's structure and logical flow. P. L. oversaw result validation and offered technical guidance. K.Z. translated the manuscript. All authors reviewed the manuscript.

Competing interest statement:

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data availability:

The datasets generated and/or analyzed during the current study are not publicly available due to Sensitive information involved in road traffic accidents, including the identity of the parties involved in the accident, vehicle tracks, and manufacturers' proprietary technical parameters, but are available from the corresponding author on reasonable request.