# Fuel Consumption Evaluation of Connected Automated Vehicles Under Rear-End Collisions

Qingchao LIU[1], Wenjie OUYANG[2], Jingya ZHAO[3], Yingfeng CAI[4], Long CHEN[5]

[1] lqc@ujs.edu.cn, Automotive Engineering Research Institute, Jiangsu University; School of Mechanical and Aerospace Engineering, Nanyang Technological University; Jiangsu University Research Institute of Engineering Technology

[2] scisci@126.com, Automotive Engineering Research Institute, Jiangsu University; Jiangsu University Research Institute of Engineering Technology

[3] zhaojingya26@163.com, Automotive Engineering Research Institute, Jiangsu University; Jiangsu University Research Institute of Engineering Technology

[4] caicaixiao0304@126.com, Automotive Engineering Research Institute, Jiangsu University

[5] chenlong@ujs.edu.cn, Automotive Engineering Research Institute, Jiangsu University

## ABSTRACT

Connected automated vehicles (CAV) can increase traffic efficiency, which is considered a critical factor in saving energy and reducing emissions in traffic congestion. In this paper, systematic traffic simulations are conducted for three car-following modes, including intelligent driver model (IDM), adaptive cruise control (ACC), and cooperative ACC (CACC), in congestions caused by rear-end collisions. From the perspectives of lane density, vehicle trajectory and vehicle speed, the fuel consumption of vehicles under the three car-following modes are compared and analysed, respectively. Based on the vehicle driving and accident environment parameters, an XGBoost algorithm-based fuel consumption prediction framework is proposed for traffic congestions caused by rear-end collisions. The results show that compared with IDM and ACC modes, the vehicles in CACC car-following mode have the ideal performance in terms of total fuel consumption; besides, the traffic flow in CACC mode is more stable, and the speed fluctuation is relatively tiny in different accident impact regions, which meets the driving desires of drivers.

## KEYWORDS

CAV; traffic accident; fuel consumption prediction; energy saving.

## 1. INTRODUCTION

With the continuous improvement of people's living standards and increased car ownership, traffic congestions frequently occur, especially during traffic accidents such as rear-end collisions, when the road capacity will be significantly reduced, or in the worst-case traffic will be paralysed. Congestions bring substantial economic loss to society and lead to increased energy consumption and environmental pollution, limiting the automobile industry's development.

ICVs have complete information perception and faster information interaction ability compared to manual driving vehicles. Thus, ICVs can know the road conditions more quickly and have more time and space to adjust the vehicle driving conditions to reduce the number of collisions, lessen the influence of congestion on the overall traffic flow, improve the traffic efficiency of the road, and reduce the economic losses caused by congestion [1]. In addition to improving the stability of traffic flow [2, 3], connected vehicle is also one of the effective ways to achieve energy saving and emission reduction [4−6], which has essential research significance in the era of green transportation and is a research direction that has attracted much attention in the industry.

This paper mainly studies vehicle fuel consumption performance under three car-following modes when a rear-end accident occurs and the traffic flow is congested. From the aspects of fuel consumption on the lane

and vehicle levels, a comparative analysis is carried out in combination with the characteristics of lane density, vehicle trajectory and vehicle speed. A fuel consumption prediction framework based on XGBoost is proposed to predict vehicle fuel consumption after an accident by combining vehicle speed, acceleration, space headway, and time and space position.

The remainder of this article is organised as follows. Section 2 presents the literature review. Section 3 introduces the data collection process based on simulation. In Section 4, the vehicle fuel consumption characteristics under different car-following modes are compared and analysed from the perspectives of lanes and vehicles. Section 5 proposes a prediction framework of vehicle fuel consumption and a comparison of prediction results for different car-following modes, followed by the summary of the paper in Section 6.

## 2. LITERATURE REVIEW

Many studies have evaluated the impact of autonomous vehicles (AVS) on traffic events. Mousavi et al. evaluated the safety impact of different market penetration rates of autonomous vehicles under different traffic service levels [7]. They found that with autonomous vehicles' increasing market penetration rate, rear-end collision and lane change conflict events on urban roads were significantly reduced. Zhu and Tasic found that as the market penetration of autonomous vehicles increases, the probability of collisions at expressway junction ramps is significantly reduced [8]. It had a significant effect even when the market penetration was low. Papadoulis et al. showed that CAVs could make traffic flow more efficient and significantly reduce traffic conflicts even with relatively low market penetration [9]. Petrović et al. found that there were more "rear end" accidents in automatic driving vehicle accidents and fewer "pedestrian" and "scratch" collision types based on automatic driving traffic accidents in California from 2015 to 2017 [10]. Li et al. employed time-exposed time-to-collision (TET) and time-integrated time-to-collision (TIT) to quantify rear-end collision risk and showed that the CACC system could significantly reduce the risk [11]. These scholars have studied traffic accidents from the safety perspective to seek measures to reduce the risk of accidents. However, when an accident is relatively minor and only causes congestion, the impact on the road capacity and vehicle fuel consumption becomes more important, especially when the traffic is stopped for a long time due to the accident.

Some studies have analysed the impact of ICVs on road capacity. Xiao et al. conducted traffic simulations on the market penetration rate of various CACCs and thoroughly studied the impact of CACCs on the expressway merging bottleneck area [12]. They found that increasing the CACC market penetration rate can improve road capacity. Zhu et al. considered the differences in driver characteristics. They divided the drivers of autonomous vehicles into adaptive and non-adaptive groups [13]. Their results showed that increasing the market penetration of CACC can effectively alleviate congestion. With the increase in the number of adaptive drivers, the cooperative driving strategy of CACC can play a better role and further improve road capacity. Zhou et al. proposed three cooperative driving strategies and studied their impact on mixed four-lane traffic flow [14]. The results showed that the proposed cooperative driving strategy could more effectively realise the formation of vehicles in CACC mode and improve traffic capacity and stability by CACC market penetration. Qin et al. established a framework for heterogeneous traffic flow stability, derived a basic heterogeneous flow model and performed numerical simulations [15]. They found that increasing the proportion of vehicles in the CACC car-following mode can improve the stability of heterogeneous flow and road capacity. Their simulation also showed that a more considerable expected space headway in the CACC car-following mode induced a larger heterogeneous flow stability area. However, the traffic capacity would be reduced.

Other studies have analysed the impact of ICVs on fuel consumption. Yang et al. proposed a collaborative driving framework for urban arterial roads, which benefited vehicle mobility and fuel economy based on simulation [16]. Ma et al. proposed an Eco-CACC and verified that the proposed Eco-CACC had better energy-saving characteristics than manual driving [17]. In a study by Almannaa et al., an Eco-CACC system was evaluated by field experiments in different scenarios at intersections [18]. The results showed that the Eco-CACC significantly reduced fuel consumption and driving time. In recent years, many studies have begun to predict the energy consumption of vehicles using machine learning methods. Zhang et al. proposed a machine learning-based energy consumption prediction framework for electric vehicles based on driving data [19]. They showed that the machine learning-based prediction framework had higher accuracy than traditional methods. Sun et al. also built a machine-learning model to predict the fuel consumption of hybrid cars based on actual driving data, and their results also showed good prediction accuracy [20]. Yao et al. used three different machine learning models to predict the fuel consumption of cars based on driving behaviour data. Their results

showed that all three models could accurately predict fuel consumption [21]. Wang et al. proposed a fuel consumption prediction model based on machine learning and achieved good prediction accuracy [22].

The literature above shows that extensive research has been conducted for ICVs on preventing traffic conflicts/collisions, improving road capacity and reducing fuel consumption, respectively. However, evaluating ICV fuel consumption with respect to accident occurrence is rarely researched. The fuel consumption prediction for accidents, especially for rear-ends causing single-lane blocking that frequently occurs on urban expressways, can be used to formulate appropriate fuel consumption indicators and facilitate vehicle path planning strategies and cost budgets. To fill the gap, this paper investigates the fuel consumption performance of vehicles in IDM, ACC and CACC car-following modes when a rear-end has blocked a single lane to examine whether the CACC mode can reduce fuel consumption and improve road capacity under accident conditions. Specifically, homogeneous traffic flow under three car-following modes is simulated for more straightforward comparative analysis. Based on the vehicle driving and accident environment parameters, a machine learning model is also established to predict the vehicle fuel consumption and divide the impact regions of the accident to facilitate vehicle path planning from an energy-saving perspective.

## 3. DATA

To study the impacts of ICVs on fuel consumption under accident conditions, a rear-end collision that often emerges on the urban expressways is simulated using Simulation of Urban Mobility (SUMO) [23]. SUMO is an open-source, highly portable, micro and continuous traffic simulation package that allows for multimodal traffic simulation, including pedestrians, with many tools for creating various traffic scenes. As shown in *Figure 1*, the simulated scene is a 2 km straight three-lane urban expressway, with a rear-end collision set up as a vehicle stopping on the middle lane (i.e. Lane_1) 400 m away from the end-point of the simulated urban expressway. We choose the middle lane as the accident lane to observe the different performances more clearly between different car-following modes because, in this case, vehicles in the accident lane have more options. Traffic flow characteristic data are obtained by setting three induction loop detectors 150 m upstream of the accident vehicle. The embedded emission estimation module retrieves the fuel consumption data in SUMO.
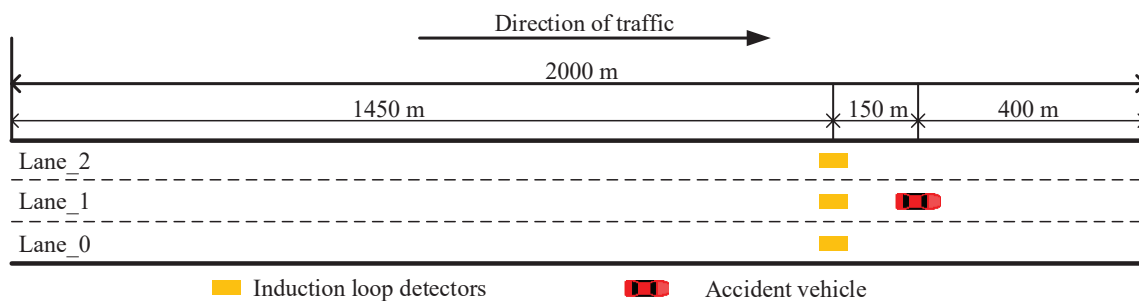


*Figure 1 – Schematic diagram of simulation scene*

As presented in *Table 1*, the simulation lasts for 2500 s, with vehicles starting to enter the network randomly from 400 s at a flow rate of 8000 veh/h. The accident vehicle departs at 500 s and travels for 1.6 km before it stops (to simulate a rear-end collision). In 1200 s (i.e. 20 min) after the collision occurs, the accident vehicle leaves the crash site and the traffic gradually returns to normal.

To explore the fuel consumption characteristics of ICVs under accident conditions, the car-following model is taken as the primary control variable in the simulation. All car-following models in our simulation are acceleration control models. German scholar Treiber et al. proposed the IDM car-following model, and it can better reflect the car-following characteristics of artificial vehicles [24]. The CACC and ACC car-following models are calibrated by the real vehicle test data in the PATH laboratory of the University of California, Berkeley, which can reflect the real car-following characteristics of CACC and ACC [25]. These car-following models are described by *Equations 1–3*, respectively:

$$\dot{v}_{IDM} = a\left\{1 - \left(\frac{v}{v_0}\right)^4 - \left[\frac{s_0 + vT + v\Delta v(2\sqrt{ab})^{-1}}{h - l}\right]^2\right\} \tag{1}$$

$$\dot{v}_{ACC} = k_1(h - l - s_0 - t_a v) + k_2 \Delta v \tag{2}$$

$$\dot{v}_{CACC} = \frac{k_p(h - l - s_0 - t_c v) + k_d \Delta v}{\Delta t + k_d t_c} \tag{3}$$

where $\dot{v}$ is the target acceleration; $v$ is the current speed; $a$ is the maximum acceleration; $v_0$ is the speed of free flow; $s_0$ is the stopping distance; $T$ is the safe time headway; $\Delta v$ is the speed difference between the vehicle ahead and the vehicle behind; $\Delta t$ is the speed update interval; $b$ is the maximum deceleration; $h$ is the space headway; $l$ is vehicle length; $k_1, k_2, k_p, k_d$ are the control factors; $t_a$ is the desired time headway under ACC: and $t_c$ is the desired time headway under CACC.

*Table 1 – Simulation scene parameter setting*

| Parameters | Typical value |
|---|---|
| The total length of the road | 2 km |
| Simulation duration | 2500 s |
| Vehicle entry time | 400 s |
| Accident vehicle departure time | 500 s |
| Road speed limit | 22.22 m/s (80 km/h) |
| Car intake rate | 8000 veh/h |
| Simulation step | 0.1 s |
| Accident duration | 1200 s |

The parameters of these car-following models are set to default values by SUMO and are listed in *Table 2*. Car is the most common vehicle type on urban expressways in daily life, so we set vehicle parameters as *Table 2*. No difference in vehicle size is assumed between the conventional vehicles and ICVs, with detailed vehicle parameter settings presented in *Table 3*.

*Table 2 – Car following parameters used in the simulation*

| Parameters | Typical value |
|---|---|
| Vehicle length | 5 m |
| Vehicle width | 1.8 m |
| Vehicle height | 1.5 m |
| Maximum vehicle speed | 33.3 m/s |
| Maximum acceleration | 2.6 m/s$^2$ |
| Maximum deceleration | -4.5 m/s$^2$ |
| Emergency braking | -9.0 m/s$^2$ |
| Minimum space gap | 2.5 m |
| Desired time headway | 1 s |

Given the selected vehicle type, the fuel consumption of vehicles is estimated by the default emission model (i.e. PC_G_EU4 emission class of HBEFA3 model) in SUMO. The model is implemented by extracting the data from HBEFA and fitting them to a continuous function by simplifying the function of the power the vehicle engine must produce to overcome the driving resistance force [26]:

$$Fuel = c_0 + c_1 va + c_2 va^2 + c_3 v + c_4 v^2 + c_5 v^3 \tag{4}$$

where $v$ is the speed of vehicle and $a$ is the acceleration of vehicle. $c_0$ to $c_5$ do the HBEFA data evaluate the linear coefficients. The same functional form has been used for all emission types, and only the parameters change per emission type and vehicle.

## 4. FUEL CONSUMPTION CHARACTERISTICS ANALYSIS

Based on the obtained traffic flow and fuel consumption data in Section 3, fuel consumption characteristics analysis will be carried out from the perspectives of lanes and vehicles, respectively.

### 4.1 Aggregate analysis of fuel consumption by lane

In this section, combined with traffic flow and vehicle trajectory characteristics analysis, the total fuel consumption performance of the lane will be analysed from an aggregated level.

As for the relationship between traffic flow characteristics and lane-level fuel consumption, flow-density curve is first examined for different car-following modes. The flow-density data are derived by setting the detection step of the induction loop detectors to 30 s. For the lane directly affected by the accident (i.e. the middle lane on which the rear-end collision occurs), the flow-density relationships for the three car-following modes are shown in *Figures 2a–2c*). It can be seen that the flow-density scatter points of different car-following modes in a rear-end collision accident all match well with the primary flow-density curve derived by Qin et al. through the car-following model formula [15, 27].

The flow-density plots of the three modes are integrated to better observe the difference in the flow-density curves among different car-following modes, as shown in *Figure 2d*. Three noticeable patterns can be observed:

1) In the free-flow region, the scatter points of the three car-following modes overlapped relatively well, and there is no significant difference in either the density of the scatter points or the maximum flow values.
2) In the congestion flow region, the scatter points of CACC and ACC following modes can be well combined into an inverse $\lambda$ curve, where the CACC and ACC points occupy the upper left part and the lower right part of the curve, respectively. This indicates a lower speed performance of ACC compared with CACC in congestion, which does not comply with drivers' desires.
3) In the congestion flow region, the scatter points of IDM are more dispersive and present higher density characteristics than CACC and ACC, indicating the instability of traffic flow in IDM mode during congestion. Such instability may lead to unnecessary energy consumption, further discussed in the following section.

The distribution characteristics of the flow-density plots above indicate that compared with IDM and ACC, the vehicles in the CACC mode have better flow stability and speed performance in the rear-end accident
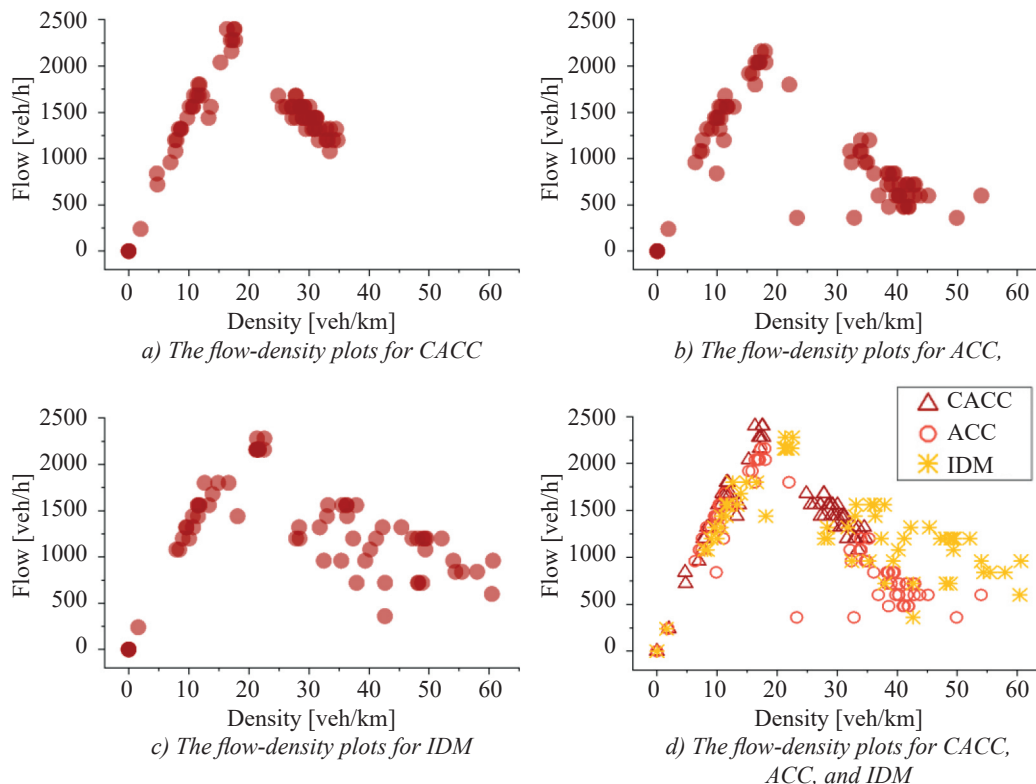


*a) The flow-density plots for CACC*

*b) The flow-density plots for ACC,*

*c) The flow-density plots for IDM*

*d) The flow-density plots for CACC, ACC, and IDM*

*Figure 2 – Fundamental diagrams at different car-following modes*

scenario. To further explore the internal relationship between traffic flow characteristics and fuel consumption for the lane directly affected by accident (Lane_1), Pearson's correlation is measured between the lane-level traffic flow data (including queue length, lane density etc. as shown in *Table 3*) and the fuel consumption data, with all data retrieved every 10s in simulation. Pearson correlation, also known as the product difference correlation or product-moment correlation, is a method to calculate the linear correlation between variables proposed by British statistician Karl Pearson [28]. For any two features, X and Y, their Pearson correlation coefficients can be calculated by the following formula:

$$\rho_{X,Y} = \frac{\Sigma(X - \overline{X})(Y - \overline{Y})}{\sqrt{\Sigma(X - \overline{X})^2 \Sigma(Y - \overline{Y})^2}} \tag{5}$$

In general, the correlation strength of variables can be judged by the following value ranges: 0.8–1.0 strong correlation; 0.6–0.8 strong correlation; 0.4–0.6 moderate correlation; 0.2–0.4 weak correlation; 0.0–0.2 very weak or no correlation. The correlation test results are shown in *Table 3*.

*Table 3 – Correlation analysis of vehicle characteristic parameters on the lane and total fuel consumption of lane*

| Car-following mode | | CACC | ACC | IDM |
|---|---|---|---|---|
| Queue length | Correlation | 0.1977 | 0.4325 | 0.6541 |
| | Significance | 0.0017 | 8.12e-13 | 6.61e-32 |
| Lane density | Correlation | 0.7152 | 0.6628 | 0.9662 |
| | Significance | 1.81e-40 | 5.35e-33 | 4.45e-148 |
| Mean speed | Correlation | 0.6026 | 0.2930 | 0.1105 |
| | Significance | 4.19e-26 | 2.44e-06 | 0.0812 |
| Mean time loss | Correlation | 0.6119 | 0.6061 | 0.9455 |
| | Significance | 4.52e-27 | 1.81e-26 | 7.49e-123 |
| Number of stop-starts | Correlation | 0.2850 | 0.4927 | 0.7867 |
| | Significance | 4.65e-06 | 1.08e-16 | 7.06e-54 |

As can be seen from *Table 3*, among the five traffic flow characteristics, lane density had the strongest correlation with the total fuel consumption of the lane, which can be explained as the higher the density is, the larger the number of vehicles in the counted fuel consumption for a certain length of the lane. For vehicles in the IDM car-following mode, the vehicle mean time loss and the number of stop-starts of vehicles are also highly positively correlated with fuel consumption. This indicates that during the accident, the fuel consumption of the vehicles in the IDM car-following mode increases significantly with increasing vehicle time loss and stop-starts. However, these parameters have weak impact on the CACC car-following mode, which indicates CACC car-following mode has better stability and is similar to the stability performance results from the flow-density plots analysis above. On the other hand, the correlation between average vehicle speed and fuel consumption is moderate for the CACC car-following mode, while it is weak for the IDM and ACC car-following modes. This may be related to the relatively high-speed performance of vehicles in the CACC mode, as shown in the flow-density plots, which will be further analysed at the individual vehicle level in the following sections.

In this paper, we will further analyse the performance of total fuel consumption of the lane in combination with individual vehicle driving trajectories. Taking the accident lane (Lane_1) as the target lane, the driving trajectories of the vehicles on the lane are extracted to observe the stability performance of the homogenous traffic flows in the three car-following modes and their total lane fuel consumptions. *Figure 3* shows the spatial-temporal trajectory diagrams of vehicles in the CACC, ACC and IDM car-following modes on Lane_1. The horizontal axis is the simulation time, and the vertical axis is the longitudinal displacement of the vehicle from the simulation starting point. Each line corresponds to the trajectory of a vehicle. *Figures 4a–4c* show the total fuel consumption curve of the vehicles in the CACC, ACC and IDM car-following modes on Lane_1 from the beginning to the end of the accident, respectively. The horizontal axis is the simulation time, and the vertical axis is the total fuel consumption of all vehicles on the lane at each time. Besides, *Figure 4d–4f* shows the number of traveling vehicles of different car-following modes on Lane_1 during the accident. Three main points can be noted from *Figure 3* and *Figure 4* as follows:

*a) Global trajectory of the vehicle in CACC*



*b) Global trajectory of the vehicle in ACC*



*c) Global trajectory of the vehicle in IDM*



*d) Local trajectory of the vehicle in CACC*



*e) Local trajectory of the vehicle in ACC*
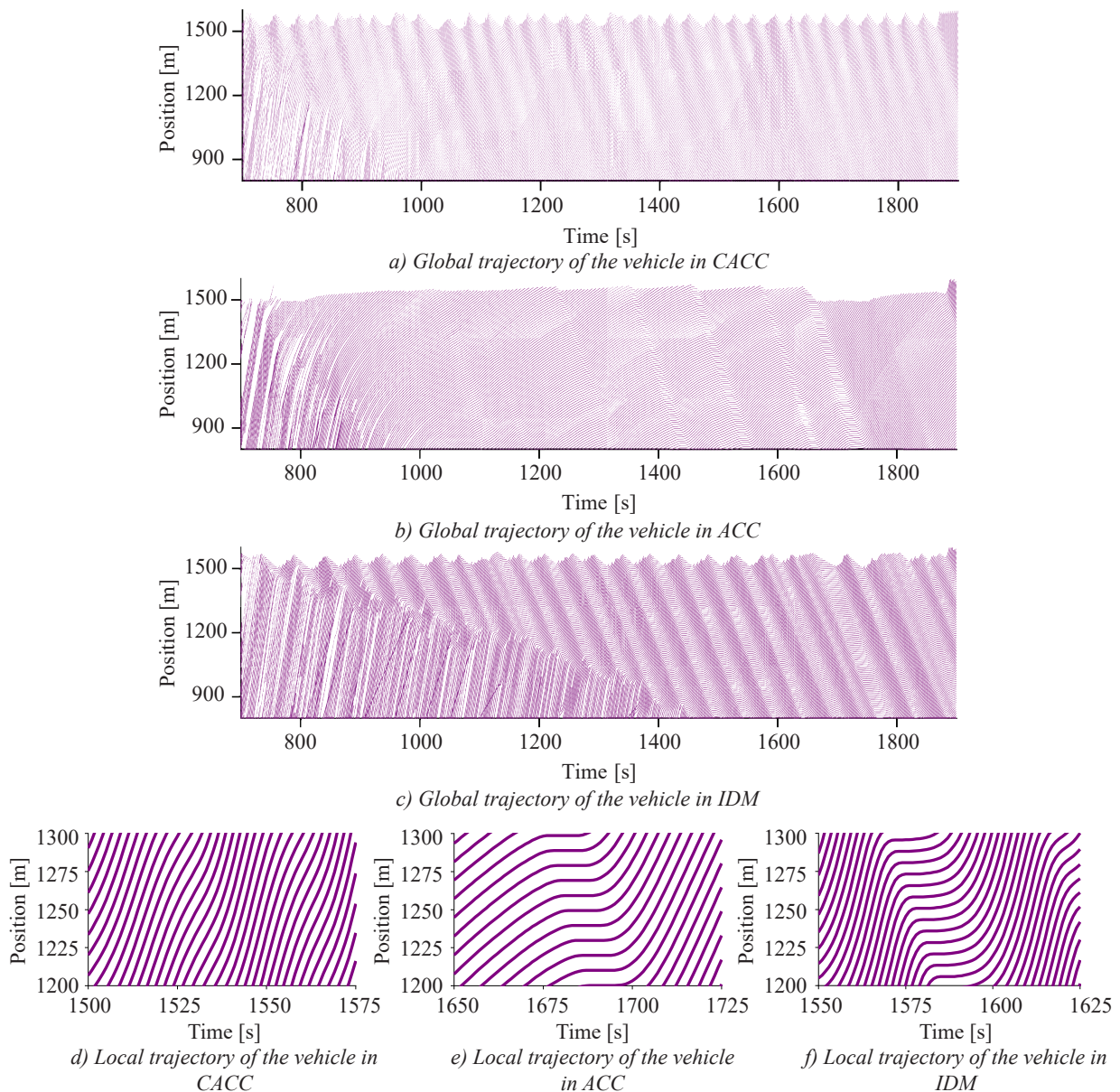


*f) Local trajectory of the vehicle in IDM*

*Figure 3 – Vehicle trajectory map at Lane_1*

1) IDM: unstable vehicle trajectories and high fuel consumption level. Unlike the CIVs with relatively strong environment perception abilities, manual driving vehicles tend not to adjust their driving state until they are very close to the accident site or the affected vehicles. Thus, compared with the CACC and ACC modes, the accident impact back-propagation speed is much lower for the IDM mode (see vehicle trajectories in *Figures 3a–3c*, resulting in a much smaller amount of time and space for the following IDM vehicles to perceive and react to the accident. In addition, once IDM vehicles are affected by an accident, they will enter a stop-and-go state (see *Figure 3f*). In this state, an affected vehicle accelerates and decelerates frequently, and its speed fluctuates greatly, generally raising the vehicle's fuel consumption at a certain speed level. With the increasing number of affected vehicles, the total fuel consumption of the entire road for the IDM mode also rises rapidly. It will likely result in a much higher fuel consumption level than the other two car-following modes when the accident impact propagates throughout the simulation scene (see *Figures 4c and 4f*). Furthermore, as the space headway of vehicles in the stop-and-go state is smaller than that of free flow, the lane density will increase continuously, leading to apparent congestion and queuing phenomenon.

2) ACC: relatively stable vehicle trajectories and medium fuel consumption level. Thanks to the strong environment perception capability, vehicles in the ACC mode have sufficient time and space to make adjustments from the initial stage of the accident. During this stage, the affected vehicles slow down gradually in a relatively stable driving state, and the total fuel consumption level of the lane is relatively
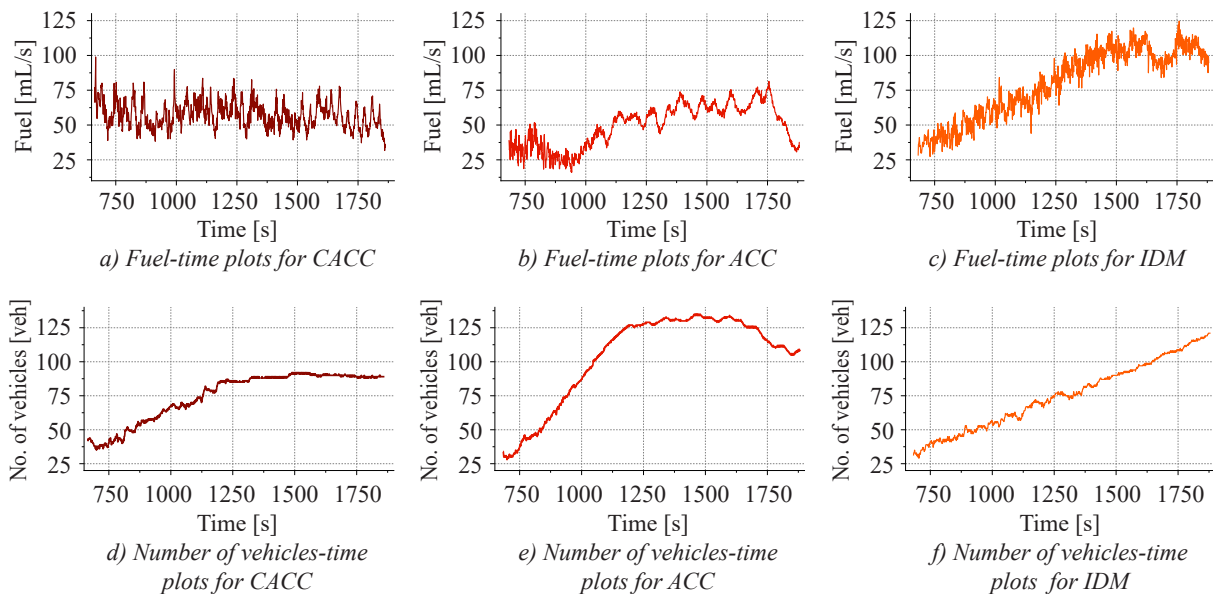
*a) Fuel-time plots for CACC*

*b) Fuel-time plots for ACC*

*c) Fuel-time plots for IDM*

*d) Number of vehicles-time plots for CACC*

*e) Number of vehicles-time plots for ACC*

*f) Number of vehicles-time plots for IDM*

*Figure 4 – Fundamental diagrams at Lane_1*

low. However, shock waves are formed as the accident car continues blocking the lane when the following vehicles wait to change lanes. Such shock waves will be transmitted backward, and the affected vehicles will gradually exhibit an unstable driving state (stop-and-go state, see *Figure 3e*). However, it is not as frequently observed as in the IDM mode. In this stage, the total fuel consumption of the lane increases but still is far lower than in the IDM mode at the same stage. Around 1700 s, in the simulation, a gap in the traffic stream appeared by coincidence. Many vehicles performed lane changes easing the lane's congestion situation before the accident ended. However, one can speculate that the congestion will likely worsen if the accident continues due to the accumulated accident impact. In addition, although the ACC mode shows a low fuel consumption level, its relatively low-speed performance (shown in the flow-density plot) reflects a low traffic efficiency, which does not adapt to the driving needs of drivers.

3) CACC: stable vehicle trajectories and low fuel consumption level. It can be seen from *Figure 3d* that the driving state of the vehicles in the CACC mode is relatively stable under accident conditions. Compared with the driving state in free flow, only decelerations of a certain degree are observed when the accident impact propagates. The stop-and-go data from the detector show that the vehicle stopping time is all zero for the CACC mode, indicating that no vehicle has come to a full stop during the accident. Additionally, it can be observed from the local trajectory diagram that the vehicles in the CACC mode have a relatively small space headway and a higher driving speed, which demonstrates that the CACC car-following mode has a larger road capacity during the accident. The fuel consumption curve also shows a stable performance for the vehicles in the CACC mode. The total fuel consumption curve of the lane fluctuates in the range of 50–75 mL/s and does not increase or decrease over time under accident conditions.

In summary, in the three-lane scene with a rear-end collision, vehicles in the CACC car-following mode perform better in terms of road capacity, driving stability and total fuel consumption than IDM and ACC. The vehicles affected by the shock wave (due to accident-caused lane changes) are still able to maintain a small space headway as well as faster driving speed in the CACC mode, indicating that the CACC mode has a strong ability to dissipate the impact of the accident on traffic flow stability, which is of great significance for further in-depth study. The results of this section prove that intelligent networking driving technology is one of the most important means to save energy, reduce emissions and improve road capacity even in an extreme driving environment such as rear-end collision. In the following section, we will analyse the fuel consumption performance of the three car-following modes based on individual vehicles' driving data.

## 4.2 Disaggregate analysis of fuel consumption by vehicle

In the previous section, the total fuel consumption performance of the lane was analysed from a macro perspective. In this section, we will focus on the individual fuel consumption performance of vehicles. The

middle lane (Lane_1, directly affected by the rear-end collision) is again taken as the target research area. The instantaneous parameters of each vehicle in this area are extracted, including time, position, speed and fuel consumption. They are presented as contour maps, as shown in *Figures 5–6*.

It could be noted that the temporal-spatial region affected by accident can be more intuitively identified in the contour map. From the speed contour maps, it can be seen that for both the affected and unaffected regions, the vehicles in the CACC mode have a significantly higher driving speed than the other two car-following modes. The vehicles in the ACC mode will drive at a very low speed after entering the affected region, and no significant fluctuation is observed in driving speed. The vehicles in the IDM mode will be very unstable in the affected region, where their speed will change frequently. The vehicles in the CACC car following mode have a more significant advantage in speed performance and can better meet drivers' needs. In contrast, the frequent changes in driving state in the IDM mode and the continuous low-speed driving in the ACC mode tend to cause drivers' impatience to a certain degree.
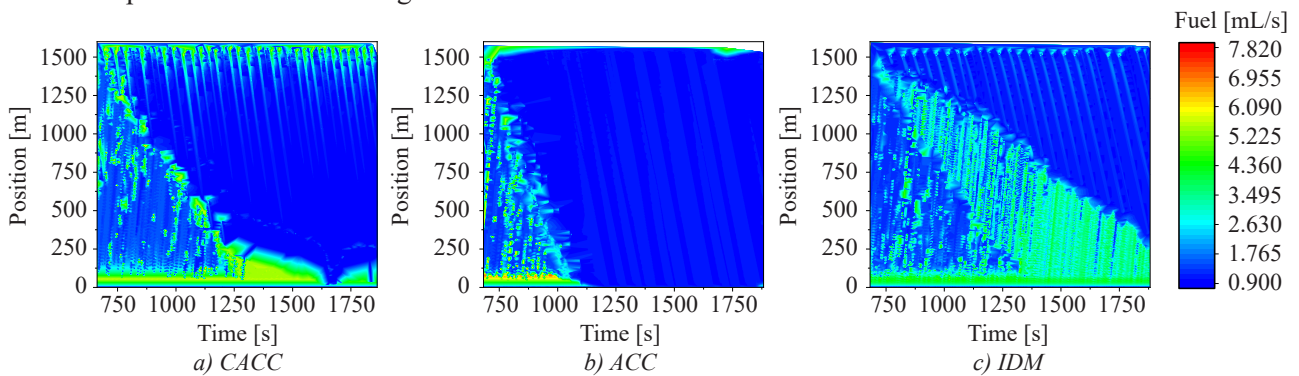


*a) CACC*        *b) ACC*        *c) IDM*

*Figure 5 – Time-space distribution diagram of fuel consumption*



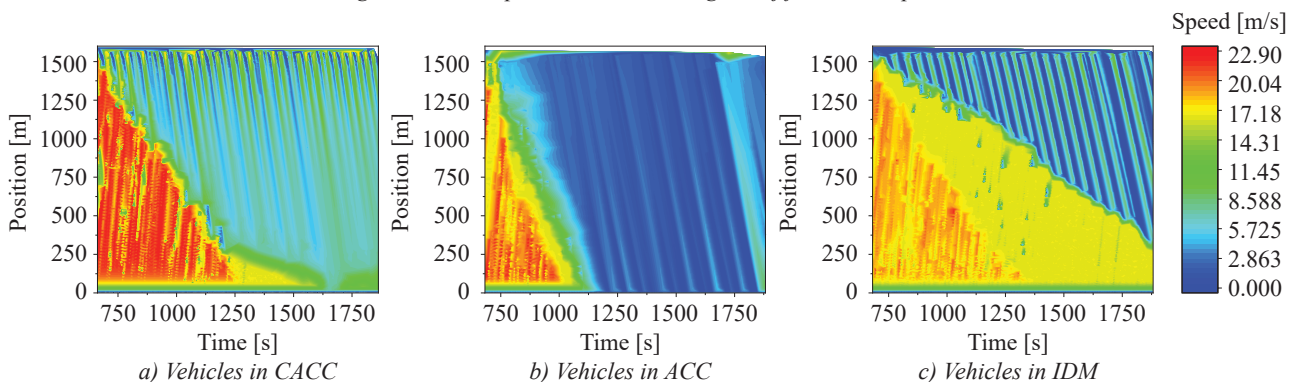*a) Vehicles in CACC*        *b) Vehicles in ACC*        *c) Vehicles in IDM*

*Figure 6 – Time-space distribution diagram of vehicle speed*

In the contour maps of fuel consumption, it can be noted that individual vehicles' fuel consumption in the accident-affected region is at a very low level for all three modes. Note that this does not contradict the results from the previous section, which shows an increasing trend in the total fuel consumption of the lane when the accident blocking continues for the IDM and ACC modes. This is because under the same lane length, the space headway of vehicles at a low speed is smaller than that in free flow, and thus the total number of vehicles on the lane is much larger for a low-speed flow than a free flow. Consequently, even when the fuel consumption of an individual vehicle is reduced due to the lower driving speed on average, the total fuel consumption of the lane will still increase. This also explains the highest correlation between the total lane fuel consumption and lane density in *Table 3*. From the perspective of instantaneous fuel consumption of individual vehicles, the ACC mode seems to exhibit the best performance among the three car-following modes. For the CACC car-following mode, the fuel consumption of an individual vehicle increases in the area near the accident-affected region. Still, it then drops to a very low level in the lower part of the accident-affected region. For the IDM car-following mode, the individual vehicle's fuel consumption characteristics are similar to that of driving speed in the accident-affected region, both exhibiting an unstable state.

Note that the analysis above is solely based on the instantaneous fuel consumption level. However, the fuel consumption level of the entire lane is also affected by the passing time of vehicles. Thus, the driving time and

total fuel consumption of each vehicle on this lane are extracted and shown in the box line diagram in *Figure 7* (marked with the specific mean values). It can be seen that the average total fuel consumption of vehicles in the ACC mode is 97.32 mL, which is far lower than 141.12 mL in the ACC mode and 142.05mL in the IDM mode.

Moreover, the total fuel consumption chart of vehicles in the CACC mode is flatter than in the other two modes, indicating that the fluctuation of the total fuel consumption of vehicles in the CACC mode is relatively small. Regarding the passing time, the average driving time of vehicles in the CACC mode is 127.49 s, slightly lower than 134.36 s in the IDM mode and far lower than 294.02 s in the ACC mode. In general, even though the fuel consumption of individual vehicles in the CACC mode is not the lowest in a congested area, individual vehicles in the CACC mode have minimal total fuel consumption due to great traffic efficiency and energy-saving ability. In detail, vehicles in CACC modes perform significantly better than vehicles in the ACC mode in terms of total driving times (*Figure 7b*) and perform significantly better than vehicles in the IDM mode in terms of fuel consumption rate (*Figure 5*).
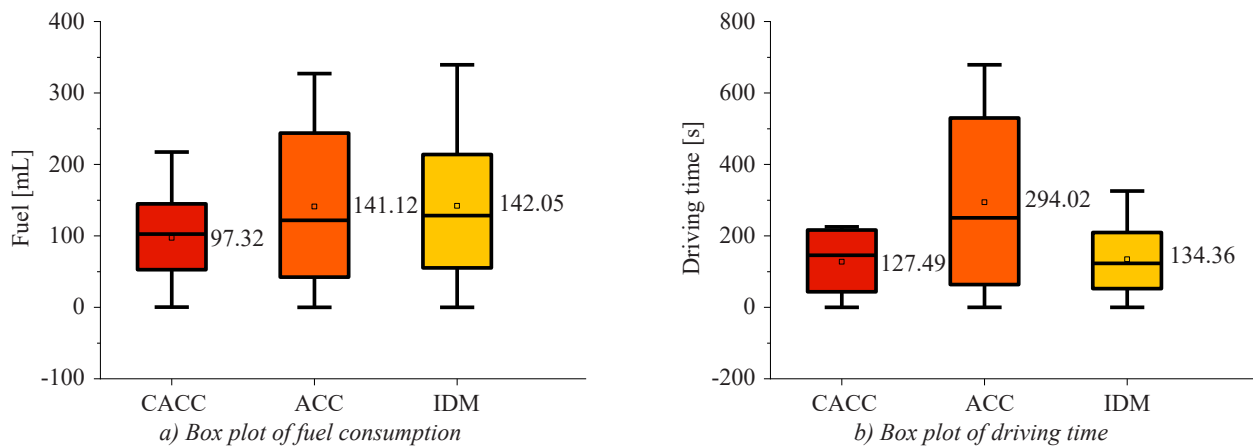


a) Box plot of fuel consumption

b) Box plot of driving time

*Figure 7 – Average parameters of vehicles passing through the accident section in different car-following modes*

## 4.3 Summary of vehicle fuel consumption characteristics

As analysed in the lane-level fuel consumption section, the curve of total fuel consumption of Lane_1 for the CACC mode is found to be the most stable, while it shows an increasing trend when the duration of the accident increases for the IDM and ACC modes. Specifically, the total lane fuel consumption in the IDM mode rises far higher than that in the CACC and ACC modes. Such increase is largely attributed to the increasing number of vehicles on the lane due to a minor space headway in the accident-affected region. However, the instantaneous vehicle fuel consumption after entering the accident-affected region is found to be lower than that in the free flow state, as shown in the contour maps of *Figure 5*. As to the instantaneous fuel consumption characteristics of individual vehicles in the accident-affected region, (1) vehicles in the ACC mode always maintain a low fuel consumption characteristics; (2) vehicles in the CACC mode also have a very low instantaneous fuel consumption rate when they are in the lower half part of the accident-affected region (with a relatively sizeable longitudinal distance from the crash site), and the rate is kept low until they get close to the crash site; (3) vehicles in the IDM mode in the accident-affected region show an unstable changing state in instantaneous fuel consumption, similar to their speed characteristics in the region. Finally, from the perspective of the driving process of an individual vehicle, the vehicle's passing time and the total fuel consumption on Lane_1 are assessed, and the results show that vehicles in the CACC mode show an obvious advantage in fuel consumption. The average value and interquartile range of total fuel consumption of vehicles in the CACC mode are much smaller than that of vehicles in the other two car-following modes, indicating that the CACC mode has the best and the most stable performance in fuel consumption at both the lane and individual vehicle levels.

From the above research on the characteristics of vehicle fuel consumption in a rear-end collision scene, we believe that intelligent network technology is vital for energy saving and emission reduction of vehicles and improving traffic efficiency in congested environments. In rear-end collision scenes, the relatively weak environment perception ability of vehicles in the IDM mode leads to a slower propagation of the impact of the accident, resulting in insufficient time and space for vehicles to adjust their driving state after entering the accident-affected region, where frequent stop-and-go vehicle actions are observed and generally lead to

relatively high fuel consumption. On the contrary, thanks to the relatively strong environmental awareness of vehicles in the ACC mode, following vehicles can learn about the accident occurrence earlier and thus have sufficient time and space to adjust the speed, as reflected by the general deceleration of vehicles in the accident-affected region. Although the individual ACC vehicle's instantaneous fuel consumption rate is very low in this region, the total fuel consumption for the vehicle to pass the accident section is still very high due to the long passing time (low driving speed). In contrast, the CACC mode has a better ability to dissipate congestion. The CACC vehicles in the accident-affected region decelerate by a certain extent compared with the free flow. Their speed does not show significant fluctuation until they approach the crash site. Such fluctuation is due to the mandatory lane change behaviour, and its back-transmission distance is very short for the CACC mode. Thus, the driving state of the CACC vehicle is relatively stable and yields good performance in all fuel consumption analyses.

## 5. VEHICLE FUEL CONSUMPTION PREDICTION AND EVALUATION

To increase vehicle awareness of the fuel consumption of subsequent driving in advance and to improve path planning, we have attempted to build a fuel consumption prediction framework based on machine learning, with details elaborated in the following section.

### 5.1 Framework of vehicle fuel consumption prediction and evaluation

As demonstrated in previous studies, the fuel consumption prediction can be treated as a regression problem [6, 19, 29, 30], where vehicle fuel consumption rate can be estimated by both vehicle motion parameters and prior environment knowledge. This study selects vehicle speed, acceleration, space headway, time and space position as the model inputs, with vehicle fuel consumption at the exact moment as the model output. Specifically, the time and space position can reflect the influence degree of the accident on the vehicle. The speed and acceleration capture the driving behaviour of the driver. The distance between vehicle heads indicates the level of congestion on the road. XGBoost, an efficient implementation of gradient boosting, establishes the fuel consumption regression model.

The process of constructing a fuel consumption prediction framework based on vehicle driving data under the influence of rear-end collision is depicted in *Figure 8*. Input and output variables are first extracted based on the collected data in Section 3, which are randomly divided into training, validation, and test sets for developing the XGBoost regression model, the performance of which is evaluated by indicators including R-squared, root mean squared error (RMSE), and mean square error (MAE). The established model will be used for fuel consumption prediction based on historical driving data and further serve for path planning, and will be elaborated in the accident impact region analysis of Section 5.3.

XGBoost is a supervised learning algorithm implemented by gradient tree boosting, which is used to deal with machine learning problems such as classification and regression. This study uses XGBoost based on tree structure to predict vehicle fuel consumption due to rear-end accidents. The model of XGBoost is defined as follows [29]:

$$\hat{y}_i = F_K(x_i) = F_{K-1}(x_i) + f_K(x_i) \tag{6}$$

where $x_i$ is the $i^{th}$ training sample, $f_K(x_i)$ represents the $k^{th}$ decision tree, and the decision tree will map the sample features so that each sample falls on a particular leaf node of the tree. Specifically, each leaf node has a weighted score representing the predicted value $\omega$ of the sample in the tree. Then, the sum of the predicted value $\omega$ of the sample in each tree is calculated as the final predicted value of this sample. The target function of XGBoost is defined as follows:

$$Obj = \sum_{i=1}^{n} L(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{7}$$

where $\sum_{i=1}^{n} L(y_i, \hat{y}_i)$ is the loss function, which is applied to evaluate the loss or error between the predicted value of the model and the true value. $\sum_{k=1}^{K} \Omega(f_k)$ is the regularisation term used to control the complexity of the model. The regularisation term tends to choose simple models to avoid overfitting, which is defined as follows:

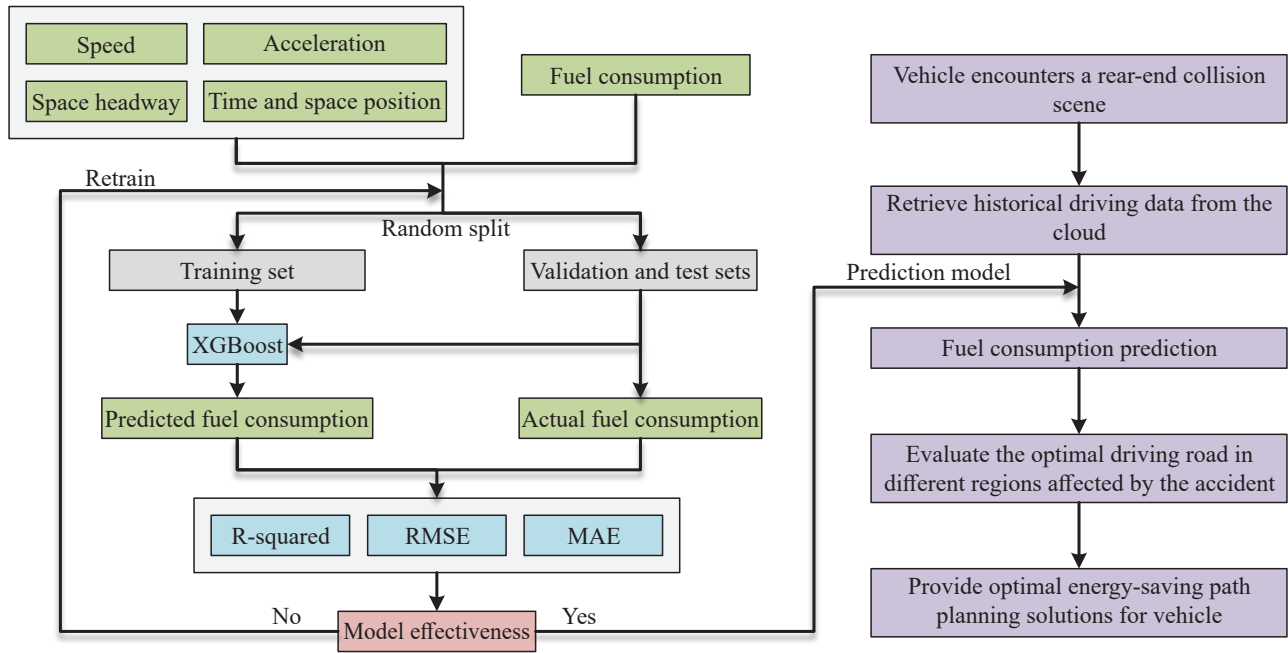$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda \|\omega\|^2 \tag{8}$$

*Figure 8 – Framework diagram of fuel consumption prediction model*

where $\gamma$ and $\lambda$ are the constant coefficients; $T$ is the total number of leaf nodes of the tree; $\gamma T$ is used to control the complexity of the tree so as to limit the complexity of the model. Meanwhile, $\frac{1}{2}\lambda\|\omega\|^2$ is used to control the weight score of the leaf node. In the objective function, $\hat{y}_i^s = \hat{y}_i^{(s-1)} + f(x_i)$, , where $s$ is the number of iterations. The model of $s^{th}$ iteration equals to the model of the earlier $(s-1)^{th}$ iteration plus a new sub-model. Taylor's expansion of the objective function is depicted in the following equation:

$$Obj^{(s)} \approx \sum_{i=1}^{n}\left[ L(y_i,\hat{y}^{(s-1)}) + g_i f_s(x_i) + \frac{1}{2}h_i f_s^2(x_i) \right] + \Omega(f_s) \tag{9}$$

where $g_i = \partial_{\hat{y}^{(s-1)}}L(y_i,\hat{y}^{(s-1)})$ represents the one-step statistics of the loss function. $h_i = \partial_{\hat{y}^{(s-1)}}^2 L(y_i,\hat{y}^{(-1)})$ is the second-order gradient statistics. Since the constant term does not affect the optimisation result, after removing the constant term, the objective function can be rewritten to unify the weight of the sub-model and leaf node:

$$Obj^{(s)} = \sum_{j=1}^{T}\left[ \left(\sum_{i\in I_j} g_i\right)\omega_j + \frac{1}{2}\left(\sum_{i\in I_j} h_i + \lambda\right)\omega_j^2 \right] + \gamma T \tag{10}$$

where $I_j$ is the sample set of leaf node $j$, i.e. all the samples landed on the leaf node $j$. $f_s(x_i)$ divides the samples onto leaf nodes and calculates the score $\omega$ of the leaf node. Therefore, $f_s(x_i)$ is replaced with $\omega_j$ when $i\in I_j$.

Through the formula above, the optimal weight $\omega_j$ and objective function $Obj^{(s)}$ of a specific tree structure can be calculated. Then the greedy algorithm can be used to evaluate its node splitting. Finally, the optimal splitting can be identified, and the best tree structure can be obtained. As XGBoost has fast computation speed and can effectively avoid overfitting, it can be well applied to the real-time fuel consumption prediction problem in the study and thus is selected.

To assess the performance of the proposed fuel consumption regression model based on XGBoost, the following indicators are selected for evaluation:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - p_i)^2}{\sum_{i=1}^{N}(y_i - \overline{y})^2} \tag{11}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - p_i)^2}$$ (12)

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - p_i|$$ (13)

where $N$ is the sample size, $y_i$ represents the true value of the $i$th sample and $p_i$ represents the predicted value of the $i$th sample. Specifically, R-squared evaluates the overall fitting degree of the XGBoost model. The closer its value is to 1, the higher the fitting degree of the model. RMSE and MAE measure the accuracy of the model prediction, and the lower the value, the higher the model's accuracy.

### 5.2 Fuel consumption prediction under the impact of accidents

This section randomly selects 10 vehicles that departed in 750~1650 s of simulation. Their driving data are extracted at an interval of 100 s, 80% of which are selected as the training sample and the remaining 20% as the validation sample. In addition, the driving data of 3 other vehicles are randomly selected as the test set. As illustrated in *Figure 8*, the vehicle's speed, acceleration, space headway, time and space position are utilised as inputs, and the instantaneous fuel consumption of the vehicle serves as the output. Examples of model input and output based on the CACC vehicle data for model training and validation are listed in *Table 4*.

*Table 4 – Input and output examples of CACC vehicle model training and validation*

| Input/ Output | Parameters | Unit | Training | | | ... | Validation | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Sample$_1$ | Sample$_2$ | Sample$_3$ | | Sample$_1$ | Sample$_2$ | Sample$_3$ |
| Input | Speed | m/s | 5.67 | 2.4 | 3.74 | ... | 20.69 | 22.21 | 2.68 |
| | Acceleration | m/s$^2$ | -0.2 | -1.2 | 0.10 | ... | 0.0 | 0.0 | 0.0 |
| | Space headway | m | 15.65 | 12.38 | 13.77 | ... | 41.23 | 43.33 | 12.68 |
| | Time position | s | 1353.3 | 1219.8 | 1742.9 | ... | 1484.8 | 1592.3 | 1888.8 |
| | Space position | m | 838.0 | 1505.28 | 1375.4 | ... | 1852.52 | 1928.32 | 1110.24 |
| Output | Fuel consumption | mL/s | 0.0 | 0.0 | 1.02 | ... | 1.42 | 1.57 | 1.0 |

The XGBoost model is developed in the PyCharm platform with Python version 3.7. The optimal values of the hyper-parameters of XGBoost, such as the maximum tree depth, the minimum leaf node sample weight and learning rate etc., are obtained based on the selected evaluation indexes R-squared, RMSE and MAE. Given the validation data, the optimal maximum tree depth is 6, the minimum leaf node sample weight is 1, and the learning rate is 0.3.

Finally, based on the test set, the predicted fuel consumption results are presented in *Figure 9*, demonstrating the degree of approximation between the predicted and the true fuel consumption of vehicles in the three car-following modes. In *Figure 9*, the actual fuel data represent the fuel consumption of vehicles in our simulation, which were exported from SUMO. The predicted fuel data are the results of our fuel prediction model. The prediction results for all three modes are generally distributed on both sides of $y=x$, indicating a significantly high goodness-of-fit of the established regression model.
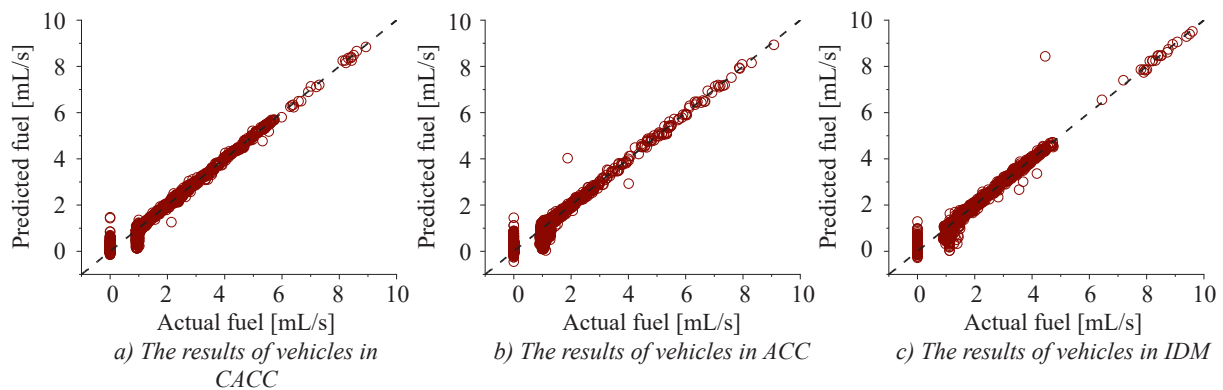


*a) The results of vehicles in CACC*   *b) The results of vehicles in ACC*   *c) The results of vehicles in IDM*

*Figure 9 – Fuel consumption prediction results*

The calculated evaluation indexes R-squared, RMSE and MAE for the validation and test sets are presented in *Table 5*.

*Table 5 – Model evaluation results*

| Evaluation index | | R-squared | RMSE | MAE |
|---|---|---|---|---|
| CACC | Validation set | 0.988544 | 0.127761 | 0.059297 |
| | Test set | 0.990474 | 0.128486 | 0.069068 |
| ACC | Validation set | 0.929834 | 0.200155 | 0.122696 |
| | Test set | 0.941739 | 0.264130 | 0.143675 |
| IDM | Validation set | 0.983515 | 0.131002 | 0.048952 |
| | Test set | 0.950419 | 0.219820 | 0.108931 |

From the evaluation results in *Table 5*, it can be seen that the XGBoost model can be well applied to the prediction of fuel consumption of vehicles in different car-following modes in a rear-end collision accident. The R-squared value of the test set for the three modes all exceed 0.92, indicating that the established model has a high goodness-of-fit of prediction. Meanwhile, RMSE and MAE are small for all three car-following modes, where the maximum instantaneous fuel consumption can reach 7 mL/s.

## 5.3 Accident impact region analysis based on predicted fuel consumption

In addition to evaluating the accuracy of the XGBoost fuel consumption prediction model, the three vehicles of the test set are utilised to draw the fuel consumption time series curve, the driving trajectory and the accident-affected region. The results are depicted in *Figures 10–12*. Note that the three impact regions (i.e. 1–3) in the figures are divided according to the situation of the vehicle affected by accident:

1) In region 1, the vehicle has just entered the simulated lane and is less affected by the blocking caused by the rear-end collision. The vehicle speed is gradually increased close to the free-flow speed. The driving state of the vehicle in this region is relatively stable.
2) In region 2, the blocking affects the vehicle's driving state. Specifically, CACC vehicles have relatively small fluctuations in speed, ACC vehicles are in a continuous low-speed driving state, and IDM vehicles are in a stop-and-go state. As a result, the vehicle's state is unstable at this stage.
3) In region 3, the vehicle drives out of the accident-affected region by changing lanes and exits the simulation scene. The vehicle speed is gradually increased back to the free-flow speed, and the driving state is gradually stabilised again.

From the prediction curves of vehicle fuel consumption, for the CACC car-following mode, the instantaneous fuel consumption of the vehicles is relatively stable in regions 1 and 3. Still, it fluctuates with no apparent pattern in region 2, where fuel consumption is maintained at a low level for almost half the time. In the ACC car-following mode, the fuel consumption curves in the three regions are all relatively stable, with certain fluctuations at the end of region 2. In the IDM car-following mode, the fuel consumption curves in the three regions all demonstrate unstable characteristics, and the fluctuations in region 2 show certain cyclic regularity.

It should be noted that due to the different impact of the accident on each vehicle, the distribution length of the three regions for each vehicle can vary, as well as the driving lanes for each vehicle in each region, which provides space for path planning from the perspective of energy saving. As illustrated in *Figure 8*, when a rear-end collision occurs, given the relative spatiotemporal information between the vehicle and the accident car, the corresponding historical driving data of the same accident scene can be retrieved (for example, from cloud storage). The fuel consumption of vehicles in different driving paths can be predicted using the proposed XGBoost model based on historical data. The optimal combination of driving lanes, the best time for lane changes and other recommended driving behaviour data in each of the three impact regions can be identified, which will help plan the vehicle path from an energy-saving perspective. The details of path planning will not be further discussed as it is not the main focus of this work.
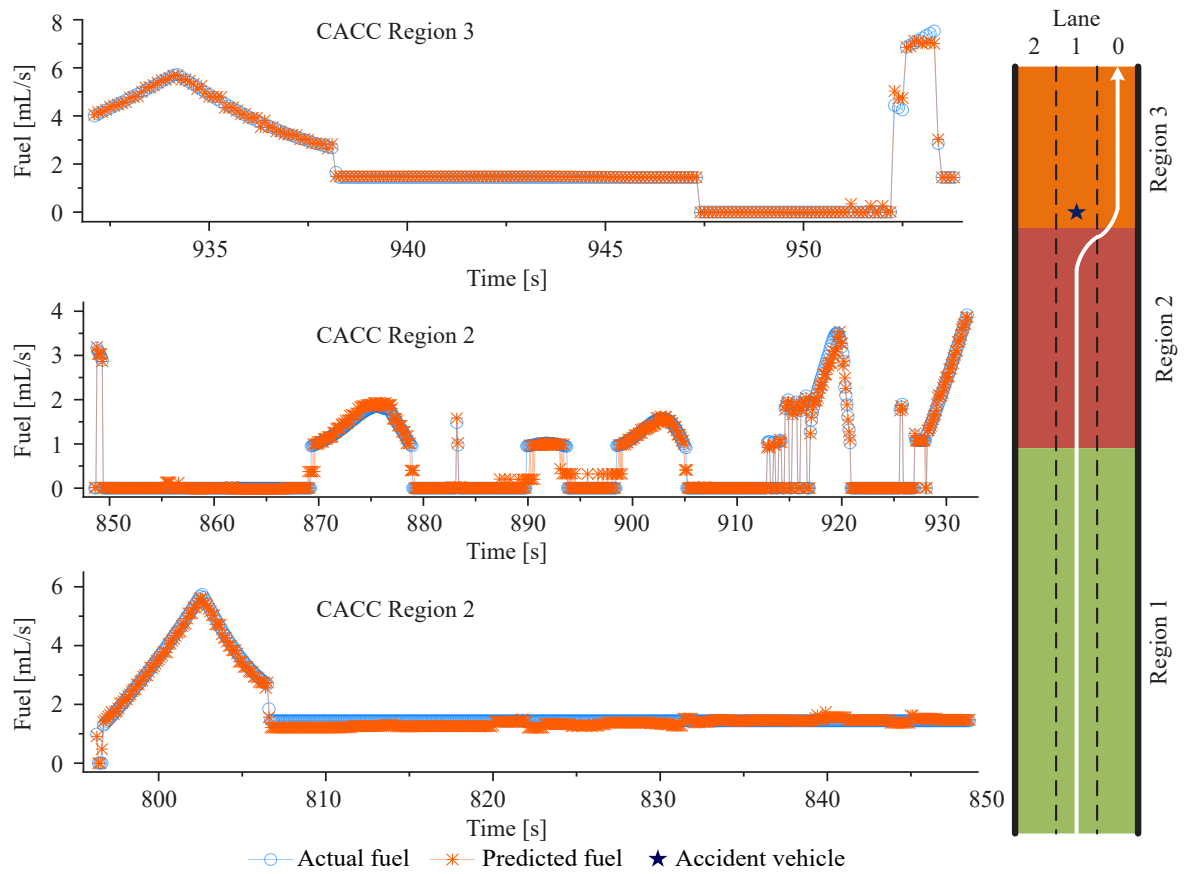
*Figure 10 – The regional division of vehicles in CACC and the fuel consumption forecast map of each region*
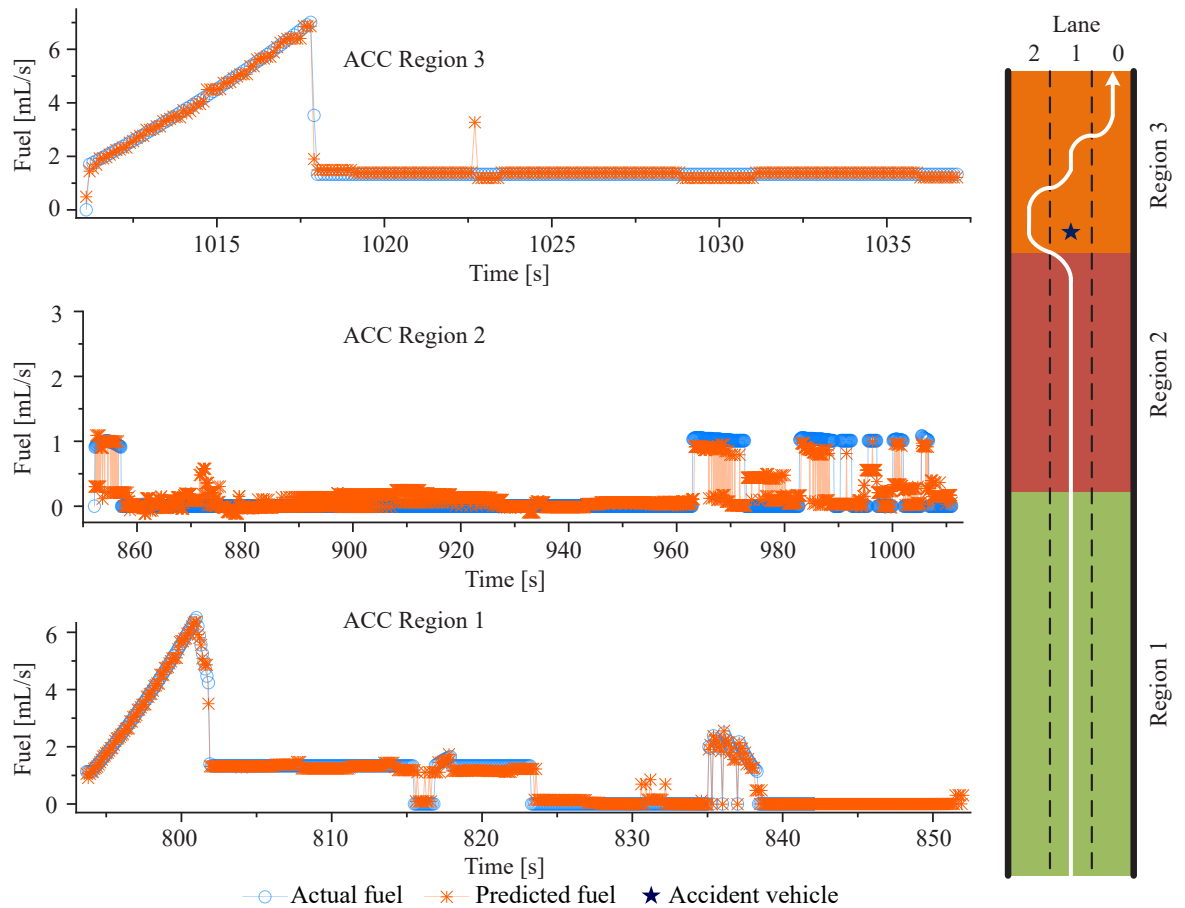


*Figure 11 – The regional division of vehicles in ACC and the fuel consumption forecast map of each region*
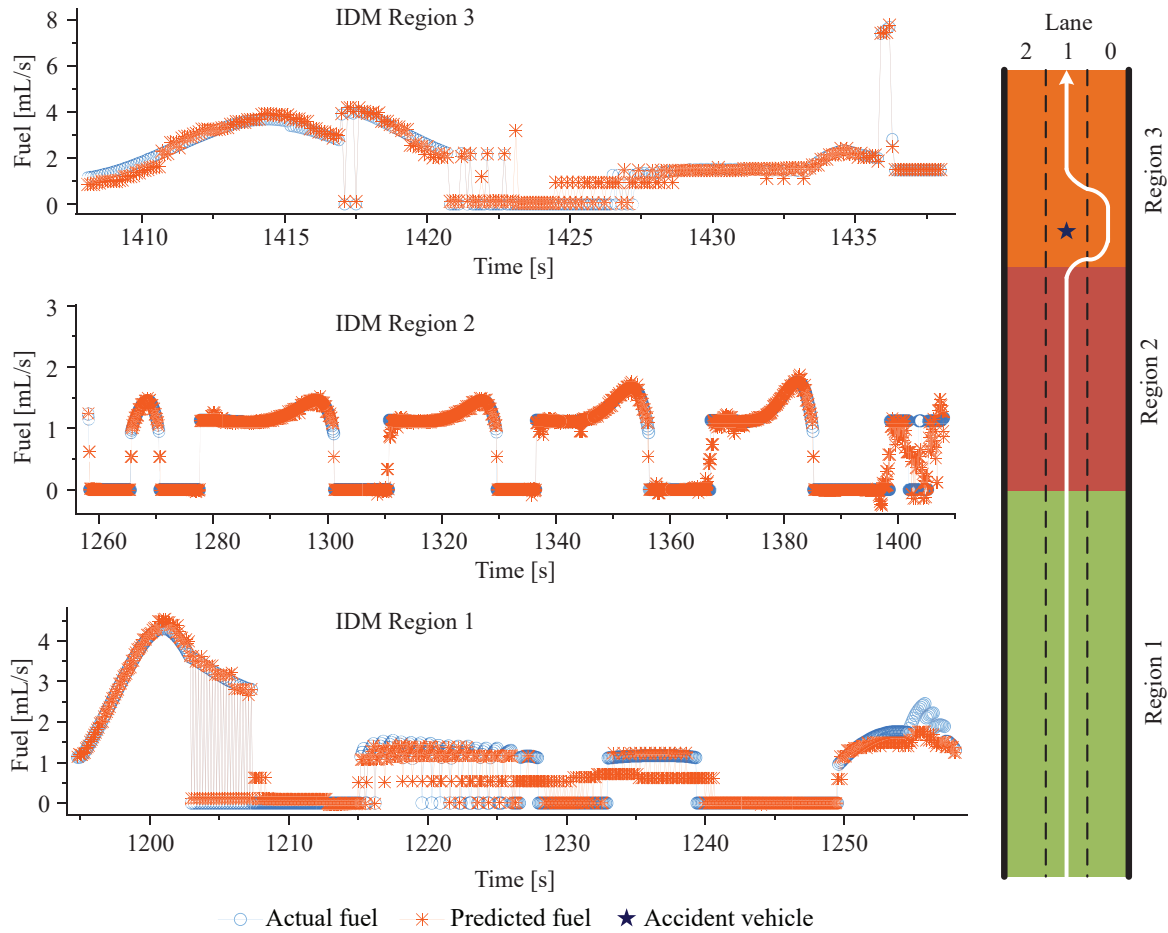
*Figure 12 – The regional division of vehicles in IDM and the fuel consumption forecast map of each region*

## 6. DISCUSSION AND CONCLUSIONS

In this paper, we built a road environment in which a rear-end collision blocks the middle lane of the three-lane urban expressways. From the lane and the individual vehicle levels, we compared and analysed the lane-level fuel consumption characteristics, the fuel consumption rate and driving efficiency of individual vehicles in CACC, ACC and IDM car-following modes. Then we proposed a fuel consumption prediction framework based on XGBoost to predict the fuel consumption rate of vehicles under the influence of rear-end collision and verified its feasibility.

On the one hand, from the results of the total fuel consumption of lanes, vehicles in the CACC car-following mode have the most stable fuel consumption curve. In contrast, the curves for the IDM and ACC modes increase significantly with time under the accident condition. The average total fuel consumption of an individual vehicle in CACC mode is far lower than that of the vehicles in the other two car-following modes. Additionally, the passing time of vehicles in the CACC mode is the shortest.

In summary, due to the relatively weak environment perception ability, vehicles in the IDM mode do not have sufficient time and space to adjust their driving state when exposed to the accident's impact. Thus, they show unstable fuel consumption and speed performance in the accident-affected region. Vehicles in the ACC mode have very low driving speeds in the accident-affected region, resulting in longer passing time and increased fuel consumption. Compared with vehicles in the IDM and ACC modes, vehicles in the CACC mode exhibit low fuel consumption and high traffic efficiency in the simulated accident environment.

Furthermore, based on the vehicle driving data in the three car-following modes, the proposed XGBoost fuel consumption regression model has a high fit and good prediction accuracy. Besides combining vehicle fuel consumption time series with driving trajectories, different vehicle fuel consumption characteristics are found in the three regions with different accident impacts. The CACC and ACC modes' fuel consumption curves

fluctuate in region 2 but are relatively stable in regions 1 and 3. In contrast, the IDM mode curves fluctuate in all three regions, with certain cyclic regularity identified in region 2.

As this study primarily focuses on the fuel consumption characteristics of vehicles in different car-following modes under accident conditions, only homogeneous traffic flow of the CACC, ACC and IDM modes were simulated and analysed. However, the market penetration rates of different car-following modes in the heterogeneous traffic flow may impact fuel consumption. In future work, the market penetration rates of different car-following modes can be used as the main control variable to evaluate its impact on vehicle fuel consumption. Meanwhile, as only rear-end collision condition is considered in this work, subsequent research can be extended to more accident scenarios to validate the proposed fuel consumption forecasting framework.

## REFERENCES

[1] Han X, et al. Energy-aware trajectory optimization of CAV platoons through a signalized intersection. *Transportation Research Part C: Emerging Technologies*. 2020;118. DOI: 10.1016/j.trc.2020.102652.

[2] Flores C, et al. Fractional-order-based ACC/CACC algorithm for improving string stability. *Transportation Research Part C: Emerging Technologies*. 2018;95:381-393. DOI: 10.1016/j.trc.2018.07.026.

[3] Talebpour A, et al. Influence of connected and autonomous vehicles on traffic flow stability and throughput. *Transportation Research Part C: Emerging Technologies.* 2016;71:143-163. DOI: 10.1016/j.trc.2016.07.007.

[4] Ma F, et al. Predictive energy-saving optimization based on nonlinear model predictive control for cooperative connected vehicles platoon with V2V communication. *Energy.* 2019;189. DOI: 10.1016/j.energy.2019.116120.

[5] Vahidi A, et al. Energy saving potentials of connected and automated vehicles. *Transportation Research Part C: Emerging Technologies.* 2018;95:822-843. DOI: 10.1016/j.trc.2018.09.001.

[6] Yao Z, et al. Reducing gasoline consumption in mixed connected automated vehicles environment: A joint optimization framework for traffic signals and vehicle trajectory. *Journal of Cleaner Production.* 2020;265. DOI: 10.1016/j.jclepro.2020.121836.

[7] Mousavi SM, et al. Investigating the safety and operational benefits of mixed traffic environments with different automated vehicle market penetration rates in the proximity of a driveway on an urban arterial. *Accident Analysis and Prevention.* 2021;152:105982. DOI: 10.1016/j.aap.2021.105982.

[8] Zhu J, et al. Safety analysis of freeway on-ramp merging with the presence of autonomous vehicles. *Accident Analysis and Prevention.* 2021;152:105966. DOI: 10.1016/j.aap.2020.105966.

[9] Papadoulis A, et al. Evaluating the safety impact of connected and autonomous vehicles on motorways. *Accident Analysis and Prevention.* 2019;124:12-22. DOI: 10.1016/j.aap.2018.12.019.

[10] Petrović Đ, et al. Traffic accidents with autonomous vehicles: type of collisions, manoeuvres and errors of conventional vehicles' drivers. *Transportation Research Procedia.* 2020;45:161-168. DOI: 10.1016/j.trpro.2020.03.003.

[11] Li Y, et al. Evaluation of the impacts of cooperative adaptive cruise control on reducing rear-end collision risks on freeways. *Accident Analysis and Prevention.* 2017;98:87-95. DOI: 10.1016/j.aap.2016.09.015.

[12] Xiao L, et al. Unravelling effects of cooperative adaptive cruise control deactivation on traffic flow characteristics at merging bottlenecks. *Transportation Research Part C: Emerging Technologies*. 2018;96:380-397. DOI: 10.1016/j.trc.2018.10.008.

[13] Zhu HB, et al. Modeling traffic flow mixed with automated vehicles considering drivers' character difference. *Physica A: Statistical Mechanics and its Applications*. 2020;549. DOI: 10.1016/j.physa.2020.124337.

[14] Zhou YJ, et al. Impact of CACC vehicles' cooperative driving strategy on mixed four-lane highway traffic flow. *Physica A: Statistical Mechanics and its Applications*. 2020;540. DOI: 10.1016/j.physa.2019.122721.

[15] Qin YY, et al. Stability analysis and fundamental diagram of heterogeneous traffic flow mixed with cooperative adaptive cruise control vehicles. *Acta Physica Sinica*. 2017;66(9).

[16] Yang Z, et al. A cooperative driving framework for urban arterials in mixed traffic conditions. *Transportation Research Part C: Emerging Technologies*. 2021;124. DOI: 10.1016/j.trc.2020.102918.

[17] Ma F, et al. Eco-driving-based cooperative adaptive cruise control of connected vehicles platoon at signalized intersections. *Transportation Research Part D: Transport and Environment*. 2021;92. DOI: 10.1016/j.trd.2021.102746.

[18] Almannaa MH, et al. Field implementation and testing of an automated eco-cooperative adaptive cruise control system in the vicinity of signalized intersections. *Transportation Research Part D: Transport and Environment*. 2019;67:244-262. DOI: 10.1016/j.trd.2018.11.019.

[19] Zhang J, et al. Energy consumption analysis and prediction of electric vehicles based on real-world driving data. *Applied Energy*. 2020;275. DOI: 10.1016/j.apenergy.2020.115408.

[20] Sun R, et al. Hybrid electric buses fuel consumption prediction based on real-world driving data. *Transportation Research Part D: Transport and Environment*. 2021;91. DOI: 10.1016/j.trd.2020.102637.

[21] Yao Y, et al. Vehicle fuel consumption prediction method based on driving behavior data collected from smartphones. *Journal of Advanced Transportation*. 2020;1-11. DOI: 10.1155/2020/9263605.

[22] Wang Q, et al. Open-pit mine truck fuel consumption pattern and application based on multi-dimensional features and XGBoost. *Sustainable Energy Technologies and Assessments*. 2021;43. DOI: 10.1016/j.seta.2020.100977.

[23] Alvarez Lopez P, et al. Microscopic traffic simulation using SUMO. *2019 IEEE Intelligent Transportation Systems Conference (ITSC), 4-7 Nov. 2018, Maui, HI, USA*. 2018. p. 2575-2582. DOI: 10.1109/ITSC.2018.8569938.

[24] Treiber M, et al. Congested traffic states in empirical observations and microscopic simulations. *Physical Review E*. 2000;62:1805-1824. DOI: 10.1103/PhysRevE.62.1805.

[25] Milanés V, et al. Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data. *Transportation Research Part C: Emerging Technologies*. 2014;48:285-300. DOI: 10.1016/j.trc.2014.09.001.

[26] Krajzewicz D, et al. Second generation of pollutant emission Models for SUMO. *SUMO2014 - Second SUMO User Conference*. 2014. DOI: 10.1007/978-3-319-15024-6_12.

[27] Yao Z, et al. Stability analysis and the fundamental diagram for mixed connected automated and human-driven vehicles. *Physica A: Statistical Mechanics and its Applications*. 2019;533. DOI: 10.1016/j.physa.2019.121931.

[28] Liang JY, et al. A survey on correlation analysis of big data. *Chinese Journal of Computers.* 2016;39(1):1-18.

[29] Chen T, et al. XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016. p. 785-794. DOI: 10.1145/2939672.2939785.

[30] Tien-Anh T. Comparative analysis on the fuel consumption prediction model for bulk carriers from ship launching to current states based on sea trial data and machine learning technique. *Journal of Ocean Engineering and Science*. 2021. DOI: 10.1016/j.joes.2021.02.005.

刘擎超，欧阳文杰，赵晶娅，蔡英凤，陈龙
追尾场景下智能网联车辆燃油消耗特性分析

摘要

智能网联车辆（CAV）是交通拥堵状态下实现节能减排的关键因素。针对追尾事故引起的交通拥堵状态，本文对三种跟驰模型（包括智能驾驶模型（IDM）、自适应巡航控制（ACC）和协同自适应巡航控制（CACC））进行了系统性交通仿真。从车道密度、车辆轨迹和车速的角度，对比分析了三种跟驰模型下车辆的燃油消耗特性。在此基础上，提出了基于XGBoost算法的车辆燃油消耗预测模型。模型结果表明，与IDM和ACC跟驰模型相比，CACC跟驰模型下的车辆总燃油消耗表现最佳；此外，CACC跟驰模型下的交通流更加稳定，在追尾事故不同影响区域内的速度波动相对较小，符合驾驶者的驾驶愿望。

关键词

智能网联车辆（CAV），交通事故、燃油消耗预测，节能