J Intel Connect Veh

Research Article

Long-term trajectory prediction method based on highway vehicle-following behavior patterns

Zhichao An[1,2], Yimin Wu[1], Fan Zhang[3], Dong Zhang[4], Bolin Gao[2,✉], Suying Zhang[5], Guang Zhou[6], Aoning Jia[1,2]

Cite this article: An Z C, Wu Y M, Zhang F, et al. *J Intel Connect Veh* 2025, **8**(1): 9210045. https://doi.org/10.26599/JICV.2024.9210045

ABSTRACT: To address existing shortcomings such as short time domains and low interpretability, this study proposes a long-term trajectory prediction model for leading vehicles that considers the impact of traffic flow. Through an analysis of trailing trajectory data from the HighD natural driving dataset, fitting relationships for the following behavior patterns were derived. Building upon the intelligent driver model (IDM), three long-term trajectory prediction models were established: acceleration delta velocity (ADV), space delta velocity intelligent driver model (SDVIDM), and space velocity intelligent driver model (SVIDM). These models were then compared with the IDM model through simulations. The results indicate that when there is one vehicle ahead, under aggressive following conditions, the ADV model outperforms the IDM model, reducing the root mean square errors in acceleration, speed, and position by 79.61%, 91.26%, and 87.82%, respectively. In scenarios with two vehicles ahead and conservative short-distance following, the SDVIDM model exhibits reductions of 83.42%, 92.85%, and 92.25%, while the SVIDM model shows reductions of 82.31%, 92.47%, and 94.02%, respectively, compared to the IDM model.

KEYWORDS: highways; long-term trajectory prediction; leading vehicle behavior patterns; car-following model

1 Introduction

Vehicle trajectory prediction methods forecast the future trajectories of surrounding vehicles, effectively reducing collision risks and enhancing driving safety. Moreover, in highway scenarios, long-term trajectory prediction can improve traffic flow, increase vehicle traffic efficiency, and prevent the occurrence of "phantom congestion" (Chai, 2020). In the integrated "vehicle-road-cloud" collaboration, roadside perception devices can be leveraged to observe the traffic flow status ahead, providing vehicles with beyond-line-of-sight perception and real-time dynamic information about multiple vehicles ahead. The "beyond-line-of-sight" perception, enabled by the advantage of "rich information" allows for the consideration of the influence of multiple vehicles ahead on the leading vehicle's trajectory. This approach facilitates accurate long-term prediction of the trajectories of leading vehicles, which is highly important for the development of control strategies in intelligent driving systems (Guo et al., 2021).

Currently, there is considerable research on predicting the behavioral states of vehicles. Liu et al. (2022) proposed a driving risk map-integrated deep learning (DRM-DL) method that comprehensively considers vehicle motion uncertainty, trajectory intent uncertainty, and the interaction trajectories between vehicles, lane markings, and road boundaries. This approach is employed for predicting surrounding vehicle behaviors. Gao et al. (2022) considered the uncertainties in vehicle motion and environmental changes by constructing a bidirectional long–short-term memory neural network (BiLSTM), ensuring the accuracy and effectiveness of surrounding vehicle predictions. Zhang et al. (2019) analyzed the motion characteristics during vehicle following conditions and used Bayesian networks to predict the motion speed of leading vehicles. Guo et al. (2020), through Markov chain Monte Carlo simulation, predicted the future states of leading vehicles by deriving a random motion prediction model for leading vehicles in hazardous scenarios. This effectively addressed the challenge of accurately characterizing the random motion of leading vehicles driven by human drivers in testing scenarios. Zhang et al. (2020) proposed a high-speed vehicle motion planning algorithm based on predicting leading vehicle trajectories, considering both driving intent and vehicle motion models. For intelligent driving vehicles, the stochastic nature of leading vehicle motion is substantial. While these studies consider this stochasticity, most of them focus on predicting intent based on the current motion state of the leading vehicle or constructing end-to-end data-driven prediction models using historical trajectory information. Unfortunately, these approaches lack accurate descriptions of the behavioral patterns in leading and following scenarios. Yuan et al. (2015) explored the relationship

[1] School of Mechanical Engineering, Hebei University of Technology, Tianjin 300130, China. [2] School of Vehicle and Mobility, Tsinghua University, Beijing 100008, China. [3] Xiangyang Daan Automobile Testing Center Co., Ltd., Wuhan 441004, China. [4] Department of Mechanical and Aerospace Engineering, Brunel University London, London UB8 3PH, UK. [5] Center of Research and Department, WEICHAI Power Co., Ltd., Weifang 261061, China. [6] Shenzhen Deeproute.ai Co., Ltd., Shenzhen 518000, China.

✉ Corresponding author. E-mail: gaobolin@tsinghua.edu.cn

Received: January 19, 2024; Revised: February 29, 2024; Accepted: May 10, 2024

between the desired following distance and following speed during the stable following stage. They established an expected following distance model and compared it with the safe distance model used in adaptive cruise control (ACC) systems. However, the impact of multiple vehicles ahead on the behavior of the leading vehicle was not considered.

The intelligent driver model (IDM) model, which is distinct from the aforementioned models, is derived from the following behavior patterns of vehicles and exhibits certain regularity and interpretability in its model parameters. Pandita and Caveney (2013) proposed a model-based method for predicting the future states of leading vehicles in a vehicle queue. Building upon the IDM model, they modeled driver following dynamics and employed online parameter estimation to adaptively adjust model parameters based on the characteristics of individual drivers in the queue. Zhan et al. (2023) established a speed prediction algorithm based on the IDM model, considering overall vehicle dynamics and vehicle parking-turning trends. They utilized foresight data from sensors and global positioning system (GPS) to forecast the speed trajectories of vehicles. This suggests that the IDM model is capable of long-term prediction of leading vehicle speeds. However, the model only considers the influence of the immediately preceding vehicle, leading to suboptimal performance in certain scenarios.

In summary, to enhance the long-term prediction accuracy of vehicle trajectories, this study selected and processed the HighD dataset specific to highway scenarios to extract the following trajectory data. Next, an analysis of the behavior patterns of leading vehicles was performed. Subsequently, a trajectory prediction model based on the behavioral patterns derived from the IDM model was constructed and validated through a comparative analysis with the IDM model. The main contributions of this paper can be summarized as follows:

1) Utilizing beyond-visual-range multidimensional dynamic traffic information, namely, the HighD trajectory dataset, a trajectory prediction model based on the behavior patterns of car-following in the presence of upstream traffic flow is constructed. This model integrates traffic flow characteristics, vehicle behavior patterns, and the car-following model.

2) By employing polynomial fitting, the method quantitatively analyzes the relative changes between the following distance and following speed, following distance and relative speed, and acceleration and relative speed. This analysis, combined with the existing intelligent driver model (IDM) for modeling and prediction, enhances the precision of acceleration, speed, and longitudinal position predictions for the leading vehicle in both aggressive and conservative following scenarios. This approach significantly enhances the breadth and depth of road traffic information utilization, improving the rationality of matching long-term trajectory predictions of leading vehicles with traffic flow characteristics.

This paper is organized as follows: In Section 2, the HighD natural driving dataset is processed, and trajectory data for one or two vehicles ahead are extracted to prepare for subsequent modeling and analysis. Section 3 analyzes the distribution and variation patterns of the following parameters and establishes relative relationships for the following distance-following vehicle speed, following distance-relative speed, and acceleration-relative speed. In Section 4, a trajectory prediction model based on behavior patterns is developed, and a genetic algorithm is used for parameter calibration. Section 5 involves simulation verification to validate the effectiveness of the established model. The concluding remarks are presented in Section 6.

2 Data source and extraction of the following data

2.1 HighD natural driving dataset

The HighD dataset originates from research conducted by scholars at Rheinisch-Westfälische Technische Hochschule (RWTH) Aachen University in Germany (Krajewski et al., 2018). This dataset was gathered using unmanned aerial vehicles (drones) to capture video data of vehicle movements at six locations near highways around Cologne, Germany. The locations where the data were collected are illustrated in Fig. 1. Fig. 2 provides visual representations of the unmanned aerial vehicles and the actual highway scenes used during the data collection process.

After obtaining the video data, vehicle trajectory data were extracted using neural network methods, and Bayesian smoothing techniques were applied to smooth the vehicle trajectories. The frequency of vehicle trajectory data in the HighD dataset is 25 Hz, with a total of 110,500 vehicle driving data points. The cumulative driving time is 147 h, covering a total distance of 44,500 km. The trajectory data are stored in 60 comma separated values (CSV) files.

2.2 Extraction of following vehicle data

Due to the greater number of recorded vehicles at location 1, we use the 14th CSV data file as an example to analyze the characteristics of the following behavior. This file has a sampling time on Thursday at 6:28 p.m., recording data from 2,844 vehicles across 6 lanes, including 2,580 cars and 264 trucks. We focus on the trajectory data of cars in lanes 6, 7, and 8, which constitute 3 lanes in the forward direction. A data extraction method is proposed. First, data on lane-changing vehicles are removed from the trajectory file. The moments where the "preceding" is 0 are identified, corresponding to the first vehicle with the respective "ID" at that time. Based on this "ID", the data for the following two vehicles are retrieved, forming a scenario in which three vehicles follow each other. The data of the first two vehicles

Fig. 1 Six data collection points of the HighD dataset.

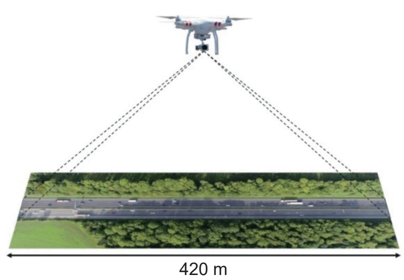


Fig. 2 Data collection scenes of the HighD dataset.

constitute a scenario where there is only one preceding vehicle in front, and the data of the last two vehicles constitute a scenario where there are two preceding vehicles in front. This approach facilitates subsequent analysis and modeling studies. The illustrations of both scenarios are shown in Fig. 3, where the dashed line represents the ego vehicle, the red and blue lines represent the preceding vehicles, and the black line represents the vehicle in front of the preceding vehicle.

3 Analysis of highway vehicle-following characteristics

3.1 Distribution and variation patterns of the following parameters

After obtaining the following trajectory data for the two scenarios, the analysis focused on key parameters such as the driver's following speed, relative speed, acceleration, and following distance.

Fig. 4 shows that the following vehicle speeds in both scenarios follow a normal distribution, with average speeds reaching 32 m/s. The speeds are mostly distributed within the range of [25 m/s, 40 m/s], and some following vehicle speeds exceed 40 m/s. The mean following vehicle speed when there are two preceding vehicles is slightly lower than the mean following vehicle speed when there is only one preceding vehicle. This indicates that when the number of preceding vehicles increases, following vehicles tend to reduce their speeds to ensure driving safety.

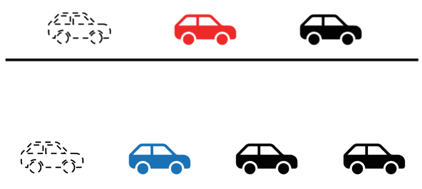


Fig. 3 Illustrations of the two scenarios.

The definition of relative speed is given by $\mathrm{d}v = v - v_{\mathrm{p}}$, where v is the speed of the ego vehicle (following vehicle) and v_{p} is the speed of the preceding vehicle. As shown in Fig. 5, when the driver is following, the relative speed follows a normal distribution. When there is one preceding vehicle, the relative speed is concentrated within the range of [−5 m/s, 5 m/s]. In contrast, when there are two preceding vehicles, the distribution of relative speed is even more concentrated, with most values falling within [−2 m/s, 2 m/s]. Some individual speed differences reach 8 m/s. The mean relative speed is negative in both scenarios, and the mean relative speed is greater when there is only one preceding vehicle than when there are two preceding vehicles. This indicates that drivers tend to maintain a smaller speed difference with the preceding vehicle when following, and this tendency is more pronounced when there are two preceding vehicles.

Fig. 6 shows that the acceleration follows a normal distribution with a mean greater than 0, concentrated within the interval [$-1\ \mathrm{m/s^2}$, $1\ \mathrm{m/s^2}$]. The distribution of acceleration is greater than that of deceleration, indicating that during the process of following other vehicles, drivers tend to exhibit a greater inclination toward acceleration than deceleration. When there is only one preceding vehicle, the tendency for acceleration is greater than when there are two preceding vehicles.

The following distance, defined as $S = X_{\mathrm{p}} - X - L$, where X_{p} is the preceding vehicle's position, X is the self-vehicle's position, and L is the length of the preceding vehicle, follows a log-normal distribution. As observed in Fig. 7, the average following distance is 67.16 m when there is one preceding vehicle and 57.78 m when there are two preceding vehicles. One of the reasons for this difference is that traffic safety and driver expectations are the main factors determining subsequent behavior. To avoid collisions with the preceding vehicle, the following vehicle should maintain a sufficient distance. Driver expectations, including speed, following distance, and environmental judgment, encourage drivers to

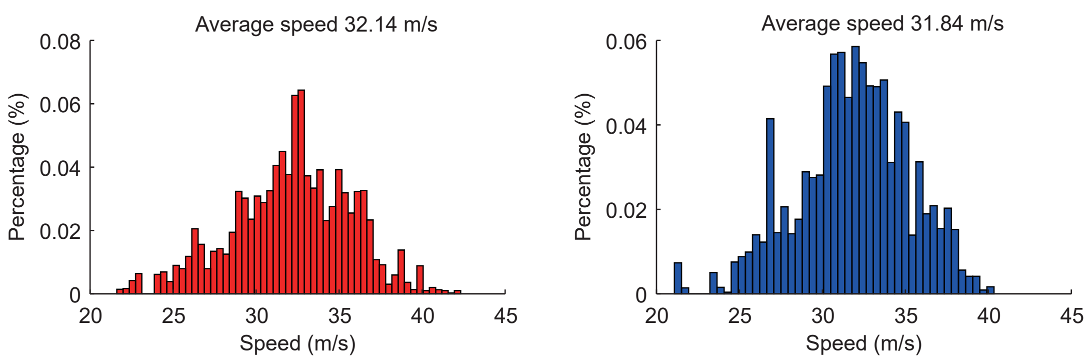


Fig. 4 Comparison of the following vehicle speed distributions.

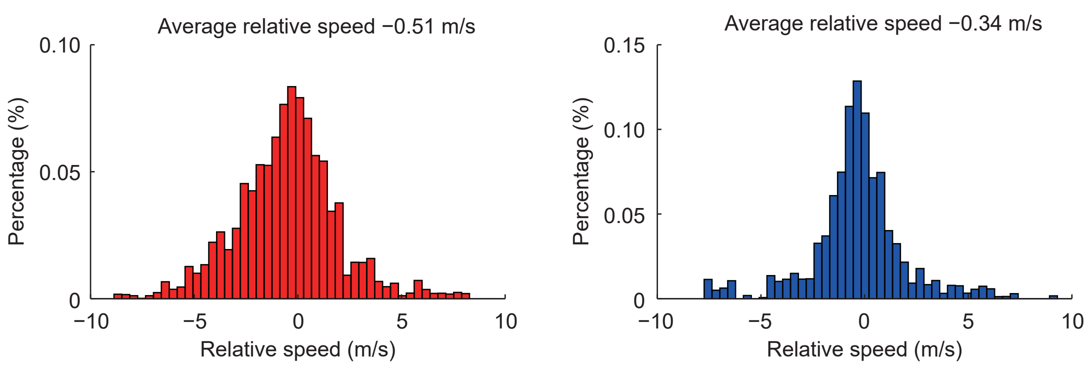


Fig. 5 Comparison of relative speed distributions.

https://doi.org/10.26599/JICV.2024.9210045

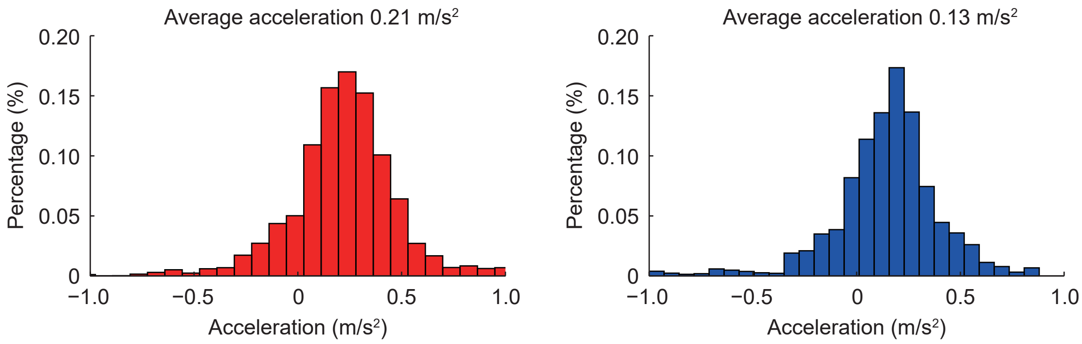


Fig. 6 Comparison of acceleration distributions.

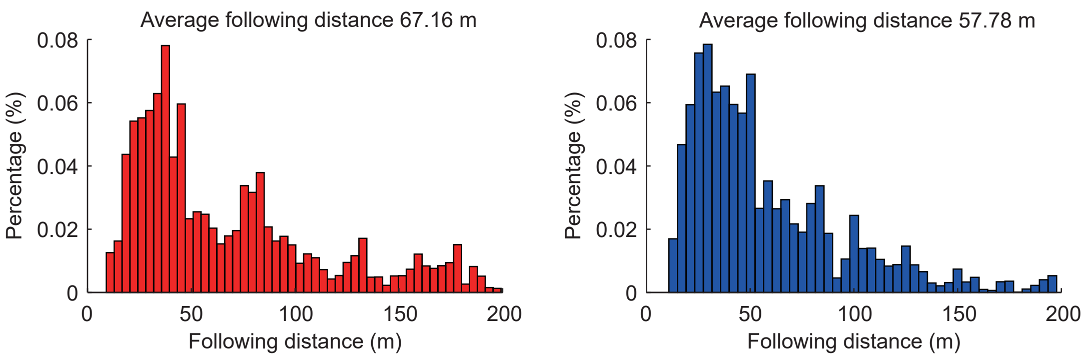


Fig. 7 Comparison of the following distance distributions.

maintain a smaller distance from the preceding vehicle. Another reason is that when there are two preceding vehicles, the traffic flow is greater. Without decreasing the following distance, neighboring vehicles are likely to enter, which could further impact the driver's traffic efficiency. Therefore, when there are two preceding vehicles, drivers tend to maintain a smaller following distance, even at the cost of sacrificing safety (Brackstone et al., 2002; Liu et al., 2017).

By analyzing the distribution patterns of vehicle parameters such as acceleration, time headway, and following distance, these data not only reflect the behavioral characteristics of drivers but also embody their decision-making objectives. The fundamental reason behind this lies in the decision-making mechanism of drivers, characterized by the principle of "seeking benefits and avoiding harm" (Wang et al., 2020). The basic strategy and principles involve achieving greater maneuverability and safety while avoiding collisions and violations.

3.2 Relative change relationships of the following parameters

A single metric cannot comprehensively describe the behavioral characteristics of the following vehicles. In traffic flow modeling, the following speed, relative speed, acceleration, and following distance are all crucial parameters. As indicated in Sun et al. (2020) and Zhu et al. (2020), the following distance is most positively correlated with the following speed, while there is a negative correlation between the relative speed and acceleration. The relationship between them can be fitted with polynomials. This study focuses on analyzing the relationships between the following distance-following speed, following distance-relative speed, and acceleration-relative speed. The following speed and relative speed are divided into 50 intervals. Corresponding data for the following distance and following car acceleration are obtained for each interval, and the data are averaged. The relative change relationships for each parameter are illustrated in Figs. 8–13.

As depicted in Fig. 8, with an increase in the following speed, the following distance exhibits an upward trend. This indicates that an increase in the following speed leads to an elevated collision risk. Due to the decision-making mechanism of drivers, who prioritize safety by "avoiding harm and seeking benefit", drivers tend to increase the following distance to ensure an adequate safety margin. The growth rate of the following distance with increasing following speed is slightly greater when there is

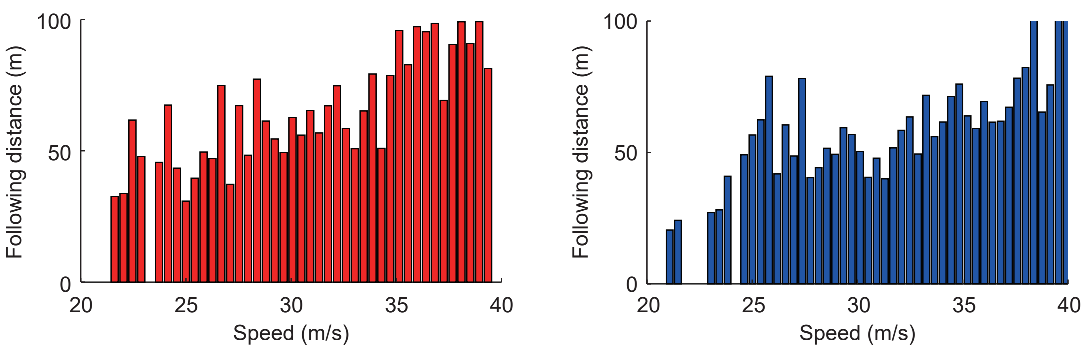


Fig. 8 Relative relationship between following distance and following speed.

only one preceding vehicle compared to the case with two preceding vehicles. This is because the following distance when there is only one preceding vehicle is greater than that in the scenario with two preceding vehicles. To prevent vehicles from adjacent lanes from merging in and to reduce the impact on the efficiency of self-driving, drivers are more inclined to control the gradual increase in the following distance while increasing the following speed when there are two preceding vehicles. This ensures efficient driving.

As shown in Fig. 9, the relationship between the following distance and relative speed tends to decrease in the middle and increase on both sides. When the relative speed is less than 0, which indicates that the following vehicle's speed is lower than the preceding vehicle's speed, the collision risk is relatively low. As the relative speed continuously increases, the following distance also increases, causing the vehicles to gradually move apart. When the relative speed is greater than 0, indicating that the following vehicle's speed is higher than the preceding vehicle's speed, the collision risk increases with increasing relative speed. In this scenario, the following vehicle gradually approaches the preceding vehicle. According to the statistical results of the following distances under the two conditions, the average following distance when there are two preceding vehicles is smaller than the average following distance when there is only one preceding vehicle. Therefore, as the relative speed increases, drivers with two preceding vehicles tend to gradually increase the following distance to ensure both efficient driving and safe driving.

As depicted in Fig. 10, the relationship between the acceleration and relative speed is approximately symmetric about the origin $(0, 0)$. When the relative speed is less than 0, indicating that the following vehicle's speed is lower than the preceding vehicle's speed, the acceleration of the following vehicle shows an initial increase followed by a decrease, with the peak occurring in the interval $(-4, -2)$. This suggests that in this interval, drivers are more sensitive to differences in speeds between the preceding and following vehicles, resulting in a greater inclination toward acceleration. As the speed of the following vehicle gradually approaches that of the preceding vehicle, the inclination toward acceleration decreases, reaching its lowest point in the interval $(0, 2)$. Notably, this lowest point does not occur at a relative speed of 0, indicating that drivers do not start decelerating immediately when the speeds of the preceding and following vehicles are the same; instead, there is a certain degree of hysteresis. When the relative speed is greater than 0, indicating that the following vehicle's speed is higher than the preceding vehicle's speed, the collision risk increases. When there is only one preceding vehicle, the distribution of acceleration is greater than the distribution of deceleration. However, when there are two preceding vehicles, the situation is reversed.

4 Trajectory prediction model based on preceding vehicle behavioral patterns

From the analysis of the highway vehicle following the abovementioned characteristics, it can be observed that there are certain correlation patterns among follow-spacing to follow-vehicle speed, follow-spacing to relative speed, and acceleration to relative speed. In this study, a polynomial fitting method is employed to represent their relationships, and the fitting results are illustrated in the following figures.

Based on Fig. 11, the relationship between the following distance and relative speed in both scenarios can be well fitted with a quadratic polynomial. Additionally, as shown in Figs. 12 and 13, the relationships between the following distance and following speed, as well as acceleration and relative speed, can be effectively modeled using cubic polynomials. The corresponding mathematical expressions are provided by Eqs. (1)–(6), and the coefficients for the fits are presented in Table 1.

1) In the scenario where there is one preceding vehicle in front of the ego vehicle:

$$S1 = A_1 \mathrm{d}v^2 - B_1 \mathrm{d}v + C_1 \tag{1}$$

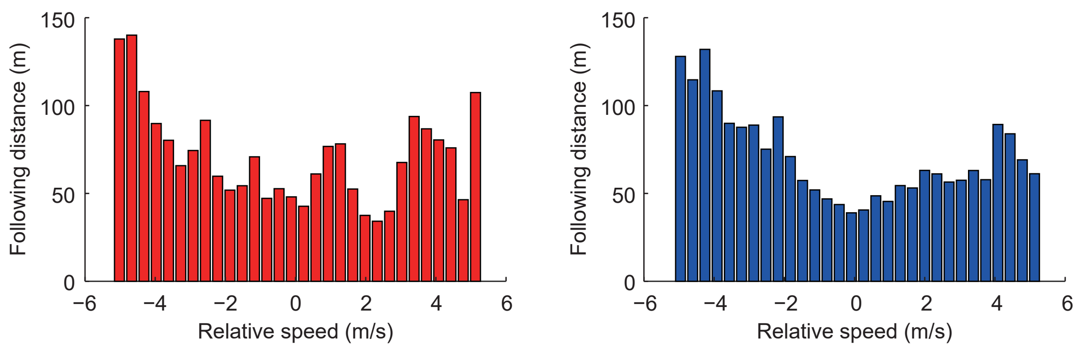


Fig. 9 Relative relationship between the following distance and relative speed.

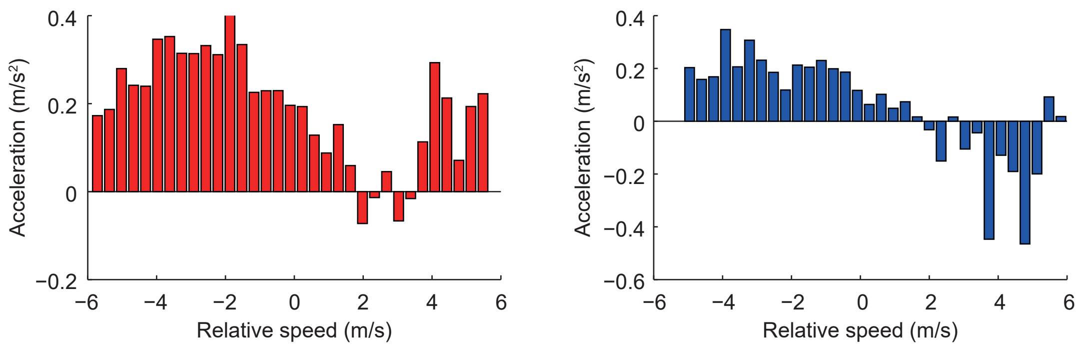


Fig. 10 Relative relationship between the acceleration and relative speed.

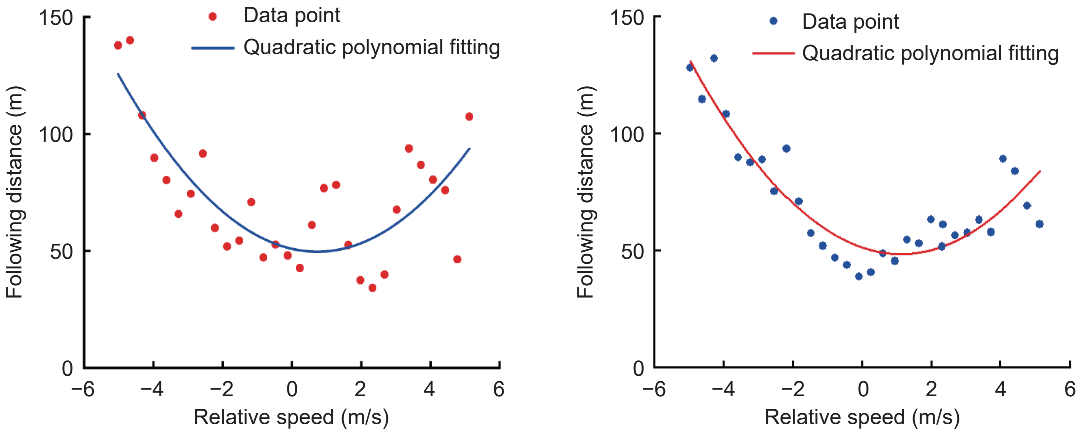


Fig. 11 Fitted relationship of the following distance and relative speed.

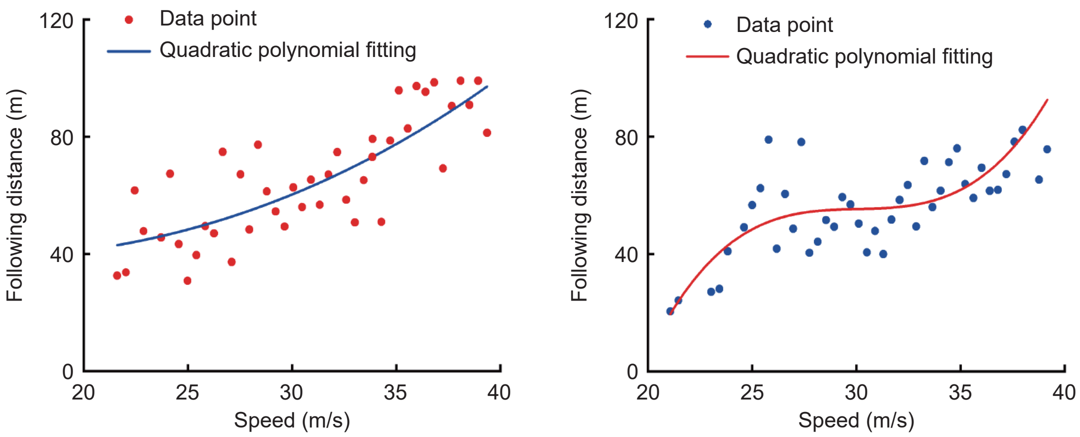


Fig. 12 Fitted relationship of the following distance and following speed.

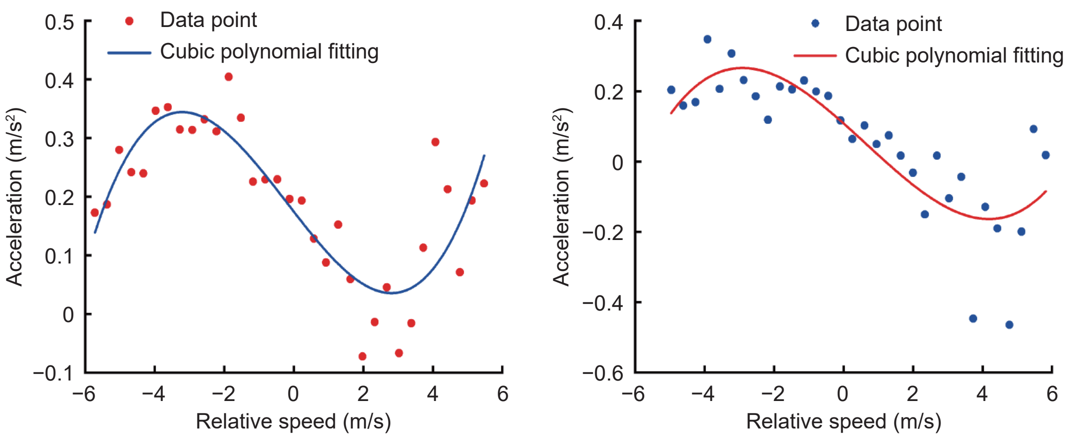


Fig. 13 Fitted relationship of the acceleration and relative speed.

$$S2 = A_2 \mathrm{d}v^3 - B_2 v^2 + C_2 v - D_1 \tag{2}$$

$$A1 = A_3 \mathrm{d}v^3 - B_3 \mathrm{d}v^2 - C_3 \mathrm{d}v + D_2 \tag{3}$$

2) In the scenario where there are two preceding vehicles in front of the ego vehicle:

$$S3 = A_4 \mathrm{d}v^2 - B_4 \mathrm{d}v + C_4 \tag{4}$$

$$S4 = A_5 \mathrm{d}v^3 - B_5 v^2 + C_5 v - D_3 \tag{5}$$

$$A5 = A_6 \mathrm{d}v^3 - B_6 \mathrm{d}v^2 - C_6 \mathrm{d}v + D_4 \tag{6}$$

The IDM model, due to its high accuracy and wide applicability in various scenarios, is widely used in microscopic traffic flow simulations. The IDM model was proposed by Treiber et al (2000).

$$a = a_0 \left[1 - \left(\frac{v}{v_0}\right)^4 - \frac{\tilde{S}}{S}\right] \tag{7}$$

$$\tilde{S} = s_0 + vT + \frac{v\mathrm{d}v}{2\sqrt{a_0 b}} \tag{8}$$

where a represents the acceleration of the ego vehicle, v is the speed of the ego vehicle, $\mathrm{d}v$ is the speed difference between the ego vehicle and the preceding vehicle, S is the gap to the

Table 1 Fitting relationships for both cases

Single preceding	Vehicle scenario	Two precedings	Vehicle scenario
A_1	2.2890	A_4	2.2272
A_2	0.0328	A_5	0.0788
A_3	0.0028	A_6	0.0024
B_1	3.3916	B_4	5.0347
B_2	2.8795	B_5	7.0115
B_3	0.0017	B_6	0.0046
C_1	50.9626	C_4	51.3215
C_2	86.3722	C_5	206.8337
C_3	0.0765	C_6	0.0878
D_1	824.6381	D_3	1,968.4845
D_2	0.1750	D_4	0.1086

preceding vehicle, a_0 is the maximum acceleration of the ego vehicle, v_0 is the maximum speed in the traffic flow, s_0 is the minimum safety gap, T is the safe time headway, and b is the absolute value of the comfortable deceleration.

$$v(t) = v(t-1) + a(t)\,\mathrm{d}t \tag{9}$$

$$y(t) = v(t)\,\mathrm{d}t + \frac{1}{2}a(t)\,\mathrm{d}t^2 \tag{10}$$

Under the known initial conditions and the preceding vehicle's speed, the self-vehicle acceleration can be calculated using Eq. (7). By utilizing Eqs. (9) and (10), the model predicts the speed and the predicted position. Based on the composition of the IDM model, Eqs. (3), (9), and (10) can be combined to form a new ADV model represented by Eq. (11). By replacing Eqs. (1) and (2) in the IDM model with the gaps, the improved SDVIDM model is given by Eq. (14), and the SVIDM model is expressed as Eq. (17). The following Eqs. (11)–(19) present the improved model for the scenario where there is only one preceding vehicle:

The ADV model:

$$A1 = A_3\mathrm{d}v^3 - B_3\mathrm{d}v^2 - C_3\mathrm{d}v + D_2 \tag{11}$$

$$v(t) = v(t-1) + a(t)\,\mathrm{d}t \tag{12}$$

$$y(t) = v(t)\,\mathrm{d}t + \frac{1}{2}a(t)\,\mathrm{d}t^2 \tag{13}$$

The SDVIDM model:

$$a = a_0\left[1 - \left(\frac{v}{v_0}\right)^4 - \frac{\tilde{S}}{S1}\right] \tag{14}$$

$$v(t) = v(t-1) + a(t)\,\mathrm{d}t \tag{15}$$

$$y(t) = v(t)\,\mathrm{d}t + \frac{1}{2}a(t)\,\mathrm{d}t^2 \tag{16}$$

The SVIDM model:

$$a = a_0\left[1 - \left(\frac{v}{v_0}\right)^4 - \frac{\tilde{S}}{S2}\right] \tag{17}$$

$$v(t) = v(t-1) + a(t)\,\mathrm{d}t \tag{18}$$

$$y(t) = v(t)\,\mathrm{d}t + \frac{1}{2}a(t)\,\mathrm{d}t^2 \tag{19}$$

After establishing the trajectory prediction model, it is necessary to calibrate the model parameters. In this study, the MATLAB Genetic Algorithm Toolbox was utilized for optimization. Due to the stochastic nature of the algorithm, to find the optimal solution, the calibration process was repeated 10 times, and the average value was taken as the final result. The calibration results for the two scenarios are presented in Tables 2 and 3.

5 Simulation validation

To validate the long-term prediction performance of the established trajectory prediction model, this study selected the root mean square error (RMSE) as the evaluation metric (E_{RMSE}). Below are the results of the experimental analysis, with the evaluation metric defined as in Eq. (20):

$$E_{\mathrm{RMSE}} = \sqrt{\frac{\sum_{i=1}^{N}\left(y_i^{\mathrm{tru}} - y_i^{\mathrm{pre}}\right)^2}{N}} \tag{20}$$

where N represents the number of samples and y_i^{pre} denote the true and predicted values.

In addition to comparing with the IDM model, considering that the condition where there are multiple vehicles in front of the preceding vehicle is very common in actual car-following scenarios, this paper includes a comparison with an LSTM neural network in Section 4.2. The network structure is a single-layer LSTM network, and the prediction method is single-step iterative prediction.

5.1 One vehicle in front of the ego vehicle

1) Aggressive car-following behavior

Fig. 14 shows the results of Simulation 1 for a prediction horizon of 10 s. When the vehicle exceeds the maximum traffic flow speed, the IDM, SDVIDM, and SVIDM models predict that the vehicle will decelerate first and then accelerate, eventually reaching the maximum speed. However, some aggressive drivers, in pursuit of higher traffic efficiency, may choose to continuously accelerate rather than adopt this driving strategy. The ADV model can effectively predict this driving behavior. As shown in Table 4, the proposed model outperforms the IDM model in predicting acceleration, speed, and longitudinal position, with reductions of 79.61%, 91.26%, and 87.82%, respectively.

2) Conservative car-following behavior

The results for Simulation 2, with a 5-s prediction horizon, are

Table 2 Parameter calibration results for the condition of one preceding vehicle

Parameter	IDM	SDVIDM	SVIDM
a (m/s²)	1.5273	1.4358	1.0978
b (m/s²)	3.1172	3.2295	3.5511
v_0 (m/s)	33.5781	33.0136	30.4095
T (s)	0.7992	1.2282	1.0203
s_0 (m)	5.3008	5.3528	5.1585

Table 3 Parameter calibration results for the condition of two preceding vehicles

Parameter	IDM	SDVIDM	SVIDM
a (m/s²)	0.7791	1.0360	1.0978
b (m/s²)	3.4131	3.6839	3.5511
v_0 (m/s)	30.3307	31.2342	30.4095
T (s)	1.2965	1.1557	1.0203
s_0 (m)	4.7935	5.3203	5.1585

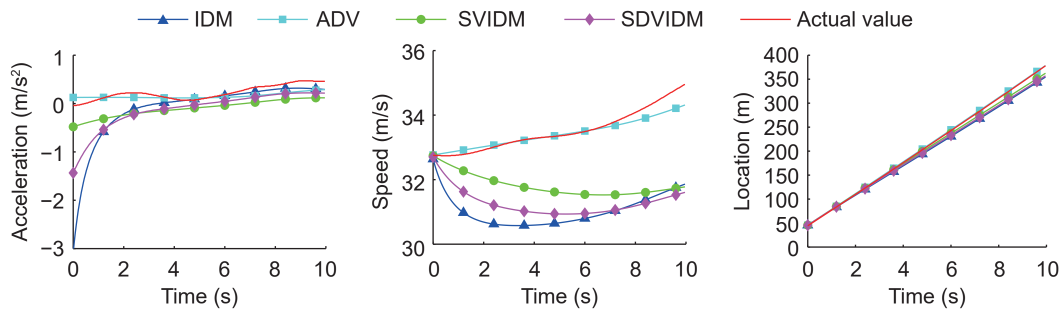


Fig. 14 Prediction results for Simulation 1.

Table 4 Comparison of the prediction errors of each model in Simulation 1

RMSE	IDM	ADV	SDVIDM	SVIDM
a (m/s^2)	0.5990	0.1221	0.4330	0.3321
v (m/s)	2.5424	0.2222	2.3673	1.8943
y (m)	12.2276	1.4893	10.2476	6.9071

depicted in Fig. 15. In this scenario, the initial speed of the subject vehicle is 21.5 m/s, which is considerably lower than the maximum vehicle speed. Under these conditions, the IDM, SDVIDM, and SVIDM predict relatively high vehicle accelerations. However, some more conservative drivers tend to gradually increase their speeds, allowing the vehicle to reach the maximum speed during the following. The prediction results revealed that the ADV model exhibited the smallest prediction error. According to Table 5, the errors of the ADV model are 41.80% lower than those of the IDM model, 68.60% lower than those of the SDVIDM model, and 77.45% lower than those of the SVIDM model.

5.2 Two vehicles in front of the ego vehicle

1) Aggressive car-following behavior

Fig. 16 presents the 10-s prediction results of Simulation 3. Under this condition, the vehicle travels faster than the maximum speed of the traffic flow, and the acceleration predictions of the IDM, SDVIDM, and SVIDM models are all negative, indicating that the vehicle slows down and eventually stabilizes near the maximum speed for car-following. To ensure driving safety, vehicles must decelerate to stabilize at a speed close to the maximum for subsequent operation. However, for aggressive drivers, when the speed exceeds the maximum, they still choose to accelerate to maintain a high-speed state to pass through the road

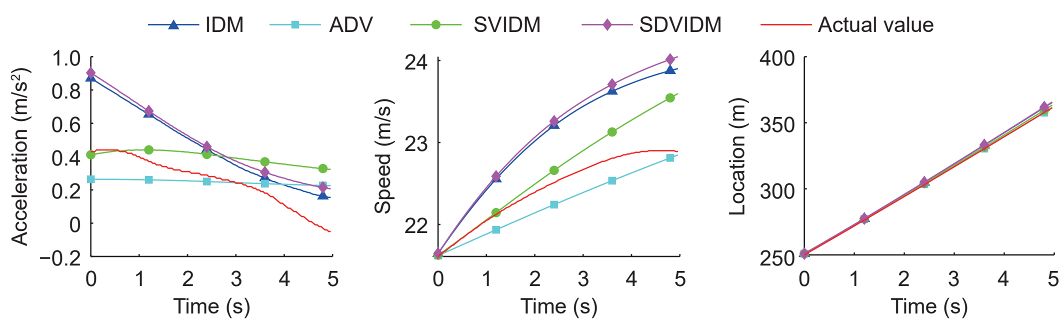


Fig. 15 Prediction results for Simulation 2.

Table 5 Comparison of the prediction errors of each model in Simulation 2

RMSE	IDM	ADV	SDVIDM	SVIDM
a (m/s^2)	0.2249	0.1309	0.2509	0.1787
v (m/s)	0.6810	0.2138	0.7581	0.2932
y (m)	2.2343	0.5038	2.3643	1.1049

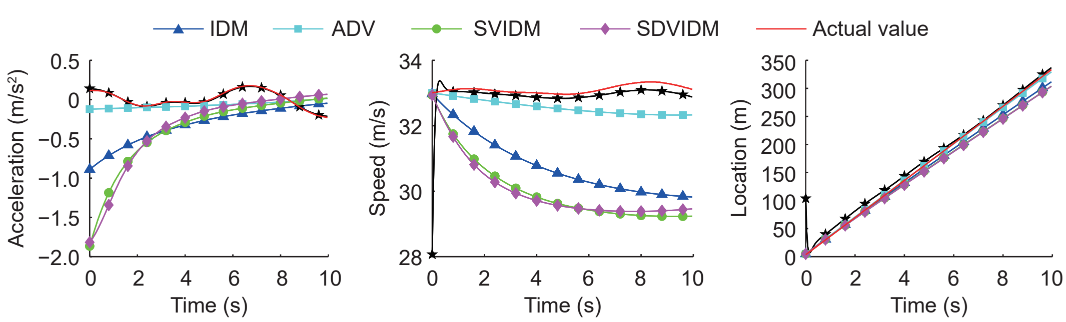


Fig. 16 Prediction results for Simulation 3.

section as quickly as possible. Due to the LSTM's use of single-step iterative prediction, the initial frames have larger prediction errors in speed and longitudinal position, but the acceleration prediction is relatively accurate. The ADV model significantly outperforms the other models in long-term prediction under these conditions. According to Table 6, the ADV model reduces the error by 73.33%, 95.85%, and 86.65% compared to the IDM model, respectively, and reduces the speed error by 75.04% and longitudinal position error by 85.99% compared to the LSTM model.

2) Conservative close car-following behavior

Fig. 17 shows the 10-s prediction results of Simulation 4. In this scenario, the vehicle starts at a speed of 24.5 m/s and a following distance of 23 m, representing a conservative close-following mode. The IDM model assumes that the following distance is close to the desired following distance, and in the initial few seconds, the vehicle will brake to maintain a comfortable following distance; hence, the vehicle will decelerate and then accelerate. However, the remaining models outperform the IDM model because the following distance in these models is calculated based on vehicle following patterns, making the calculated following distance more reasonable and the predicted acceleration more accurate and in line with drivers' following habits. Overall, the SDVIDM and SVIDM models have the highest prediction accuracy. The prediction errors are shown in Table 7. According to Table 7, the SDVIDM model reduces the acceleration error by 83.42%, speed error by 92.85%, and longitudinal position error by 92.25% compared to IDM, respectively, and reduces the speed error by 45.93% and longitudinal position error by 92.98% compared to LSTM; the SVIDM model reduces the acceleration error by 82.31%, speed error by 92.47%, and longitudinal position error by 94.02% compared to IDM, respectively, and reduces the speed error by 43.12% and longitudinal position error by 94.60% compared to LSTM.

3) Conservative long-distance car-following behavior

Fig. 18 displays the 10-s prediction results of Simulation 5. In this scenario, the vehicle starts at a speed of 23.5 m/s and an initial following distance of 58 m, which is indicative of a conservative long-distance following mode. The prediction results show that both the IDM and the improved models predict larger accelerations in the first few seconds, leading to a greater tendency for the vehicle to accelerate. This indicates the driver's inclination to speed up to catch up with the preceding vehicle and reduce the following distance. As the prediction horizon increases, the acceleration gradually approaches the true value, and the rate of increase in the speed slows. Overall, the ADV model exhibits relatively high prediction accuracy. According to Table 8, the ADV model reduces the errors by 9.90%, 11.06%, and 14.77% compared to the IDM model, respectively, and reduces the longitudinal position error by 49.75% compared to the LSTM model.

Table 6 Comparison of the prediction errors of each model in Simulation 3

RMSE	IDM	ADV	SDVIDM	SVIDM	LSTM
a (m/s^2)	0.4214	0.1124	0.6151	0.5254	0.0099
v (m/s)	2.5091	0.1042	3.2502	2.9953	0.4174
y (m)	10.3357	1.3797	15.0057	13.1077	9.8529

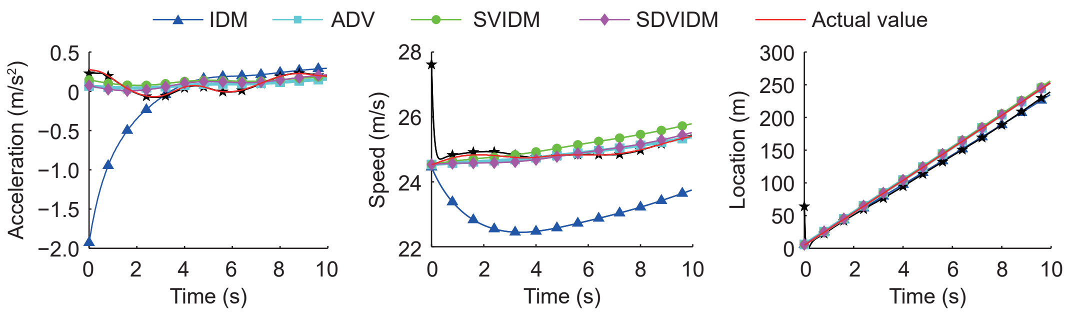


Fig. 17 Prediction results for Simulation 4.

Table 7 Comparison of the prediction errors of each model in Simulation 4

RMSE	IDM	ADV	SDVIDM	SVIDM	LSTM
a (m/s^2)	0.5490	0.1044	0.0910	0.0971	0.0158
v (m/s)	1.9400	0.3696	0.1388	0.1460	0.2567
y (m)	10.1417	2.0426	0.7861	0.6063	11.2229

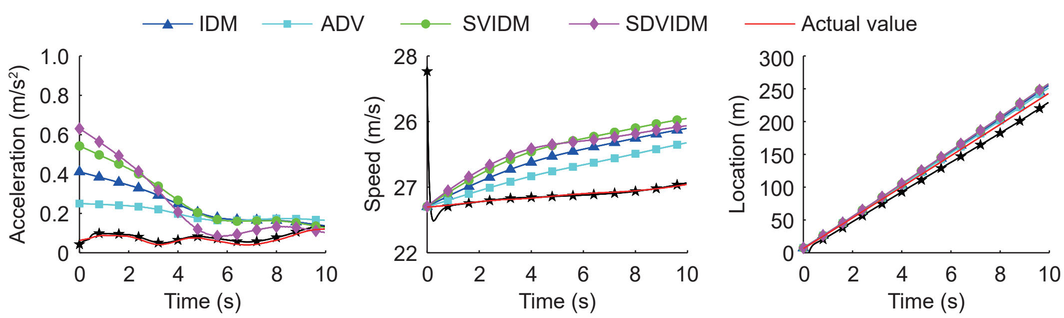


Fig. 18 Prediction results for Simulation 5.

清華大学出版社 Tsinghua University Press | Available on IEEE Xplore

Table 8 Comparison of the prediction errors of each model in Simulation 5

RMSE	IDM	ADV	SDVIDM	SVIDM	LSTM
a (m/s^2)	0.1930	0.1739	0.2460	0.2191	0.0182
v (m/s)	1.2157	1.0812	1.4166	1.3320	0.3013
y (m)	6.2711	5.3405	7.7426	7.0710	10.6286

6 Conclusions

The long-term trajectory prediction method proposed in this study, which is based on the regularities of vehicle following behavior on highways, considers the influence of upstream traffic flow characteristics on the leading vehicle's following behavior. This study summarizes the behavioral patterns exhibited by the leading vehicle during the following process and incorporates these insights into an improved long-term domain trajectory model using the IDM. Compared to the baseline model, this model reduces the RMSE in both aggressive and conservative following scenarios while extending the prediction horizon, thereby enhancing the model's interpretability. For future work, we plan to consider the influence of a greater number of surrounding vehicles on the leading vehicle's behavior and to integrate the proposed model with deep learning models to further refine the accuracy and optimize the long-term trajectory prediction model.

Replication and data sharing

The data and codes that support the findings of this study are available at https://doi.org/10.26599/ETSD.2025.9190033. The MATLAB program code within this research can be made accessible upon request via email to the corresponding author.

Acknowledgements

We gratefully acknowledge the financial support provided by the Hetao Shenzhen–Hong Kong Science and Technology Innovation Cooperation Zone (HZQB-KCZYZ-2021055) and the support from the National Natural Science Foundation of China (52172389).

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

References

Brackstone, M., Sultan, B., McDonald, M., 2002. Motorway driver behaviour: Studies on car following. *Transp Res Part F Traffic Psychol Behav*, **5**, 31–46.

Chai, J., 2020. On vehicle trajectory prediction model based on combined clustering and LSTM algorithm. Ph.D. Dissertation. Beijing: Beijing Jiaotong University. (in Chinese)

Gao, Y., Zhu, J., Gao, H., 2022. The surrounding vehicles behavior prediction for intelligent vehicles based on Att-BiLSTM. *JUSTC*, **52**, 6.

Guo, J., Li, W., Luo, Y., Chen, T., Li, K., 2021. Driver car-following model based on deep reinforcement learning. *Autom Eng*, **43**, 571–579.

Guo, J., Li, K., Wang, J., Chen, T., Li, W., WANG, B., 2020. Prediction of preceding car motion under car-following scenario in the internet of vehicles based on bayesian network. *Autom Eng*, **42**, 847–853.

Krajewski, R., Bock, J., Kloeker, L., Eckstein, L., 2018. The highD dataset: A drone dataset of naturalistic vehicle trajectories on German highways for validation of highly automated driving systems. In: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2118–2125.

Liu, X., Wang, Y., Jiang, K., Zhou, Z., Nam, K., Yin, C., 2022. Interactive trajectory prediction using a driving risk map-integrated deep learning method for surrounding vehicles on highways. *IEEE Trans Intell Transport Syst*, **23**, 19076–19087.

Liu, X., Shi, J., Chen, Y., 2017. *Traffic Behavior*. Beijing: Science Press. (in Chinese)

Pandita, R., Caveney, D., 2013. Preceding vehicle state prediction. In: 2013 IEEE Intelligent Vehicles Symposium (IV), 1000–1006.

Sun, H., Liu, X., Huang, K., 2020. Research on driver's follow-up characteristics based on natural high-speed driving data. *Journal of Shandong Jiaotong University*, **28**, 23–29. (in Chinese)

Treiber, M., Hennecke, A., Helbing, D., 2000. Congested traffic states in empirical observations and microscopic simulations. *Phys Rev E*, **62**, 1805–1824.

Wang, J., Zheng, X., Huang, H., 2020. Decision-making mechanism of the drivers following the principle of least action. *China J of Highw Transp*, **33**, 155.

Yuan, W., Fu, R., Ma, Y., Guo, Y., Du, C., 2015. A study on driver's vehicle-following model based on high speed real driving data. *Automot Eng*, **37**, 679–685. (in Chinese)

Zhan, W., Sun, Z., Fu, J., Sun, F., 2023. Research on long-term vehicle following speed prediction based on IDM algorithm. *Automob Techn*, **9**, 27–34.

Zhang, J., Li, K., Luo, Y., Zhang, S., Li, H., 2019. Prediction of preceding car motion under car-following scenario in the internet of vehicles based on bayesian network. *Autom Eng*, **41**, 245–251. (in Chinese)

Zhang, Y., Zhou, B., Wu, X., Cui, Q., Chai, T., 2020. Motion planning of high speed intelligent vehicle based on front vehicle trajectory prediction. *Autom Eng*, **42**, 574–580.

Zhu, K., Shen, J., Wang, X., Zhou, Q., 2020. Research on the Following Charqcteristics of Expressway Vehicles Based on the Measured date and mathematical statistic. *Communications Science and Technology Heilongjiang*, **43**, 172–175. (in Chinese)

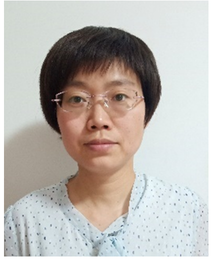
Zhichao An graduated from Shenyang University of Technology in 2021 with the B.S. degree in engineering. He is currently pursuing the M.S. degree in vehicle engineering at Hebei University of Technology. He is engaged in research work in School of Vehicle and Mobility, Tsinghua University and his research field is vehicle trajectory prediction.

Yimin Wu is a professor with the Department of Vehicle Engineering, School of Mechanical Engineering, Hebei University of Technology, Director of the Society of Automotive Engineers of China, and Vice Chairman of the Society of Automotive Engineers of Hebei Province. He has long been engaged in the research of automotive system dynamics design and automotive noise and vibration (NVH) control. In the past 5 years, he has presided over or participated in more than 10 research projects, including 1 national project of the Ministry of Science and Technology, 4 projects of the National Foundation Committee and provincial and ministerial level, and more than 10 projects of enterprises and institutions.

Fan Zhang graduated from Dublin City University with the M.S. degree in software engineering and communication in 2011. He has successively served as the director of the Central Office of the Business Planning Department of the Technology Center of Dongfeng Motor Corporation, the Director of the Intelligent Connected Testing Department of Xiangyang Da'an Automobile Inspection Center, and is currently the deputy general manager of Xiangyang Daan Automobile Testing Center Co., Ltd. (presiding over the work), with research interests in the development and testing of intelligent networked vehicles.

Dong Zhang is currently a lecturer in the Department of Mechanical and Aerospace Engineering, Brunel University London, UK. He joined Brunel in 2021 after having spent one year at the Nanyang Technological University, Singapore as research fellow in School of Mechanical and Aerospace Engineering, and the School of Electrical and Electronic Engineering. He received the M.Sc. degree in 2015 from Jilin University, Changchun, China and the Ph.D. degree in 2019 from University of Lincoln, Lincoln, UK. He joined Brunel University London and founded the Intelligent Driving and Transportation Research Group in September 2021. He has published more than 30 articles in international journals as well as numerous conference articles, and about 10 granted patents mostly in the area of intelligent driving and transportation control and active safety systems for road vehicles.

Bolin Gao is currently an associate research professor at the School of Vehicle and Mobility, Tsinghua University. His research interest is the dynamic design and control methodology of Intelligent and Connected Vehicles. He currently serves as a Secretary General of The Youth Committee of China-Society of Automotive Engineer. He has won special prize of China Automotive Industry Science and Technology Progress, and first prize of Science and Technology Progress of China Intelligent Transportation Systems Association, etc. He has been selected in the China Association for Science and Technology "Young Elite Scientists Sponsorship Program", Tianjin "131" Innovative Talent Training Project, etc.



Suying Zhang received the M.S. degree from Wuhan University of Technology in 2010. She is currently working with a research direction of engine after-treatment technology.

Guang Zhou graduated from Tsinghua University with a basic science class, and the Ph.D. degree in artificial intelligence from the University of Texas at Dallas. He is the "Young Leader" of strategic emerging industries in the Greater Bay Area, and a member of the Young Entrepreneur Committee of the Asia-Pacific Economic Cooperation China Business Council. With more than 20 years of experience in robotics and embedded systems, he has won several world-class awards in robotics competitions. He has more than 10 years of experience in computer vision and in-depth, machine learning research.

Aoning Jia received the B.S. degree in vehicle engineering from Hebei Agricultural University in 2020. Currently, he is pursuing the M.S. degree in Vehicle Engineering at School of Mechanical Engineering, Hebei University of Technology, China. He is engaged in research work in School of Vehicle and Mobility, Tsinghua University, China. His research interests include digital mapping and cloud control system.