Research Article

Dynamic data collection of staggered-following behavior in non-lane-based traffic streams

Sanhita Das[1] · **Akhilesh Kumar Maurya[1]** · **Anuj Kishor Budhkar[2]**

Abstract

Different maneuvering patterns of vehicles, absence of lane discipline and interactions of a large number of vehicles with each other and with roadway features make the traffic phenomena of non-lane-based traffic streams more complex. Vehicles' movement in weak lane discipline traffic is rather two-dimensional because they always tend to evaluate possible available gaps on the road while progressing longitudinally. Recent literature has underlined the importance of centerline separation of traffic in modeling the staggered-following behavior. However, the requirements for staggered-following trajectory data are indeed stringent. Although recent advancements in new digital technology have expedited new horizons in the field of traffic engineering, proper estimation of microscopic car-following data has still proven to be challenging. Understanding that a proper evaluation of car-following behavior in non-lane-based traffic environments requires an accurate characterization of the microscopic traffic variables and reliable experimental data, this study describes an image-based in-vehicle trajectory data collection system to process the microscopic variables (such as longitudinal gap, centerline separation, vehicle speeds and accelerations), using camera calibration and in-vehicle GPS information on straight roads. Improved accuracy in the experimental data collection, proper extraction and estimation of data can substantially enrich the understanding of riders' behavioral phenomena from a microscopic perspective and the realism of traffic sub-models, which will result in a better prediction of microscopic simulation models.

Keywords Heterogeneous · Car-following · Centerline separation · Data collection

1 Introduction

Recent advancements in the development of microscopic simulation models and intelligent transportation systems (ITSs) have spurred a growing interest in the transport modelers of many countries. Achieving a detailed understanding of how drivers react to the surrounding traffic, how they control the vehicles in the car-following process and different factors affecting their behavior will essentially enhance the realistic replication of riders' behavior in simulation modeling.

The peer-reviewed literature has supported the development of car-following theories and its subsequent sub-models since decades [4, 11, 12]. In particular, car-following condition refers to that state in which the subject vehicle assigns full leadership to the immediate vehicle in front. Unlike in homogeneous and lane-based traffic conditions where car-following behavior mostly prevails, vehicles in non-lane-based traffic environments not only interact with the front leading vehicle but also with the surrounding vehicle in the lateral direction. Due to the differences in the static and operational characteristics of diverse vehicle types in non-lane-based traffic streams, they often tend to look for possible available gaps in the surrounding traffic while progressing longitudinally. Therefore, the subject vehicles do not always fully follow

the leading vehicles in the longitudinal direction, but they maintain some lateral separation with the preceding vehicle (or centerline separation—CS) either to perceive the forward visual field with more confidence or to anticipate the behavioral response of the front vehicles. This behavior is commonly termed as 'staggered-following behavior' in the literature [2, 5, 8].

Modeling the staggered-following behavior of vehicles in non-lane-based traffic environments, however, requires proper collection, extraction and an accurate estimation of reliable experimental data. Although the advancements in new digital technology have expedited new horizons in the field of traffic engineering, the collection and processing of accurate unbiased, time-series data for an empirical verification of staggered-following behavior in non-lane-based traffic streams is still a challenging task.

Wolshon and Hatipkarasulu [18] showed the viability of GPS-based technology in the collection and processing of vehicle movement data for car-following study. Shekleton [17] also demonstrated the suitability of differential GPS in collecting accurate car-following data in typical urban environments. Considering the advantages of satellite tracking in terms of high resolution and excellent accuracy, Ranjitkar et al. [15] demonstrated the superiority of GPS and utilized the collected data to compare several car-following models. In the context of mixed-traffic environments, Ravishankar and Mathew [16] modified the Gipp's car-following model and incorporated vehicle-type-dependent parameters by conducting a series of experiments using GPS-equipped vehicles with 10 drivers. In another study, Jiang et al. [10] used high-precision GPS devices to investigate the evolution of traffic flow such as propagation, growth, dissipation and merge of disturbances in controlled car-following experiments. Different researchers have demonstrated the advantages of satellite-based Global Positioning System (GPS) technology in modeling the car-following behavior [9, 13, 14]. However, the data requirements of staggered-following behavior are indeed stringent. While the longitudinal spacing between the two vehicles can be obtained from the recorded GPS positions with 1 m accuracy, the estimation of lateral separation from the GPS receivers may produce unreliable results. Therefore, it is envisaged that an integration of satellite-based GPS technology with an image-based trajectory extracting system can prove to be suitable in the collection and estimation of staggered-following data.

This paper therefore attempts to establish an image-based in-vehicle data collection methodology using camera calibration and satellite-based GPS receivers on straight roads of non-lane-based traffic streams, to process the microscopic parameters (such as longitudinal gap, centerline separation and speed) in staggered-following scenario.

2 Methodological framework

A discussion on the experimental setup, data collection process and calibration of the video data is provided in this section.

2.1 Data collection process

The in-vehicle experimental setup consists of a GPS device (Racelogic Video VBOX) with 10 Hz data logging frequency and a video camera attached at the windshield of the following vehicle, allowing real-time monitoring of the forward visual field. The GPS device provides the vehicle position and speed with 1 m and 0.1 km/h accuracy, respectively, while the video recorder captures continuous video data at 25 frames/s. For the purpose of the study, two cars were equipped with GPS receivers and experiments were conducted along straight rural roads on NH-27 in Guwahati. Both the vehicles were then allowed to move in the following state so that the car-following (or staggered-following to be more precise) data could be further processed and analyzed. In this process, it was ensured that the vehicles were not involved in any lane-changing or overtaking processes, and for any intrusions in between the two GPS-equipped vehicles, the corresponding data were discarded from further analysis.

2.2 Calibration

While information on speeds and accelerations of both the vehicles is obtained from the GPS receivers, the longitudinal and lateral separations between the two successive vehicles can be extracted from the video footage. Camera calibration of the recorded video is conducted by a vanishing point technique, where image coordinates of the four corners forming a perfect rectangle in the real field need to be captured [6]. With an aim to extract and analyze the longitudinal and lateral spacing, a semiautomated trajectory extractor is developed, where the image/screen coordinates of the video footage are obtained by using mouse clicks at different time stamps and the corresponding screen coordinates are then converted to real-world coordinates by utilizing [6] calibration equations. The camera calibration technique used in this study is depicted in Fig. 1.

Before starting the experiment, four endpoints of the road (ABCD as shown in Fig. 1a) representing an exact rectangle of known dimensions are marked on the road. The corresponding endpoints are further extracted from the video footage by manual mouse clicks on the screen to obtain its respective screen coordinates. For the estimation
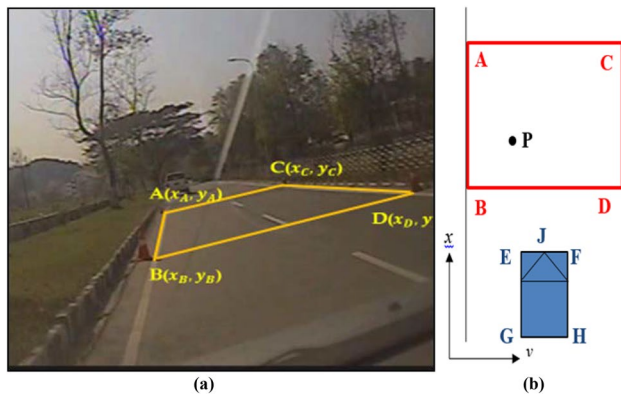
Fig. 1 Camera calibration technique: **a** rectangle *ABCD* viewed from the camera, **b** representation of the calibration pattern

Fig. 2 Variables considered in the study [5]

of inter-vehicle longitudinal and lateral separations, let us consider the test vehicle (EFGH) with its front center (point J) parked parallel to the road edge (AB). The lateral distance and longitudinal distance of one of the marked endpoints from one corner of the test vehicle also need to be known (say point B from point E). For any point P lying in the plane ABCD (i.e., plane of the road), its camera coordinates (x_P, y_P) can be calculated and subtracted from coordinates of point B and E so as to obtain the lateral and longitudinal offset of point P. In other words, lateral offset or centerline separation $= (y_P - y_B) - (y_E - y_B) + (y_J - y_E)$ and longitudinal offset $= (x_B - x_P) + (x_E - x_B)$. The distances $(x_E - x_B)$, $(y_E - y_B)$ and $(y_J - y_E)$ are accurately measured in the field. The screen coordinates are extracted at every 0.2 s and are then transformed into real coordinates as discussed before. The recorded vehicle positions from the GPS receivers are then synchronized with the positional data extracted from the video footage at every 0.2-s intervals.

3 Results and analysis

3.1 Preliminary analysis

In the staggered-following scenario, several variables were extracted from the video recorders and the GPS receivers, such as speeds of the leading and following vehicles, acceleration, longitudinal gap and centerline separation. The dataset consists of detailed trajectory data of cars at every 0.2-s intervals. A representation of longitudinal gap (LG) and centerline separation (CS) considered in the study is presented in Fig. 2.

In total, 7200 cases of time-series data for staggered car-following scenario are obtained in this study. The trajectory data, however, exhibited some noise artifacts; hence, they were filtered by applying moving average filter
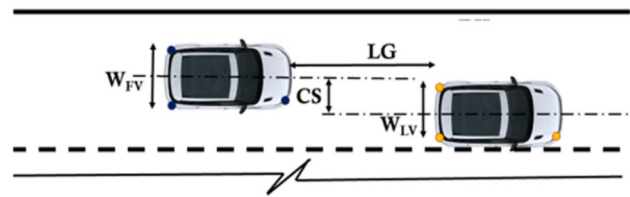
for duration of 1 s before any further analysis. A comparison of real and filtered data for acceleration and relative speed is presented in Fig. 3.

An accurate representation of car-following behavior should consider the driver's perception and its behavioral response. The driver in the following vehicle has direct control over the brakes and accelerator of his own vehicle, and therefore, he reacts in response to the actions of the leading vehicle. The variation in relative speed and acceleration forms the basis of stimulus and reaction in car-following models. Figure 4 depicts the relative speed and acceleration profiles of the following vehicle during 150-s recording time.

From the figure, it is clear that the changing tendency of following vehicle's acceleration is quite similar to the relative speed. The higher the acceleration rate offered by the leading vehicle, the larger will be the relative speed between the vehicles, and based on driver's perception, the following vehicle will react accordingly after a certain time lag. This time lag is considered as driver's reaction delay which can be obtained from the time difference between two subsequent variations in relative speed and acceleration, as indicated by arrows in the figure. Similar to relative speed and acceleration, the variations in longitudinal gap and speed are presented in Fig. 5, which also depicts the same changing pattern. This pattern clearly indicates that drivers can perceive the gap between the vehicles and adjust their speeds based on the current available gaps.

However, for a proper representation of car-following behavior in non-lane-based traffic streams, the lateral interaction (or centerline separation) data need to be extracted and processed for car-following model development. As already discussed, CS is usually considered as an indicator of lateral interaction in the car-following process of non-lane-based traffic. Essentially, CS < 0.34 m represents a car-following state [7], while 0.34 < CS < 3 m depicts a staggered-following case where the subject vehicles interact with the leading vehicles maintaining certain centerline separation between them.

The obtained trajectory data indicated a wide range of longitudinal gap, speeds and centerline separation varying from a range of 3.4–29.89 m, 10–78 km/h and
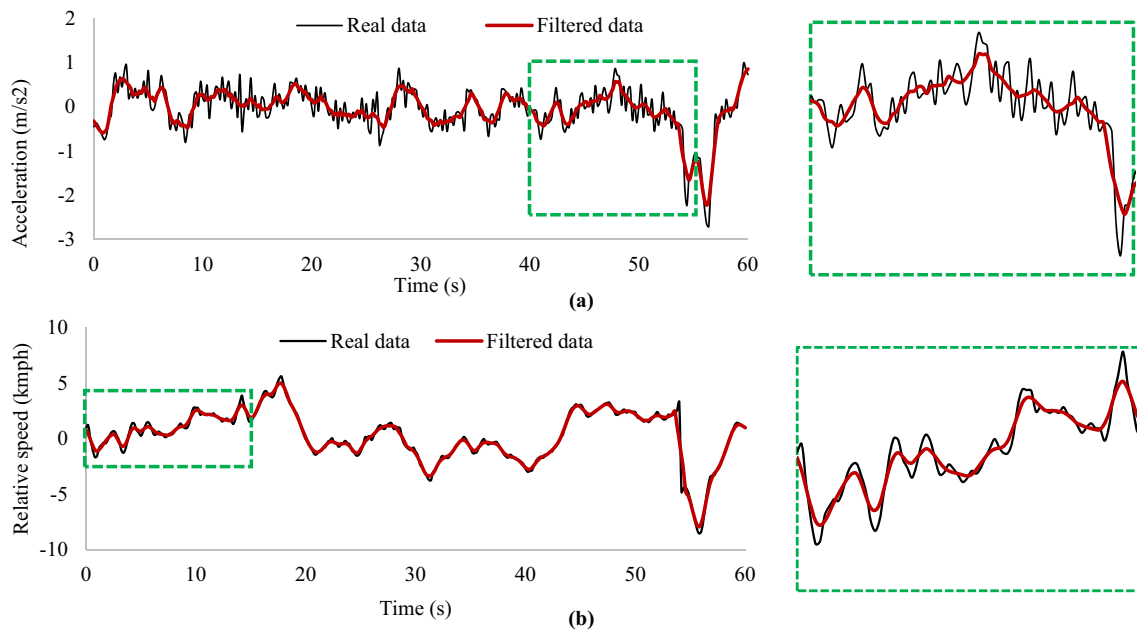
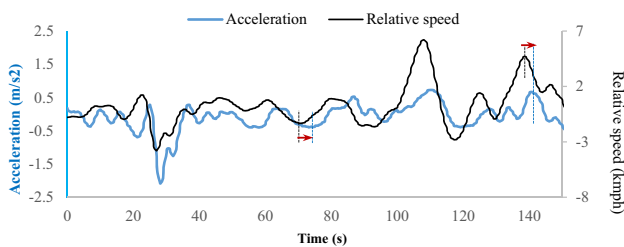**Fig. 3** Comparison of the obtained real data and filtered data for **a** acceleration and **b** relative speed

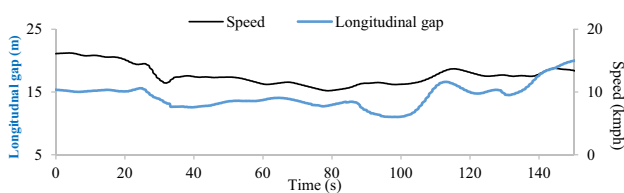**Fig. 4** Relative speed and acceleration profile

Fig. 5 Longitudinal gap and speed profile

0–3 m, respectively. The observed range for longitudinal gap (< 30 m) further justifies the longitudinal interaction region for car-following-car cases, as indicated in the previous car-following research work [1].

3.2 Univariate modeling of the traffic variables

Proper estimation of probability distribution models for longitudinal gap, vehicle speeds and centerline separation

can provide a better understanding of the stochastic uncertainties in the car-following processes. The probability distributions of these variables are also considered as dominant input parameters for generating vehicles in microsimulation modeling.

Several statistical models were used to fit longitudinal gap, centerline separation and vehicle speed data. The candidate distributions selected for the study are logistic, Weibull, lognormal, normal and gamma. The maximum likelihood technique is employed to estimate the parameters of the univariate models, while the suitability of a distribution is evaluated using log-likelihood values and two goodness-of-fit statistics such as Kolmogorov–Smirnov (K–S) and Anderson–Darling (A–D) tests. Hence, a particular distribution is considered to be the best-fitted one if the log-likelihood value is the largest and the selected distribution passes the goodness-of-fit tests (i.e., the statistic values are lower than the critical values at 5% significance level). Table 1 presents the log-likelihood (LL) values and the goodness-of-fit statistics of the selected distributions for longitudinal gap, centerline separation and vehicle speeds.

Based on the LL values and goodness-of-fit results, logistic distribution provided the best-fits for longitudinal gap and speed data, while normal distribution was found to the best-fitted one for centerline separation. The distribution profiles of the best-fitted statistical models for LG, CS and vehicle speeds are presented in Fig. 6.

The histogram plots of Fig. 6 indicate that inter-vehicle longitudinal spacing ranges from 5 m to a maximum of

Table 1 Log-likelihood and goodness-of-fit statistics for each microscopic traffic variable

Distribution type	LG			CS			Speed		
	LL	KS	AD	LL	KS	AD	LL	KS	AD
Lognormal	−6486.0	0.05	12.25	−1508.2	0.10	34.31	−7154.04	0.08	41.74
Weibull	−6469.0	0.09	25.67	−1563.3	0.11	41.56	−7352.6	0.16	91.71
Logistic	**−6426.9**	**0.04**	**7.87**	−2285.9	0.21	153.5	**−7091.07**	**0.10**	**35.8**
Normal	−6451.0	0.05	9.62	**−1494.5**	**0.06**	**24.37**	−7169.28	0.12	56.29
Gamma	−6481.2	0.05	12.25	−1899.6	0.06	24.37	−7111.13	0.08	41.74

Bold features indicate the best-fitted distribution

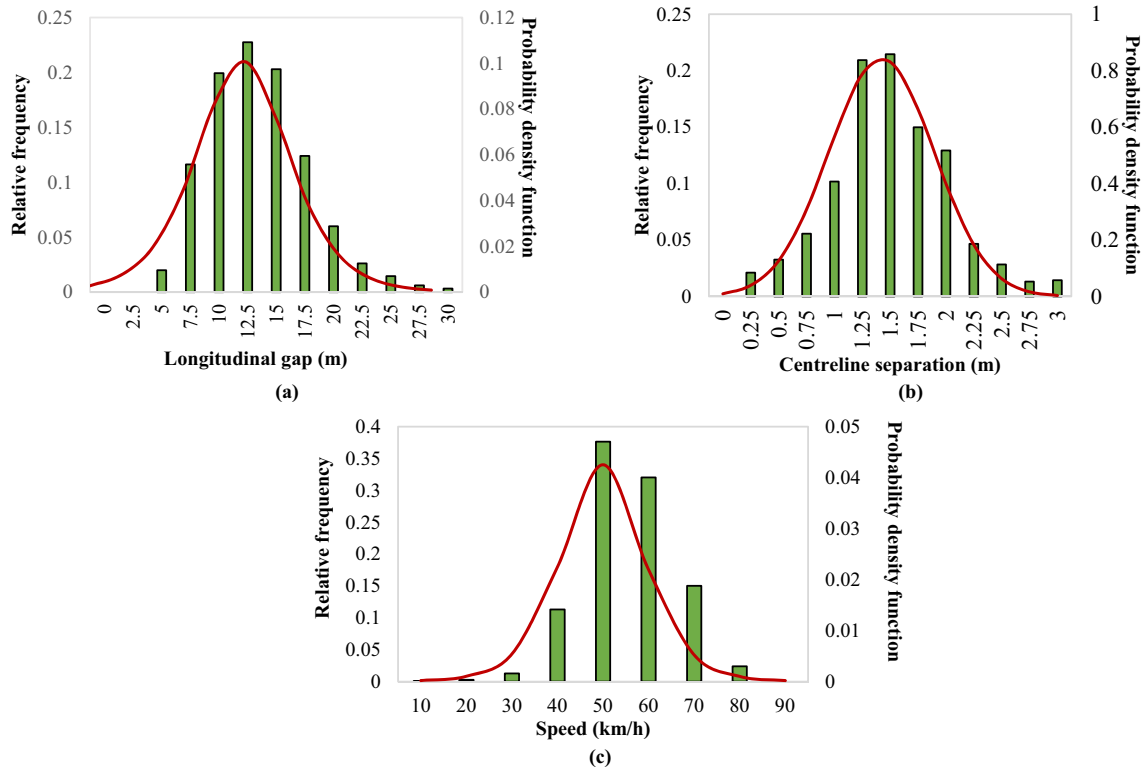*LG* longitudinal gap, *CS* centerline separation

Fig. 6 Best-fitted probability density curves for **a** longitudinal gap, **b** centerline separation and **c** vehicle speeds

30 m, while the peaks of the distributions lie at 12.5 m, 1.5 m and 50 km/h for LG, CS and speeds, respectively. As importantly, lower values of LG may not actually imply an unsafe car-following event as the following vehicles have the flexibility to move laterally maintaining large CS even at lower longitudinal gaps. The variation in LGs with CS and speeds may provide a better understanding of the car-following events of non-lane-based traffic streams.

3.3 Relationship among LG, CS and speed

Firstly, to understand the interrelationship between LG and speed in staggered-following conditions, the variations in longitudinal gaps with different speed ranges across all centerline separations are presented in a box plot as shown in Fig. 7.

The connecting line in Fig. 7 represents the mean values of LG at each speed range, and it clearly depicts an increasing relationship between the two variables. As expected, significant statistical differences in the variances of longitudinal gaps were observed across all speed ranges at 5% significance level ($Fstat = 381.89, p < 0.001$). This indicates that as longitudinal gap between the interacting vehicles increases, the following vehicles will accelerate and proceed at higher speeds to avail the gap. Moreover, the consideration of lateral descriptor of vehicle interactions (that is, centerline separation) in car-following events will provide additional insights into the variations in LG with

Fig. 7 Box plot showing variations in LG with speeds
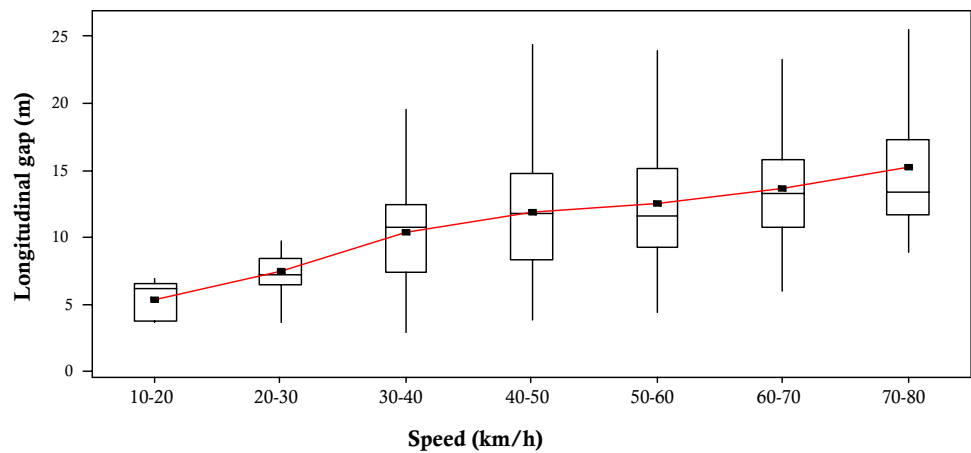
speeds. The descriptive statistics of LGs and speeds at different CS ranges are therefore evaluated, and a summary of the values is presented in Table 2.

It can be observed from the table that the mean and median values of longitudinal gaps follow a decreasing trend as centerline separation increases. This implies that vehicles at large CS feel lesser obstruction of the leading vehicles, have wide visual field of view and can anticipate the leading vehicle's behavior due to which they tend to follow the leading vehicles closely, resulting in lower longitudinal gaps. This decreasing trend of longitudinal gap with CS has also been justified in [3, 7] work where they have showed that vehicles in weak lane discipline traffic maintain shorter headways with the increase in horizontal separation between them. However, the mean and median values of speeds reveal a decreasing relationship till 2-m CS, while the speed values increase beyond 2-m CS. This can be attributed to the fact that the following vehicles reduce their speeds when they interact with the leading vehicle below 2-m CS, whereas at CS greater than 2 m, the extent of influence of the leading vehicles comparatively reduces and the following vehicles feel less constrained. As a result, they tend to proceed at higher speeds at larger CS (> 2.5 m).

The variations in average longitudinal gap with speed for different centerline separations are presented in Fig. 8a, while Fig. 8b displays the variation in longitudinal gap with centerline separation for different speed ranges.

As discussed above, the observed trend further justifies the increasing relationship of LG and speed for each CS range. With the increase in CS between the vehicles, the longitudinal gap decreases at different following vehicle's speeds. Moreover, the decreasing relationship between LG and CS is well depicted in Fig. 8b where it is clear that the decreasing trend goes up with the increase in speeds. This is quite expected because at a particular CS level, as the LG between the vehicles increases, the following vehicles tend to proceed at higher speeds to avail the gap.

4 Conclusions

Although GPS technology can prove to be a promising technique in vehicle tracking, the collection of real-time trajectory data for staggered car-following conditions has still proven to be challenging. While the longitudinal spacing between the two vehicles can be obtained from the recorded GPS positions with 1 m accuracy, the estimation of lateral separation from the GPS receivers may produce unreliable results. This study therefore attempts to provide an in-vehicle data collection methodological approach using camera calibration and satellite-based GPS receivers,

Table 2 Descriptive statistics of LG and speed for different CSs

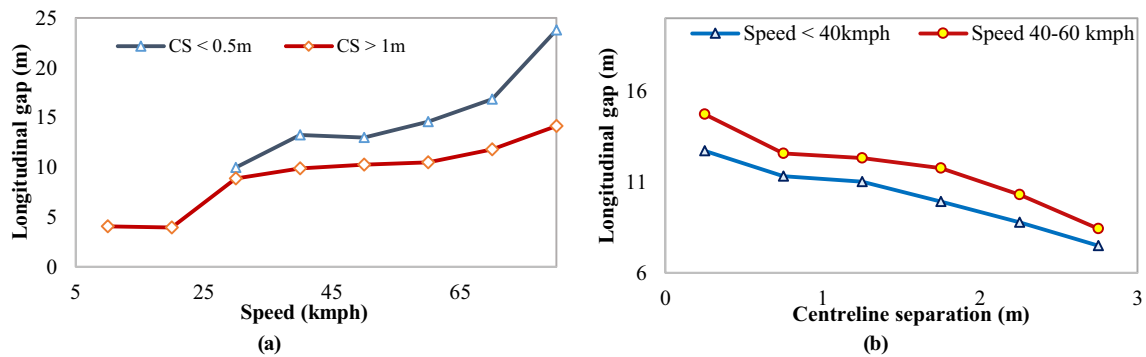| CS (m)	LG (m)			Speed (km/h)			Sample
	Mean	Median	SD	Mean	Median	SD	
0–0.5	17.76	18.51	4.99	54.66	56.3	10.29	387
0.5–1	14.36	13.84	4.71	53.84	53.09	8.72	1024
1–1.5	12.48	12.26	3.91	50.68	49.92	10.17	2424
1.5–2	11.16	11.07	3.47	47.33	44.71	9.51	2622
2–2.5	9.31	8.49	2.80	47.56	44.60	10.23	539
2.5–3	7.68	6.88	2.47	54.02	54.13	8.32	198

SD standard deviation

Fig. 8 Variation in longitudinal gap with **a** speed and **b** centerline separation

to process reliable dynamic time-series data (longitudinal gap (LG), centerline separation (CS) and vehicle speeds) for staggered car-following conditions of non-lane-based traffic streams.

Preliminary analysis on LG, CS and speed data indicated that logistic distribution provided the best-fit for longitudinal gap and speed data, while normal distribution was found to the best-fitted one for centerline separation in staggered car-following conditions of non-lane-based traffic streams. As expected, the results further indicated a positive dependent relationship between LG and CS and a reciprocal relationship between LG and CS. This implies that vehicles at large CS feel lesser obstruction of the leading vehicles, have wide visual field of view and can anticipate the leading vehicle's behavior due to which they tend to follow the leading vehicles closely, resulting in lower longitudinal gaps. It was further observed that the following vehicle speeds increase when CS exceeds 2 m.

The results of this study can substantially enrich the understanding of riders' behavioral phenomena from a microscopic perspective and the realism of traffic sub-models, which will result in a better prediction and development of microscopic simulation models.

Compliance with ethical standards

References

1. Asaithambi G, Joseph J (2018) Modeling duration of lateral shifts in mixed traffic conditions. J Transp Eng Part A: Syst 144(9):04018055
2. Asaithambi G, Kanagaraj V, Toledo T (2016) Driving behaviors: models and challenges for non-lane based mixed traffic. Transp Dev Econ 2(2):19. https://doi.org/10.1007/s40890-016-0025-6
3. Budhkar AK, Maurya AK (2017) Multiple-leader vehicle-following behavior in heterogeneous weak lane discipline traffic. Transp Dev Econ 3(2):20
4. Brackstone M, McDonald M (1999) Car-following: a historical review. Transp Res Part F: Traffic Psychol Behav 2(4):181–196
5. Das S, Maurya AK (2018) Multivariate analysis of microscopic traffic variables using copulas in staggered car-following conditions. Transportmetr A: Transp Sci. https://doi.org/10.1080/23249935.2018.1441200
6. Fung GS, Yung NH, Pang GK (2003) Camera calibration from road lane markings. Opt Eng 42(10):2967–2977. https://doi.org/10.1117/1.1606458
7. Gunay B (2003) Methods to quantify the discipline of lane-based driving. Traffic Eng Control 44(1):22–27
8. Gunay B (2007) Car following theory with lateral discomfort. Transp Res Part B: Methodol 41(7):722–735
9. Gurusinghe GS, Nakatsuji T, Azuta Y, Ranjitkar P, Tanaboriboon Y (2002) Multiple car-following data using real-time kinematic global positioning system. Transp Res Rec: J Transp Res Board 1802:166–180
10. Jiang R, Hu MB, Zhang HM, Gao ZY, Jia B, Wu QS, Yang M (2014) Traffic experiment reveals the nature of car-following. PLoS ONE 9(4):e94351
11. Kesting A, Treiber M (2008) Calibrating car-following models by using trajectory data: methodological study. Transp Res Rec: J Transp Res Board 2088(1):148–156
12. Mahapatra G, Maurya AK, Chakroborty P (2018) Parametric study of microscopic two-dimensional traffic flow models: a literature review. Can J Civ Eng 45(11):909–921
13. Mathew TV, Ravishankar KVR (2012) Neural network based vehicle-following model for mixed traffic conditions. Eur Transport/Trasporti Europei 52:1–15
14. Punzo V, Simonelli F (2005) Analysis and comparison of microscopic traffic flow models with real traffic microscopic data. Transp Res Rec: J Transp Res Board 1934:53–63
15. Ranjitkar P, Nakatsuji T, Kawamua A (2005) Car-following models: an experiment based benchmarking. J Eastern Asia Soc Transp Stud 6:1582–1596
16. Ravishankar KVR, Mathew TV (2011) Vehicle-type dependent car-following model for heterogeneous traffic conditions. J Transp Eng 137(11):775–781
17. Shekleton S (2002) A GPS study of car following theory. undergraduate honours project at UNSW, UNSW School of Civil and Environmental Engineering
18. Wolshon B, Hatipkarasulu Y (2000) Results of car following analyses using global positioning system. J Transp Eng ASCE 126(4):324–331