Validation of lane-changing behaviour in a weaving section using a driving simulator experiment

Transportation Research Part F: Psychology and Behaviour 115 (2025) 103364

Contents lists available at ScienceDirect

Transportation Research Part F: Psychology and Behaviour

journal homepage: www.elsevier.com/locate/trf

Validation of lane-changing behaviour in a weaving section using a driving simulator experiment

Suyi Mao [a,b,*], Jaeyoung Jay Lee [a,c,d], Arastoo Karimi [b], Alessandra Lioi [b], Marco Bassani [b,*]

[a] School of Traffic and Transportation Engineering, Central South University, Changsha, Hunan 410075, China
[b] Department of Environment, Land and Infrastructure Engineering (DIATI), Politecnico di Torino, Torino, Piemonte 10129, Italy
[c] Department of Civil, Environmental and Construction Engineering, University of Central Florida, Orlando, FL 32816, USA
[d] School of Civil and Environmental Engineering, Faculty of Engineering, Queensland University of Technology (QUT), Brisbane, QLD 4000, Australia

ARTICLE INFO

Keywords:
Lane change
Weaving section
Driving simulator
Validation
Fidelity

ABSTRACT

Driving simulators have been extensively utilized to investigate driving behaviour. Nevertheless, concerns persist regarding their validity and reliability in accurately replicating real-world driving. No study has directly compared lane-changing behaviour in actual expressway weaving sections with that observed in a driving simulator. To address the research gap, a test track is designed to replicate the geometric characteristics of a real-world weaving section. The experiment involves 42 participants. Data on driving behaviour and performance are collected to evaluate key traffic efficiency and lane change-related variables. Kolmogorov-Smirnov test assesses the discrepancies between field observations and simulation outcomes. The results identify the statistical similarities in speed and headway across most manoeuvres, lane change duration and accepted lead gap for mandatory lane changes, as well as generalized time-to-collision (*GTTC*) in discretionary lane changes. However, simulation drivers tend to drive faster particularly in ramp scenarios due to reduced speed perception. Some novice drivers maintain larger headways and spend less time on mandatory lane changes. Furthermore, the location of merging and diverging positions significantly influences the lane change duration. During simulated driving, poor rear visibility leads to more forced lane changes and larger accepted gaps. The *GTTC* values from the simulation are higher than those from field, except for discretionary lane changes, with younger drivers exhibiting smaller *GTTC* values. Lastly, the simulation data suggest challenges in replicating the dangerous scenarios resulting from the smaller accepted gaps observed in field. The findings offer support for the broader implementation of driving simulation experiments for weaving sections.

1. INTRODUCTION

Weaving is defined as the crossing of two or more traffic streams traveling in the same general direction along a significant length of roadway without the aid of traffic control devices. Weaving sections are formed when a merge area is closely followed by a diverge area, or when an on-ramp is closely followed by an off-ramp and the two are joined by an auxiliary lane (HCM, 2000). Weaving sections

* Corresponding authors.
E-mail addresses: suyimao@csu.edu.cn (S. Mao), marco.bassani@polito.it (M. Bassani).

https://doi.org/10.1016/j.trf.2025.103364
Received 30 March 2025; Received in revised form 8 September 2025; Accepted 8 September 2025
Available online 16 September 2025

are critical, recurrent bottlenecks that negatively affect the efficiency and safety of roadways (Hall & Agyemang-Duah, 1991; Cassidy & Bertini, 1999; Golob et al., 2004; Chung et al., 2007). Merging and diverging manoeuvres in proximity require vehicles to change lanes frequently at weaving sections, leading to a reduction in traffic capacity and an increase in the risk of collisions (Marczak et al., 2016; Sulejic et al., 2017; Chen & Ahn, 2018).

Investigating drivers' lane change behaviour at expressway weaving sections is of significance for identifying the influencing factors associated with traffic delays and conflicts. Some researchers have collected field data in weaving sections to examine key factors affecting lane change behaviour (Sarvi et al., 2011; Marczak et al., 2014; Kusuma et al., 2015; Ouyang et al., 2023). Their findings suggest that geometric design factors (the length of weaving sections, different types of weaving sections, the presence of an auxiliary lane, etc.) and surrounding traffic conditions, including traffic volume, vehicle type composition and weaving ratio, significantly influence lane change operations. Other studies have focused on modelling lane change behaviour in weaving sections through traffic simulation software (Rakha & Zhang, 2004; Bham, 2011). These studies demonstrate that fine-tuning parameters in lane change models can enhance the alignment between simulated outcomes and real-world data. Recent research has increasingly employed driving simulators to explore lane change behaviour in weaving sections (De Blasiis et al., 2017; Cai et al., 2018; Yuan et al., 2019). These investigations reveal that driver characteristics (i.e., drivers' gender and age) play a critical role in lane change decisions. Several researchers have examined lane change decision-making and mechanisms in weaving sections (Wan et al., 2017; Sulejic, 2018; Kusuma et al., 2020; Zhao et al., 2022). Other studies have explored safety and efficiency improvements as well as lane change decision-making in weaving sections under connected and automated driving conditions (Tilg et al., 2018; Hao et al., 2020; Amini et al., 2021; Li et al., 2021; Yang et al., 2022; Zhou et al., 2023).

Lane-changing behaviour significantly affects both traffic efficiency and safety. Therefore, investigating weaving sections—where frequent lane changes occur—is crucial. Such an analysis would help us understand the mechanisms behind lane changes and support the establishing effective traffic management policies to enhance traffic efficiency and safety. Furthermore, with the advancement of intelligent transportation systems, a key area for further research is determining whether connected and automated vehicles can satisfy stringent safety performance requirements in weaving sections. The research on the weaving section primarily encompasses three methods of data acquisition: field data collection, simulation modelling, and driving simulation experiments. Due to the complexity of road geometry and the intricacies involved in video processing, acquiring accurate data from real-world weaving sections poses significant challenges. Simulation model-generated data often lack robust validation against real-world datasets, particularly in the context of driving behaviours. Replicating human drivers' lane-changing behaviours in complex weaving flows is especially difficult, and variations in driving behaviours influenced by demographic characteristics are often overlooked. Driving simulator serves as an effective tool for capturing micro-driving behaviours and provides a reliable platform for investigating lane-changing behaviours in weaving sections. Although driving simulators have gained widespread use in recent years, the validity of their output remains a critical concern (Mullen et al., 2011; Wynne et al., 2019; Zhang et al., 2024). Previous studies have validated driving simulators under various driving behaviours and road conditions (Godley et al., 2002; Bella, 2005; Underwood et al., 2011; Meuleners & Fraser, 2015; Branzi et al., 2017; Bassani et al., 2018; Catani & Bassani, 2019; Karimi et al., 2020; Faschina et al., 2021; Olsson, 2023; Chen et al., 2025). Previous studies on the calibration of driving simulators have primarily focused on relatively simple road geometric conditions, such as straight sections and curves, with limited attention given to more complex configurations. Furthermore, most of these studies relied on macroscopic measures such as speed and lane position, while few explored potential differences in microscopic lane-changing behaviour. To date, no study has specifically validated lane-change behaviour in weaving sections using driving simulator and real-world data. Addressing this gap by investigating the validity of driving simulators in weaving sections is important. Such an investigation would yield insights into the effectiveness of simulators for analysing lane-changing behaviour under complex roadway conditions. This, in turn, would establish a solid foundation for future research, including the development and evaluation of intelligent traffic management systems and the safety assessment of connected and automated vehicles.

To address the identified research gap, this study aims to replicate a highly consistent driving environment of the weaving section in a driving simulator by utilizing real-world data obtained from video recordings. This study meticulously designs the driving simulation experiment to collect driving behaviour data in the weaving section, defines crucial variables for statistical testing and analysis, and uncovers the underlying mechanisms of lane-changing behaviour. Both field and simulation data were collected to compare the lane-changing behaviour characteristics of drivers in the weaving section, including the quantification of basic traffic efficiency metrics and lane-changing patterns. The structure of this study is as follows: Section 2 details the method, encompassing the study objectives, field data processing, materials and equipment used, experimental design, participants information, selected variables for analysing lane-changing behaviour, and data collection and manipulation. Section 3 compares the results from both field and simulator experiments, while Section 4 discusses these findings. Finally, Section 5 concludes the paper with implications.

2. METHOD

The video footage of a weaving section in Changsha Airport Expressway (China) was captured by three DJI Mavic 2 unmanned aerial vehicles (UAVs) for a duration of 25 min. Computer vision technology was utilized to derive the vehicle trajectory data (Section 2.1). The simulator experiment was conducted at the fixed-base driving simulator of the Road Safety and Driving Simulation Laboratory (RSDS) at Politecnico di Torino, Italy. The simulation scenarios accurately replicated the road geometry and traffic conditions observed in the field. The experimental process was meticulously designed, and participants were recruited to conduct the experiment (Section 2.2). Considering distinct data source structures and driving behaviour characteristics at expressway weaving sections, specific observed variables were defined and derived to represent lane change manoeuvres (Section 2.3). Subsequently, statistical comparison methods were employed to investigate the significance of differences between traffic efficiency and safety indexes observed in both field studies and driving simulation experiments (Section 2.4).

2.1. Field study

On May 31, 2021, from 3:00 PM to 3:25 PM under a favourable weather condition, three UAVs captured footage of a representative weaving section on the airport expressway in Changsha, Hunan Province, China. The UAVs were positioned at an altitude of 120 m above ground level. The weaving section depicted in Fig. 1 spans a length of 400 m and comprises six lanes in two directions, including two on-ramps and one off-ramp. There are three lanes in the weaving section, the main road has two lanes with a width of 3.75 m, and the auxiliary lane is 5 m wide. The video was recorded at 30 frames per second, resulting in a vehicle trajectory resolution of the same frequency. Subsequently, the individual video segments obtained from each UAV were seamlessly stitched together using advanced video stitching techniques implemented in OpenCV C++. In order to determine the precise latitude and longitude coordinates for both the road and moving vehicles within the footage, reference points such as road edges and intersections where lane lines converge on on– and off-ramps were selected. These reference points' geographical coordinates were measured onsite using a compass instrument to facilitate subsequent computer vision processing to derive trajectory data of moving vehicles (Yang et al., 2023).

In this study, the Mask Region-CNN algorithm was employed for vehicle detection. The Mask Region-CNN algorithm, an advanced version of Faster Region-CNN primarily utilized for target detection and semantic segmentation tasks, is a deep learning model. Further details on traffic video processing can be found in Zheng (2019). The vehicle track processed by the Mask Region-CNN algorithm undergoes further refinement. Specifically, a moving average method was employed to effectively mitigate data noise and ensure smoother trajectory representation. Additionally, an elimination process was implemented to exclude anomalous data instances such as stationary vehicles over prolonged periods, vehicles exhibiting abnormal speeds, and trajectories with insufficient detection time.

Due to the limitations in UAVs shooting height and angle, the upper half of the weaving section's off-ramp was beyond the range of UAVs footage. Therefore, this study focuses on analysing the lower half of the filmed video's weaving section. The final data obtained from image processing included vehicle trajectory information for both main lanes and on/off ramps at the lower half of the weaving section. Finally, the data of 1,669 vehicles were retained.

The speed of vehicle trajectories in the field weaving section, as depicted in Fig. 2, were modelled by a normal distribution function that passed the Kolmogorov-Smirnov test, with an average value of 55.9 km/h and a standard deviation of 11.3 km/h. It is important to note that the field-collected speed data reflect the instantaneous speeds of all vehicles—both crossing and lane-changing—while passing through cross-sections of the entire weaving section at a sampling frequency of 30 Hz. Despite the speed limit being set at 100 km/h on this weaving section, the presence of vehicles driving at low speeds to exit or enter the expressway from on/off ramps intensified traffic flow weaving behaviour, resulting in an overall lower travel speed. Previous studies conducted on weaving sections have demonstrated similar results to those observed in this study, with vehicle average speeds ranging from 50 km/h to 70 km/h (Sulejic et al., 2017; Yang et al., 2012; Yuan et al., 2019). Prior studies have demonstrated that the exponential, gamma, and lognormal distributions are widely used for modeling headway probabilities under different traffic conditions (Cowan, 1975; May 1990; Saha et al., 2019). Following a comparison of the goodness-of-fit with other distributions (i.e., gamma and lognormal distributions), the headways, as depicted in Fig. 3, were sampled from exponential distribution, and truncated within intervals ranging from 2 to 8 s. Given the limited duration of the experiment, headways exceeding 8 s were truncated.

2.2. Simulation study

2.2.1. Driving simulator equipment

The simulation study was conducted using the fixed base driving simulator of the Road Safety and Driving Simulation Laboratory (RSDS) at Politecnico di Torino (Italy). The simulator was equipped with three 32-inch monitors (resolution: 1920 × 1080 pixels, frequency: 60 Hz, field of view: 130° × 20°), a fully outfitted driving position that included a seat, instrument panel, force feedback steering wheel, pedals, manual transmission, and vibration pads to simulate road roughness, wheel roll, and impacts. During the experiments, the cockpit images displayed on the screens enabled drivers to visualize the vehicle's width, adding realism to the simulation. A Dolby Surround 5.1 sound system provided an authentic reproduction of the car engine, nearby traffic, and the surrounding environment. SCANeR Studio® simulation software was utilized to build the driving scenarios, run the simulation, and

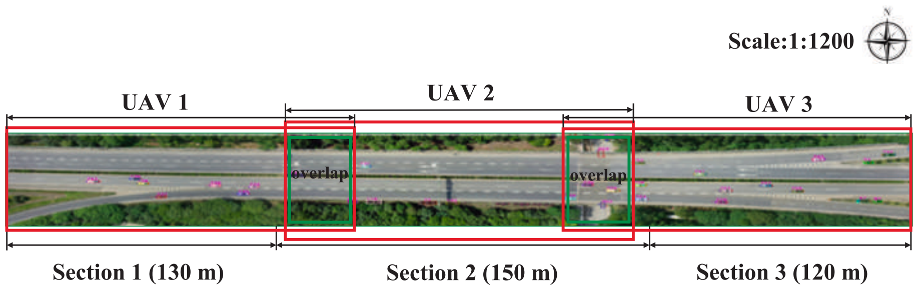


Fig. 1. A snapshot of the weaving section of field study.

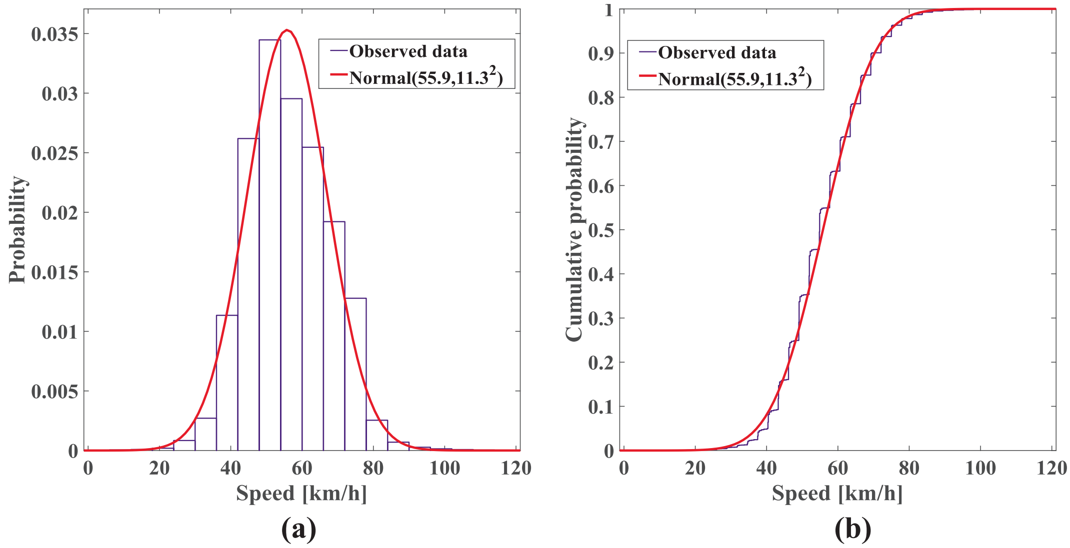


Fig. 2. Fitted normal distribution on speed of field study: (a) speed probability distribution and (b) speed cumulative probability distribution.

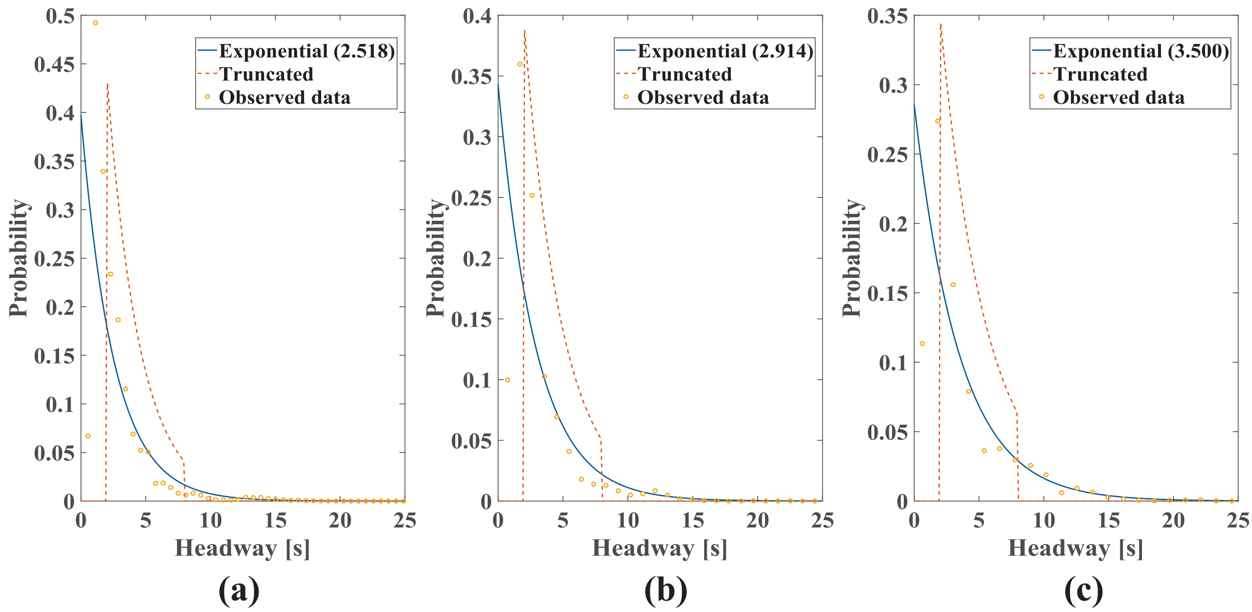


Fig. 3. Fitted exponential truncated distributions on field headway: truncated headway distribution of left lane (a), main lane (b) and auxiliary lane (c).

collect the driving data. The driving simulator was validated for longitudinal and transversal behaviour, and the passing manoeuvre (Bassani et al., 2018; Catani & Bassani, 2019; Karimi et al., 2020).

2.2.2. Test track

A meticulously designed test track was utilized, incorporating field weaving section map and road terrain data. By employing field satellite imagery and conducting on-site investigations, a comprehensive three-dimensional virtual model was developed, incorporating additional road infrastructure, vegetation, and buildings. The comparison between the field satellite imagery and the simulated three-dimensional model is depicted in Fig. 4.

2.2.3. Driving simulator experimental design

In this study, three independent variables were considered for designing the experiments: two numeric factors (i.e., speed of surrounding vehicles and participant age) and one categorical factor (i.e., participant gender). Setting appropriate traffic flow parameters for the weaving section in driving simulation experiments is crucial to ensure an accurate replication of field traffic data collection scenarios. To determine the test range of simulated surrounding vehicles speed, a response surface central composite design was employed, treating speed as a numeric factor. Based on the field data fitted to a normal distribution data as shown in Fig. 2; the 15th percentile corresponds to a speed of 44.1 km/h, while the 85th percentile corresponds to a speed of 67.7 km/h.

To control for variation due to test drivers, age and gender were included as factors. The age range of participants was set from 25 to 55 years to encompass diverse age categories, spanning from young to middle-aged drivers. A response surface central composite design was employed (Montgomery & Runger, 2020) to structure the experiment. The corresponding codes and actual values for each variable are presented in Table 1. The levels of continuous factors were selected to ensure a rotatable design space. To implement the response surface central composite design, a total of 26 experimental runs were conducted with Italian participants, as illustrated in Table 2. Five central runs for each gender were used to have reasonably stable variance for prediction. This design ensures the acquisition of effective and reliable results within a limited number of trials.

Since the field observations was collected in China, 16 young Chinese who had recently come to Italy and possessed no prior driving experience in Italy were also included. This minimizes potential confounding effects arising from variations in driving environments across different countries. The speed levels of test scenarios for Chinese drivers were determined using a random extraction method, while ensuring that the proportion remains almost consistent with the experimental runs of central composite design as depicted in Table 3.

The traffic flow in each lane and each manoeuvre (i.e., through, merging, diverging and ramp-to-ramp) in the simulation environment was considered fixed and determined based on data collected by the field study. The headway between every two vehicles for simulation was randomly extracted from the truncated headway distributions of three different lanes in weaving section shown in Fig. 3. The vehicles observed in the field were predominantly passenger cars, with a small proportion of vans. In the simulation experiment, the distribution of vehicle types closely mirrored the field data; however, all vehicles operated by participants were the same passenger car type. All scenarios were designed with daytime and good weather conditions, in accordance with the field study.

2.2.4. Participants

In accordance with the Code of Ethics of the World Medical Association (Association, 2024), forty-two licensed drivers were recruited on a voluntary basis without receiving any form of compensation or remuneration. Italian participants were recruited via email invitations sent to individuals followed by random selection from the pool of eligible candidates. Chinese participants were recruited through social media announcements. Prior to the experimental session, all participants provided informed consent by signing a consent form. Demographic characteristics of the test drivers are summarized in Table 4. Gender distribution was balanced between male and female participants. Italian test drivers represented a broader range of age groups and include more experienced drivers, whereas Chinese test drivers were predominantly young and less experienced. The differences in gender, age, and driving experience may influence driving behaviour in simulators (De Blasiis et al., 2017; Cai et al., 2018; Yuan et al., 2019). Additionally, cultural differences between Italian and Chinese test drivers may also affect their behaviour in simulated environments (Wang et al., 2019). Field data consist of vehicle trajectories collected from UAVs. Due to the lack of demographic data, the age distribution and driving experience of these field individuals remain undetermined. Nevertheless, based on citywide driver statistics, approximately 67 % of drivers in this city are male, and roughly 20 % have accumulated more than ten years of driving experience. This suggests that the Chinese drivers represented in the field data differ from the Chinese test drivers participated in simulation experiments, who typically consist of younger, more educated individuals. Consequently, the two groups of Chinese drivers may exhibit different driving behaviours due to variations in age, experience, and the data collection environments (field vs. simulation).

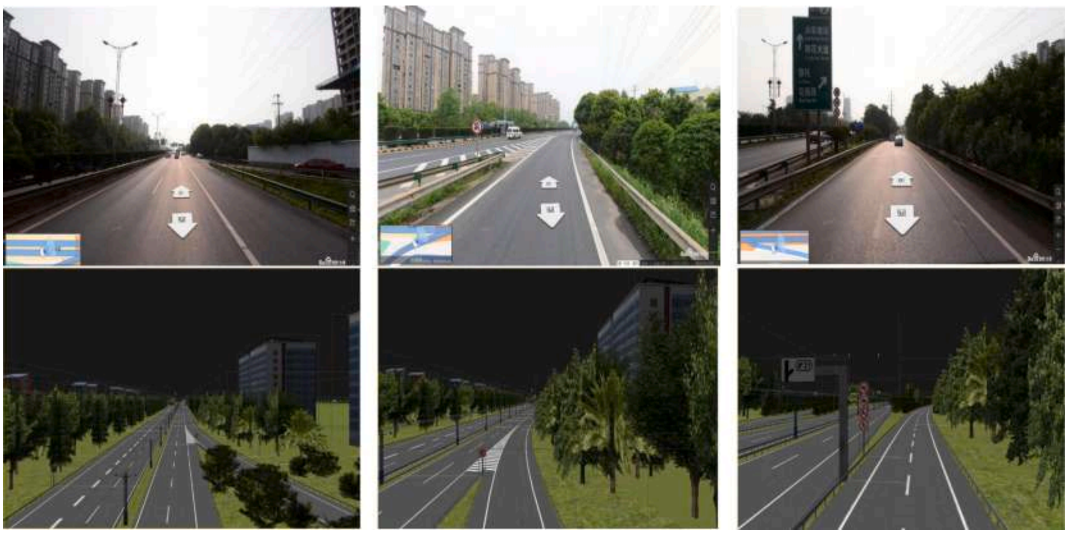


Fig. 4. Comparison between the field satellite imagery and the simulated model.

Table 1
Independent variables of test scenarios.

Variable type	Independent parameter	Level with the codes				
		−1.41	**−1**	**0**	**+1**	**+1.41**
Numeric variable	Speed (km/h)	44.1	47.6	55.9	64.2	67.7
	Age of participants	25	29	40	50	55
Categoric factor	Gender of participants	**Level 1** Male			**Level 2** Female	

Table 2
Speed levels of test scenarios for Italian drivers.

Speed level (km/h)	Age	Number of Italian test drivers
44.1	40	2 (male: 1; female: 1)
47.6	29	2 (male: 1; female: 1)
47.6	50	2 (male: 1; female: 1)
55.9	25	2 (male: 1; female: 1)
55.9	40	10 (male: 5; female: 5)
55.9	55	2 (male: 1; female: 1)
64.2	29	2 (male: 1; female: 1)
64.2	50	2 (male: 1; female: 1)
67.7	40	2 (male: 1; female: 1)

Table 3
Speed levels of test scenarios for Chinese drivers.

Speed level (km/h)	Age	Number of Chinese test drivers
44.1	25	2 (male: 1; female: 1)
47.6	29	2 (male: 1; female: 1)
55.9	25	4 (male: 2; female: 2)
55.9	29	4 (male: 2; female: 2)
64.2	25	1 (male: 1; female: 0)
64.2	29	2 (male: 1; female: 1)
67.7	29	1 (male: 0; female: 1)

Table 4
Demographic characteristics of test drivers.

Nationality	Details	Distributions	Frequencies	Percentages
Italian	Gender	Male	13	50 %
		Female	13	50 %
	Age	(Min. 22, Max. 55, Mean. 39.42, SD = 8.92)		
	Years of holding a driving license	Less than 3 years	0	0 %
		3 years to 10 years	3	11.54 %
		More than 10 years	23	88.46 %
	km drive annually	Less than 1000 km	2	7.69 %
		1000 km to 10,000 km	8	30.77 %
		More than 10,000 km	16	61.54 %
Chinese	Gender	Male	8	50 %
		Female	8	50 %
	Age	(Min. 20, Max. 35, Mean. 26.56, SD = 3.42)		
	Years of holding a driving license	Less than 3 years	2	12.50 %
		3 years to 10 years	13	81.25 %
		More than 10 years	1	6.25 %
	km drive annually	Less than 1000 km	11	68.75 %
		1000 km to 10,000 km	4	25 %
		More than 10,000 km	1	6.25 %

2.2.5. *Experiment protocol and data collection*

The experimental driving simulator protocol is illustrated in Fig. 5. Prior to the main experiment, participants were provided with an introductory session on the fundamental functions and operations of the experimental driving simulator. Subsequently, they were instructed to peruse a document elucidating the meaning of performing various manoeuvres as indicated by arrows in Fig. 6. During the experimental driving session, the participants followed the indicating arrows to perform manoeuvres. Indicating arrows were displayed upon the test driver's entry into the weaving section, ensuring sufficient time for manoeuvre execution. Afterwards, participants engaged in a brief training driving test, wherein they drove for five minutes on a trial track equipped with ramps to enhance their comprehension of indicating arrows and familiarize themselves to the simulator. Following this, participants completed a pre-drive questionnaire designed to collect basic demographic information. Then, participants proceeded with the main experiment by completing the driving path which lasted approximately 20 min. The experimental track and the sequence of these manoeuvres were meticulously designed to ensure efficient data collection while maintaining data accuracy. Each driver traversed the same weaving section 11 times, performing a total of 4 through-manoeuvres, 3 merging manoeuvres, 3 diverging manoeuvres, and 1 ramp-to-ramp manoeuvre. Finally, participants filled out a post-drive questionnaire to assess any potential effects of simulation sickness.

The longitudinal and lateral position, speed, and acceleration of both the subject and all simulated vehicles in the scenarios were recorded at a frequency of 30 Hz in accordance with the field study. Successively, specific manoeuvres performed within the weaving section were extracted from the collected raw data for subsequent analysis.

2.3. Observed variables

To validate lane change behaviour, the following variables were selected and extracted from the field and simulator studies. All variables were extracted for all vehicles in the field video acquisition data as well as for vehicles driven by participants in simulation.

2.3.1. Variables related to traffic efficiency

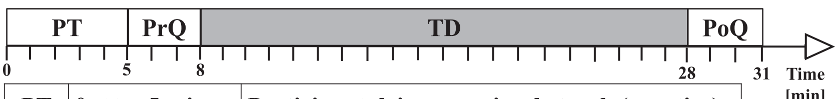
The traffic flow status of the entire weaving section was compared in both field and simulation environments using two selected variables: speed and headway. For both field observations and simulation data, this study extracted the operating speed of vehicles in the direction of travel. Headway is defined as the distance between the front bumpers of two consecutive vehicles divided by the speed of the following vehicle. Additionally, vehicles operating within the weaving sections were categorized into four distinct manoeuvres (through, merging, diverging, and ramp-to-ramp) for comparison in terms of cumulative distribution function.

2.3.2. Variables related to lane change

Based on the different motivations for lane changing, lane-changing behaviours are classified into four categories: (i) discretionary lane change, (ii) mandatory lane change, (iii) merging, and (iv) diverging. These types of lane changes are distinguished by the direction and purpose of the lane number change when vehicles switch lanes. Notably, a discretionary lane change is a spontaneous decision made by drivers without external compulsion, typically aimed at achieving more favourable driving conditions. In contrast, a mandatory lane change is compelled by traffic rules or route requirements, where drivers must change lanes to reach their intended destination. In this study, mandatory lane changing specifically refers to drivers in the leftmost lane moving to the right prior to executing a diverging manoeuvre. It is important to highlight that, given the aim of this study is to examine the microscopic characteristics of lane-changing behaviour, each individual lane change by a vehicle is documented as a single lane-changing event. For instance, if a vehicle changes lanes twice in succession, these are recorded as two distinct lane-changing events. To accurately identify and define lane-changing behaviour, a standardized extraction procedure is proposed in this study. The identification of lane-changing manoeuvres is based on variations in the vehicle's lateral position relative to the lane centreline, specifically using the lane gap metric. As illustrated in Fig. 7, the lane gap is defined as the perpendicular distance between the vehicle's centre of gravity (CoG) and the lane centreline, measured in meters. A positive lane gap indicates that the CoG is located to the left of the lane centreline, whereas a negative value indicates that it is located to the right. Since a lane change constitutes a continuous trajectory over time, three key temporal markers are identified for precise localization: (i) the start, (ii) the insertion, and (iii) end moments. An example of these temporal markers is shown in Fig. 8(a).

The start moment of lane change is indicated by "A" in Fig. 8(a). In accordance with previous studies (Ali et al., 2018; Yang et al., 2019), the lane change start moment was defined as the instance when the lane changer suddenly changes its lateral movement in the original lane. For instance, if a vehicle switches to the right lane, it experiences a continuous reduction in lane gap, whereas switching to the left lane leads to an ongoing increase in lane gap.

The insertion of the target lane is indicated by "B" in Fig. 8(a). This refers to the moment when the vehicle's CoG reached the boundary of the target lane. Mathematically, it is characterized by abrupt changes in the notation of lane gap. For instance, if a vehicle switches to the right lane, there will be a sudden transition from negative to positive values at the insertion point.



PT	$0 \leq t < 5$ min	Participant drive on a simple track (practice)
PrQ	$5 \leq t < 8$ min	Pre-drive questionnaire
TD	$8 \leq t < 28$ min	Test drive
PoQ	$28 \leq t < 31$min	Post-drive questionnaire

Fig. 5. Simulator experiment protocol.

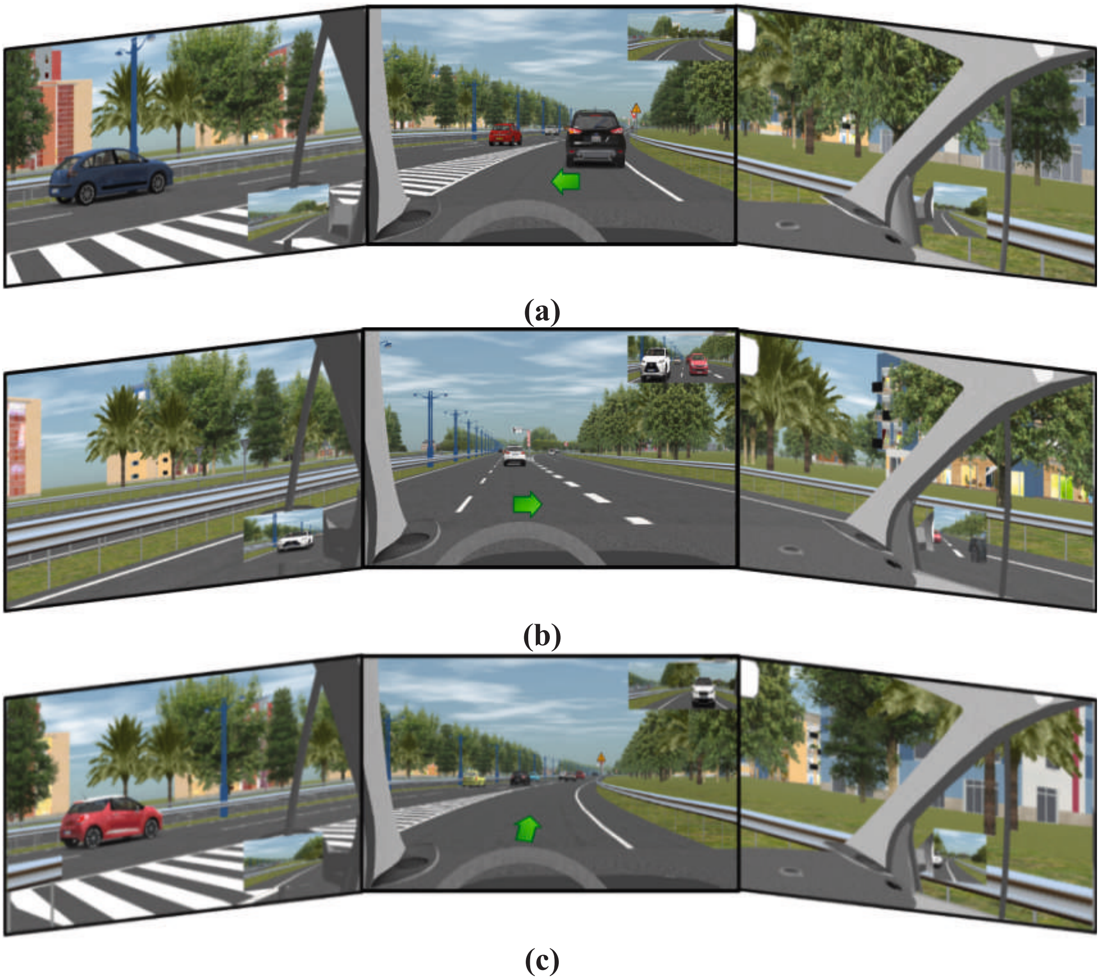


Fig. 6. Frames of three different indicating arrows. (a) merging, (b) diverging, and (c) ramp-to-ramp.

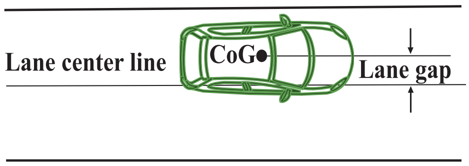


Fig. 7. Lane gap definition.

The end moment of lane change is indicated by "C" in Fig. 8(a). The lane change concluded at point "C" when the movement of the lane changer stabilizes within the target lane.

The Wavelet Transform with a Mexican hat function was employed to identify the start, insertion, and end moments of lane change. Specifically, the lane changer' lane gap was transformed into a wavelet-based curve, as depicted in Fig. 8(b). This approach has demonstrated remarkable efficiency in capturing spatial information in time series data (Zheng et al.,2011; Chen et al., 2014; Ali et al., 2020).

2.3.2.1. Lane change duration. The lane change duration represents the time required for completing the lane change process (Fig. 9), which can be calculated as the difference between the end time and start time of the lane change extracted by wavelet transform, measured in seconds. The effective start time/position refers to the moment when/where the vehicle first crossed the target lane line, while the effective end/position time corresponds to the moment when/where the vehicle's last contact point crossed the target lane line.

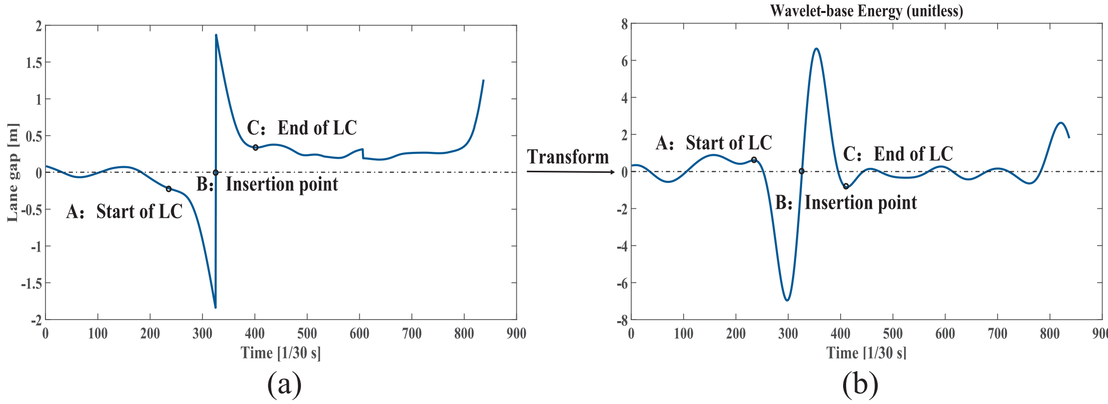


Fig. 8. Identification of three crucial time points during lane change: (a) lane changer's lane gap; (b) transformed curve derived from Wavelet-Transform.

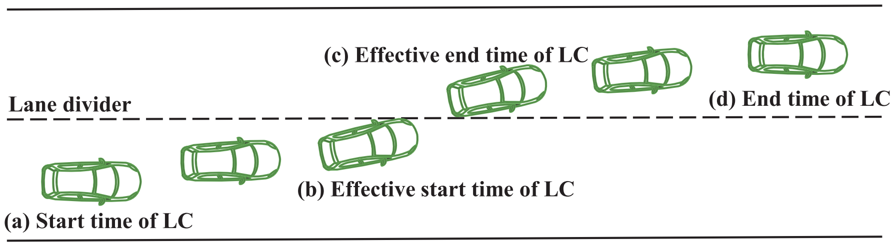


Fig. 9. The process of lane change (LC).

2.3.2.2. Accepted target lane lead and lag gap (AG_TLead and AG_TLag). As depicted in Fig. 10, the accepted lead/lag gap represents the accepted time gap between the lane-changing vehicle and the leading/following vehicle of target lane at start time of lane change, calculated as the distance gap between the lane-changing vehicle and the leading/following vehicle divided by the speed of lane-changing vehicle/following vehicle, measured in seconds. It is worth noting that a positive gap signifies the absence of vehicle body overlap in the direction of travel between the lane-changing vehicle and the leading/following vehicle, whereas a negative gap indicates the presence of such overlap. Negative gaps typically occur in forced lane-change scenarios where the conditions for lane changing are unfavourable.

2.3.2.3. Generalized time-to-collision (GTTC). The widely adopted method for assessing potential crash risk between two consecutive vehicles has been TTC (Meng & Qu, 2012; Sayed et al., 2013; Gu et al., 2019). This approach, however, assumes that vehicles are in the same lane, which may not be applicable in scenarios where collisions occur at different angles. Given this study's focus on evaluating crash risk during lane change, the conventional TTC was extended to incorporate a two-dimensional assessment (Ward et al., 2015; Zhang et al.,2023).

If vehicles are approximated as rectangles, the distance between two vehicles (denoted as $D_{i,j}$) is the shortest distance between any point i on vehicle Veh_i and any point j on vehicle Veh_j, calculated using Equation (1).

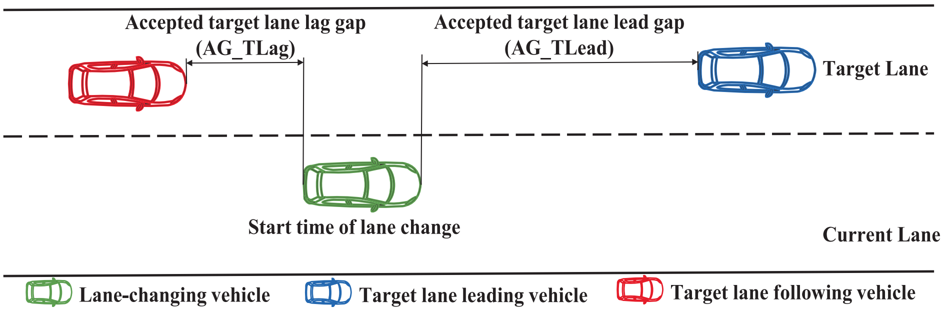


Fig. 10. Definition of accepted target lane lead and lag gap.

$$D_{i,j} = min\{d_{i,j} | i \in Veh_i, j \in Veh_j\} \tag{1}$$

Considering two vehicles with vector positions and velocities, the positions of the points on both vehicles at the given time step are defined as p_i and p_j, while their respective velocities are represented by v_i and v_j. This study assumes a scenario where the vehicles maintain a constant velocity in close proximity to each other, implying zero acceleration. Consequently, the distance between these vehicles (denoted by $D_{i,j}$) can be determined using the following equation:

$$D_{i,j}{}^2 = \|p_i - p_j\|^2 = (p_i - p_j)^T(p_i - p_j) \tag{2}$$

By differentiating both sides of Equation (2), the rate of change for the distance $D_{i,j}$ is derived, referred to as Equation (3):

$$D_{i,j}D_{i,j} = (p_i - p_j)^T(v_i - v_j) \tag{3}$$

Then the distance between Veh_i and Veh_j and its first derivatives can be computed according to Equations (4)-(5):

$$D_{i,j} = \sqrt{(p_i - p_j)^T(p_i - p_j)} \tag{4}$$

$$D_{i,j} = (p_i - p_j)^T(v_i - v_j)/D_{i,j} \tag{5}$$

The closure rate between the two vehicles is assumed to be constant in this study, thus the standard equation of motion (Equations (6) and (7) is employed to solve the *GTTC*:

$$D_{i,j} + D_{i,j} \times GTTC = 0 \tag{6}$$

$$GTTC = -D_{i,j}/D_{i,j} \tag{7}$$

In Equations (6) and (7), *GTTC* needs to be calculated only when $D_{i,j}$ is negative. If $D_{i,j} \geq 0$, the distance between Veh_i and Veh_j would either increase or remain constant, indicating a negligible risk of collision. In these cases, *GTTC* is considered as negative infinity, signifying no future collisions regardless of the time horizon.

2.4. Statistical comparison between field and simulator study

2.4.1. Two-sample Kolmogorov–Smirnov (K-S) test

The significance of the differences between two data populations was investigated using a two-sample K-S test, which compares the empirical cumulative distribution functions of the field and simulator investigations' sample distributions and calculates the maximum distance between these functions. Optimal results are obtained with sufficiently large sample sizes, preferably at least 15 (Kanji, 2006; Karimi et al., 2020).

Let $x; (x_1, x_2, \cdots, x_m)$ and $y; (y_1, y_2, \cdots, y_n)$ be two independent random samples of sizes m and n, respectively, drawn from populations characterized by their respective F and G cumulative distribution functions (CDFs). The null hypothesis (H_0) can be tested against the two-sided alternative hypothesis (H_a) (Berger and Zhou, 2014):

$$H_0 : F(t) = G(t), \ for\ every\ t \tag{8}$$

$$H_a : F(t) \neq G(t), \ for\ at\ least\ one\ value\ of\ t \tag{9}$$

The null hypothesis (H_0) can also be tested against one of the following one-sided alternative hypotheses:

$$H_a : F(t) > G(t), \ for\ all\ values\ of\ t, \ strictly\ greater\ for\ at\ least\ one\ value\ of\ t \tag{10}$$

$$H_a : F(t) < G(t), \ for\ all\ values\ of\ t, \ strictly\ smaller\ for\ at\ least\ one\ value\ of\ t \tag{11}$$

The test statistic for alternative hypothesis of $F(t) \neq G(t)$ was calculated as follows:

$$D = max|F_m(t) - G_n(t)|, \ min(x, y) \leq t \leq max(x, y) \tag{12}$$

where $F_m(t)$ and $G_n(t)$ are empirical CDFs for the samples x and y, respectively.

To test the two alternative hypotheses of $F(t) > G(t)$ and $F(t) < G(t)$ for some value(s) of t, the test statistics were calculated as Equation (13) and (14), respectively:

$$D^+ = max[F_m(t) - G_n(t)], \ min(x, y) \leq t \leq max(x, y) \tag{13}$$

$$D^- = max[G_n(t) - F_m(t)], \ min(x, y) \leq t \leq max(x, y) \tag{14}$$

To test H_0 at the significance level of 0.01, H_0 could be rejected in favour of H_a if the calculated p-value is lower than 0.01.

2.4.2. Two-sample Kolmogorov–Smirnov N(K-S) test after probability sampling

As the sample size grows, the K-S test's statistical power increases, making it easier to detect even small differences. Consequently, this leads to a higher likelihood of obtaining statistically significant results for very minor differences, thereby increasing the false positive rate (Drezner et al.,2010; Nguyen, 2017). In order to address the issue, some researchers have proposed a method involving re-sampling followed by the K-S test (Olea and Pawlowsky-Glahn, 2009; Wang et al., 2011). Referring to the methods of previous researches, this paper employed a K-S two-sample test method based on probability sampling. Continuous variables were segmented into multiple intervals according to their distribution ranges, and the probability density for each interval was computed, as presented in Equation (15). Finally, as illustrated in Equation (16), the required number of samples for each interval was determined by multiplying the total number of samples by the corresponding probability distribution:

$$p_i = \frac{m_i}{N} \tag{15}$$

$$n_i = n \times p_i \tag{16}$$

where, i represents the index of the data interval, p_i refers to the distribution probability of the i^{th} data interval, m_i refers to the amount of data in the i^{th} data interval, N represents the number of the population, n_i represents the number of samples in the i^{th} data interval, and n represents the total number of samples.

3. RESULTS

The results section provides the descriptive statistics of observed variables, compares the differences in continuous variables before and after probability sampling, and presents the statistical analysis of observed variables.

3.1. Descriptive statistics of observed variables

The observed variables from field and simulator are summarized in Table 5. The continuous variables (i.e., speed, headway and *GTTC*) in Table 5 were recorded at a frequency of 30 Hz. It is important to note that the speed, headway, and GTTC collected in the field represent the metrics of all vehicles—both crossing and lane-changing vehicles—across the entire weaving section. However, the

Table 5
Descriptive statistics of observed variables in field and driving simulator.

Group	Variable	#	Mean	SD	Min.	Max.
Field (Chinese)	Speed (km/h)	921,076	55.90	14.43	0.41	119.97
	Headway (s)	814,864	2.87	2.48	0.02	20.00
	Lane change duration (s)	1,660	6.40	3.64	0.50	20.3
	AG_TLead (s)	995	3.37	7.16	−2.24	116.63
	AG_TLag (s)	1,128	1.38	1.62	−4.57	14.82
	GTTC (s)	322,515	19.39	18.58	0.10	100.00
Simulator (All Italian)	Speed (km/h)	229,086	55.73	9.21	9.99	89.52
	Headway (s)	228,799	3.13	2.39	0.06	19.99
	Lane change duration (s)	246	6.01	1.97	0.96	15.42
	AG_TLead (s)	223	7.26	20.64	−0.33	194.71
	AG_TLag (s)	202	7.33	15.86	−37.77	114.20
	GTTC (s)	57,768	28.23	22.41	0.10	99.98
Simulator (Chinese < 40 years old)	Speed (km/h)	138,573	54.37	11.73	7.57	94.48
	Headway (s)	137,276	4.48	3.44	0.03	20.00
	Lane change duration (s)	153	5.15	1.99	1.29	12.96
	AG_TLead (s)	120	8.99	21.71	−0.43	178.51
	AG_TLag (s)	109	7.23	18.46	−23.83	160.24
	GTTC (s)	29,279	21.93	19.85	0.11	99.94
Simulator (Italian < 40 years old)	Speed (km/h)	105,732	54.92	9.49	9.99	89.52
	Headway (s)	105,634	2.76	2.15	0.06	19.99
	Lane change duration (s)	112	6.01	1.74	2.08	14.66
	AG_TLead (s)	100	6.41	19.90	−0.11	194.71
	AG_TLag (s)	86	6.13	13.04	−37.77	77.83
	GTTC (s)	24,900	28.17	22.35	0.47	99.92
Simulator (Italian ≥ 40 years old)	Speed (km/h)	123,354	56.42	8.91	19.10	85.35
	Headway (s)	123,165	3.45	2.54	0.13	19.92
	Lane change duration (s)	134	6.01	2.15	0.96	15.42
	AG_TLead (s)	123	7.95	21.29	−0.33	180.31
	AG_TLag (s)	116	8.22	17.67	−19.93	114.20
	GTTC (s)	32,868	28.27	22.46	0.10	99.98

Note: #: the number of data; AG_TLead: accepted target lane lead gap; AG_TLag: accepted target lane lag gap; *GTTC*: *GTTC* between the lane-changing vehicle and the leading and following vehicles on both the current and target lanes during a lane change.

simulation data reflect the driving behaviour of the subject vehicle operated by the participants throughout the weaving section. For lane change duration and accepted gap respectively signify distinct variables associated with each lane change performed. Additionally, outliers falling outside the predefined threshold range were excluded to facilitate subsequent probability sampling and enable easier comparison of distribution characteristics. The threshold range was determined based on both the practical significance and percentile distributions of the variables. Specifically, data points with speeds exceeding 120 km/h, headways greater than 20 s, accepted gaps less than −50 s or greater than 200 s, and GTTC values exceeding 100 s were removed. These outliers constitute approximately 5 % of the total dataset and were likely attributable to errors in video extraction trajectory coordinates or calculations.

In simulation study, there are 399 lane change manoeuvres that were recorded. Specifically, Chinese conducted a total of 153 lane changes while Italians conducted 246 (112 by those younger than 40 and 134 by those older than 40). Simulation data is categorized into three groups: (i) Chinese drivers under 40, (ii) Italian drivers under 40, and (iii) Italian 40 and over. The impact of age on driving behaviour was assessed by comparing younger and older groups. Additionally, a comparative analysis between Chinese and Italian was conducted to enhance the generalizability and robustness of the findings. A preliminary analysis of the overall simulation sickness score after driving (mean = 22.57, range = 11.11–100) suggests that the average level of simulation sickness experienced by participants was relatively low. This suggests that the impact of simulation sickness on the experimental outcomes is relatively minor.

In the field study, a total of 1,660 lane changes (including 658 discretionary lane changes, 343 mandatory lane changes, 102 merging manoeuvres, and 557 diverging manoeuvres) were observed. Descriptive statistical parameters of headway show no significant differences between the field-collected data and the simulation data. The observed speed in the field, although similar in average value to the simulated data, exhibits a wider range between the minimum and maximum values. The lane change duration observed in the field is comparable to the average value from simulations; however, the field data demonstrates a higher standard deviation and maximum value.

These variations might be due to the limited range of traffic flow conditions that can be simulated. Simulation experiments primarily represent typical field scenarios (i.e., the 15th to the 85th percentile of speed) but do not account for extreme low- or high-speed conditions. For AG_TLead, AG_TLag, and GTTC, field-collected data exhibit lower average values and smaller standard deviations compared to simulation data. Notably, the simulation data shows significantly higher maximum values for AG_TLead and AG_TLag, with AG_TLag also displaying a smaller negative minimum value. These differences could be influenced by the driving perspective during simulated lane changes. In the simulated environment, drivers may not be fully accustomed to the virtual setting, leading to difficulties in accurately judging distances—especially the distance to vehicles behind—based solely on rear-view mirrors. As a result, more forced lane changes (characterized by negative AG_TLag) and hesitant lane changes occur, which necessitate larger lane change gaps.

From the descriptive statistical information, it is evident that while the observed variables exhibit some similarity between field and simulation data, notable differences also exist. Therefore, further statistical analysis is necessary to explore these variations in greater depth.

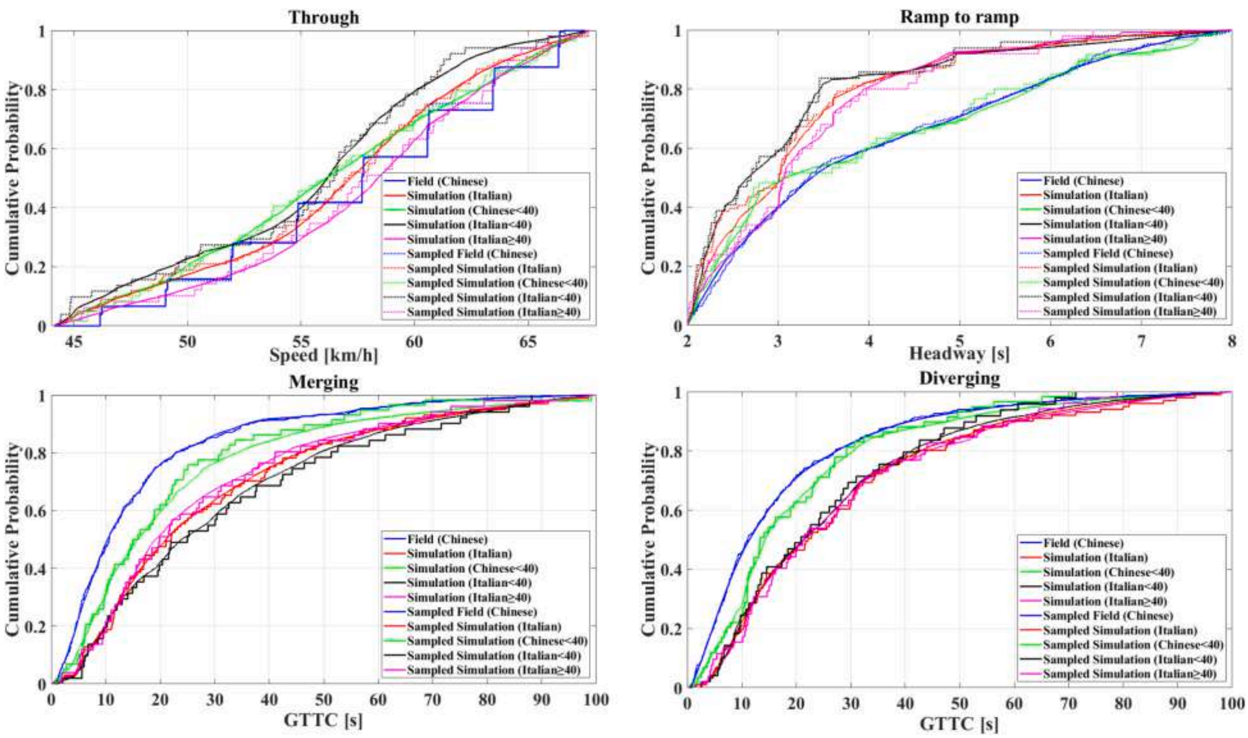


Fig. 11. Comparison of continuous variable between sample and population.

3.2. Before and after probability sampling

Probability sampling was implemented for observed continuous variables (i.e., speed, headway and *GTTC*). The proportional relationship among the grouped data was considered. A sample size of 300 was applied to field data. For Italians' data in simulation, a sample size of 100 was used (including 50 Italians < 40 years old and 50 Italians ≥ 40 years old). Furthermore, a sample size of 60 was designated for Chinese data in simulation. It is important to highlight that probability sampling was conducted based on distinct manoeuvres, considering the varying distributions of variables across different manoeuvres. To ensure the accurate representation of the distribution characteristics of the entire population, various statistical tests including χ^2 test, F-test, and *t*-test were conducted. The results indicated a high level of consistency (over 95 %) between the sampled data and population data in terms of their distribution characteristics. The cumulative distribution probability graphs for sampled and population data are illustrated in Fig. 11. The solid line represents the population data, while the dashed line represents the sampled data. It is evident that the sampled data effectively captures the distribution characteristics of the population data.

The results provide compelling evidence that the employed probability sampling method ensured a robust representation of the population dataset's distribution characteristics.

3.3. Statistical comparison of variables related to traffic efficiency

Observed variables related to traffic efficiency resulting from field and simulator investigations were assessed and compared. Four specific manoeuvres — through, merging, diverging, and ramp-to-ramp — were individually analysed.

3.3.1. Speed

Since only the 15th to 85th percentiles of the speed distribution are considered when designing the experiment, the K-S test was consequently performed only on the distribution within this interval. It is worth noting that Figs. 11–12 indicate that the cumulative probability distribution of speed derived from field data exhibits a step-wise pattern. This characteristic may result from the discretization of speed during the data acquisition and processing procedure. Furthermore, Figs. 11–12 display only a limited speed range (44.1 to 67.7 km/h), the visual appearance of discretization is amplified. The results in Table 6 indicate that, for the through manoeuvre, Italians, Chinese and old Italians exhibit similar speeds to field. Conversely, the speeds of young Italians differed significantly from field. Therefore, a one-sided alternative hypothesis was tested against the null hypothesis. The results demonstrate that the speed distribution in field was significantly lower than of young Italians ($D^{-} = 0.2502$, $p = .0034$). For merging manoeuvre, the results of two-sided K-S test indicated no significant difference in the cumulative probability distribution of speed across different pairs. This demonstrates that the speed distribution of field and simulation exhibited a high degree of similarity during merging

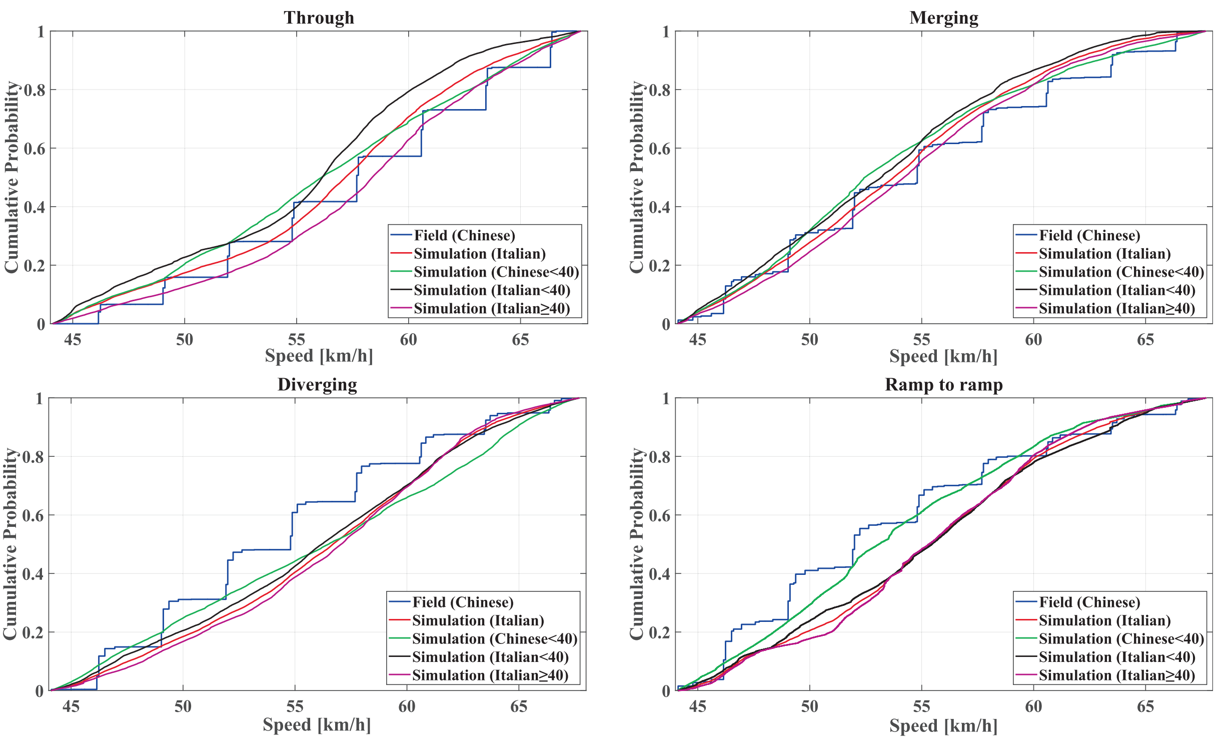


Fig. 12. Comparison of cumulative probability for speed (44.1–67.7 km/h).

manoeuvre. Nevertheless, it is important to note that despite the lack of significant differences in cumulative probability distribution, Fig. 13 reveals notable differences in the spatial distribution of average speed. During both through and merging manoeuvres, the average speed of the field data was consistently higher than that of the simulation data in weaving sections 2 and 3, with the peak difference occurring at the end of weaving section 2. This disparity can primarily be attributed to the differences in available headways between field and simulation vehicles, as the upper limit of headways in the simulation was truncated at 2–8 s. In contrast, in real-world scenarios, drivers had greater flexibility to adjust their acceleration profiles, enabling them to achieve higher speeds. Additionally, the average speed of the simulation data fell within the experimental design range (44.1–67.7 km/h). Since some field data values exceeded this range, this likely explains the observed differences in the spatial distribution of speeds during through and merging manoeuvres. Regarding the diverging manoeuvre, the cumulative probability distribution of speed for Italians and old Italians from the simulation study was significantly different from that of field study. The results of one-sided alternative hypothesis test indicate that the cumulative probability distribution of speed in the field was greater than those in the simulation. Nonetheless, no significant difference was observed between Chinese, young Italians, and field. The results for the ramp-to-ramp manoeuvre show that only Chinese had a statistically similar cumulative probability distribution of speed to the field data. The outcomes of one-sided alternative hypothesis test also indicate that the cumulative probability distribution of speed in the field was greater than those of Italians, young Italians, and old Italians.

Besides, as depicted in Fig. 13, for diverging and ramp-to-ramp manoeuvres, the average speed of field was lower than that of simulation. This discrepancy can be attributed to the deceleration typically required when entering off-ramps, compounded by the static nature of the physical motion simulation environment in simulators, which may hinder drivers' ability to accurately perceive higher speed during off-ramp manoeuvres.

3.3.2. Headway

Since only the headway distribution were truncated at 2–8 s when designing the experiment (Fig. 14), the K-S test was consequently performed only on the distribution within this interval. The findings in Table 7 suggest that, for the through manoeuvre, only cumulative probability distribution of Chinese was significantly different with field ($D = 0.2432$, $p = .0039$). Additionally, the results from the one-sided alternative hypothesis test confirm that the headway cumulative probability distribution of field was significantly higher than Chinese. Regarding the merging manoeuvre, both Italians and old Italians exhibited significant differences compared to field. A one-sided alternative hypothesis test confirms that headway for Italians and old Italians was significantly smaller than that of field. Furthermore, as depicted in Fig. 15, the relationship between the average headway of field and that of simulation varied across the weaving section, with no significant difference observed in section 2. In section 1, field's average headway was slightly larger, while at the end of section 3, it was significantly smaller. Concerning the diverging manoeuvre, there is no significant difference between field and simulation. During ramp-to-ramp manoeuvre, Italians and young Italians exhibited significantly reduced headways compared to field. Meanwhile, as depicted in Fig. 15, the average headway of field progressively decreases, closely mirroring the characteristics of merging manoeuvre. It is evident that the average headway of field exhibited significant fluctuations during merging and ramp-to-ramp manoeuvres. This phenomenon can likely be attributed to the higher speeds and increased traffic volumes observed in sections 2 and 3 downstream of the on-ramp, as vehicles needed to accelerate to merge into the main road. Additionally, frequent lane changes exacerbated traffic congestion downstream.

Nevertheless, due to the limited number of simulated vehicles, maintaining a continuous traffic flow in the downstream weaving section in simulation proves challenging. Additionally, as shown in Fig. 15, it is noteworthy that Chinese exhibited relatively high average headways during all manoeuvres, which may be attributed to their driving styles.

3.4. Statistical comparison of variables related to lane change

Observed variables for four types of lane changes, including discretionary, mandatory, merging, and diverging, were evaluated and compared separately.

3.4.1. Lane change duration

As depicted in Table 8, for discretionary lane change, only young Italians showed no significant difference from field. In contrast,

Table 6
The results of two-sided K-S test for speed (44.1–67.7 km/h).

Manoeuvre	Comparison Group Pairs Field (Chinese) vs. Simulation (Italian)	Field (Chinese) vs. Simulation (Chinese < 40)	Field (Chinese) vs. Simulation (Italian < 40)	Field (Chinese) vs. Simulation (Italian ≥ 40)
Through	$D = 0.1767$ $p = .0160$	$D = 0.1563$ $p = .1638$	$D = 0.2502$ $p = .0068$	$D = 0.1139$ $p = .6193$
Merging	$D = 0.1564$ $p = .0467$	$D = 0.1381$ $p = .2675$	$D = 0.1716$ $p = .1440$	$D = 0.1287$ $p = .4497$
Diverging	$D = 0.2051$ $p = .0030$	$D = 0.1905$ $p = .0461$	$D = 0.1851$ $p = .0935$	$D = 0.2716$ $p = .0024$
Ramp to ramp	$D = 0.2418$ $p < .0001$	$D = 0.1507$ $p = 01890$	$D = 0.2452$ $p = .0093$	$D = 0.2711$ $p = .0025$

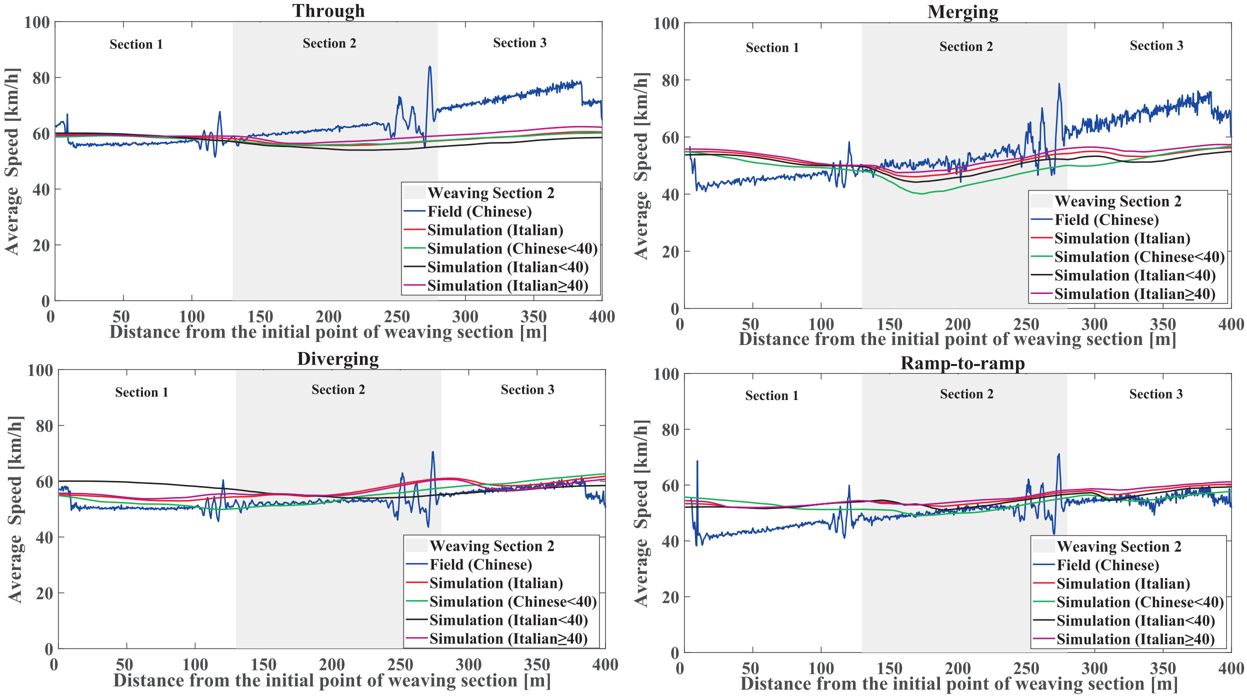


Fig. 13. Comparison of spatial distribution for average speed.

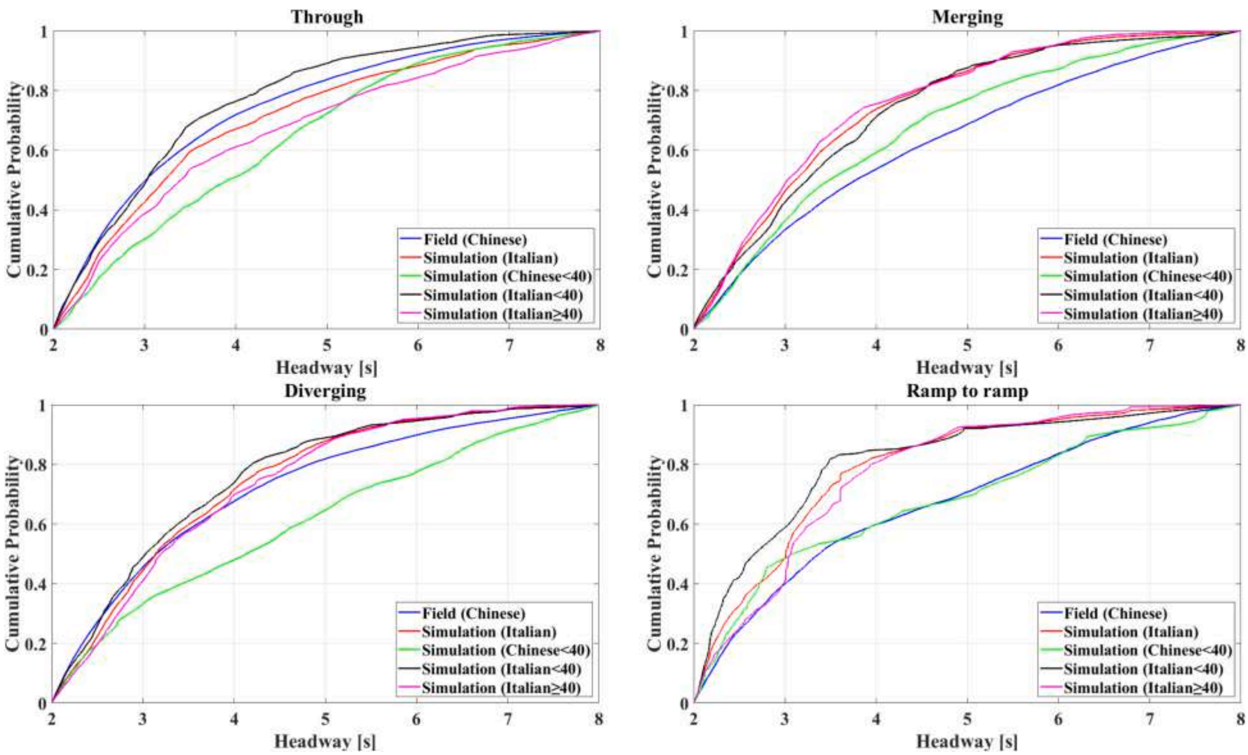


Fig. 14. Comparison of cumulative probability for headway (2–8 s).

Table 7
The results of two-sided K-S test for headway.

Manoeuvre	Comparison Group Pairs Field (Chinese) vs. Simulation (Italian)	Field (Chinese) vs. Simulation (Chinese < 40)	Field (Chinese) vs. Simulation (Italian < 40)	Field (Chinese) vs. Simulation (Italian ≥ 40)
Through	$D = 0.0900$	$D = 0.2432$	$D = 0.1100$	$D = 0.1300$
	$p = .5583$	$p = .0039$	$p = .6522$	$p = .4592$
Merging	$D = 0.2167$	$D = 0.1405$	$D = 0.2209$	$D = 0.2614$
	$p = .0014$	$p = .2499$	$p = .0275$	$p = .0050$
Diverging	$D = 0.1083$	$D = 0.2140$	$D = 0.1112$	$D = 0.0943$
	$p = .3286$	$p = .0171$	$p = .6380$	$p = .8136$
Ramp to ramp	$D = 0.2338$	$D = 0.1444$	$D = 0.3052$	$D = 0.2223$
	$p < .0001$	$p = .2286$	$p < .0001$	$p = .0241$

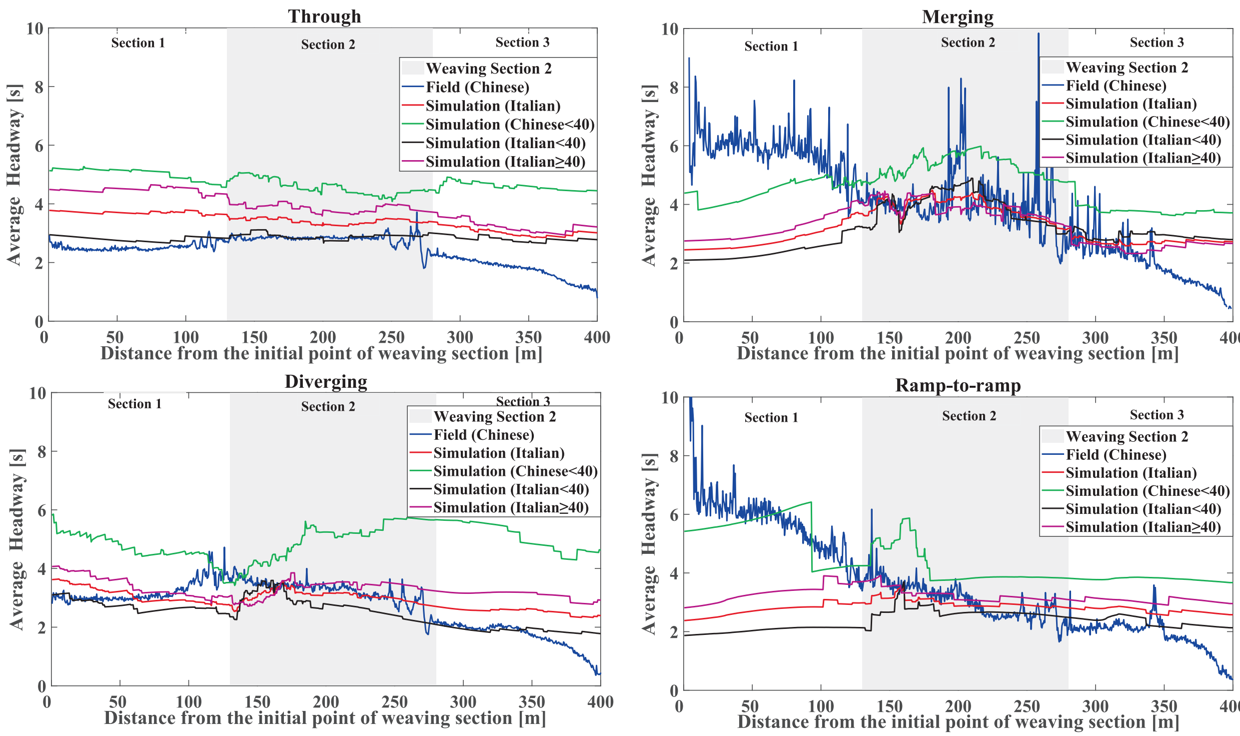


Fig. 15. Comparison of average headway spatial distribution.

other groups exhibited significantly different distributions compared to field ($p < .0001$). One-sided alternative hypothesis tests revealed that field distribution was significantly lower than that of these three groups. Furthermore, as depicted in Fig. 16, the discrepancy in distribution between simulation and field primarily resides in the right tail, where field data exhibit a higher frequency

Table 8
The results of two-sided K-S test for lane change duration.

LC[a] Type	Comparison Group Pairs Field (Chinese) vs. Simulation (Italian)	Field (Chinese) vs. Simulation (Chinese < 40)	Field (Chinese) vs. Simulation (Italian < 40)	Field (Chinese) vs. Simulation (Italian ≥ 40)
DLC[b]	$D = 0.3222$	$D = 0.4394$	$D = 0.2999$	$D = 0.4153$
	$p < .0001$	$p < .0001$	$p = .0203$	$p < .0001$
MLC[c]	$D = 0.2030$	$D = 0.3591$	$D = 0.1852$	$D = 0.2618$
	$p = .1041$	$p = .0022$	$p = .6612$	$p = .0866$
Merging	$D = 0.4917$	$D = 0.4654$	$D = 0.4690$	$D = 0.5210$
	$p < .0001$	$p < .0001$	$p < .0001$	$p < .0001$
Diverging	$D = 0.3765$	$D = 0.4093$	$D = 0.4937$	$D = 0.2776$
	$p < .0001$	$p < .0001$	$p < .0001$	$p = .0037$

Note: [a]: lane change; [b]: discretionary lane change; [c]: mandatory lane change.

of longer lane change durations. For mandatory lane change, Italian and field exhibited a high degree of similarity. On the other hand, Chinese showed notable differences from field. The results of the one-sided alternative hypothesis test indicate that the distribution of field is significantly lower than that of Chinese ($D^{-} = 0.3591$, $p = .0011$), suggesting that Chinese tended to have shorter lane change durations during mandatory lane change. For the merging manoeuvre, the results indicate that there are statistically significant differences in lane change duration distribution between simulation and field ($p < .0001$). The one-sided alternative hypothesis test reveals that the field distribution was significantly greater than that of simulation. As depicted in Fig. 16, this discrepancy is primarily observed on the left side, indicating that field exhibited more frequent merging manoeuvres with shorter lane change durations. As depicted in Fig. 17, the cumulative probability distribution of distance between on-ramp and merge effectively start position in field was significantly smaller compared to simulation. This analysis reveals that simulation drivers initiated merging at a significantly shorter distance from the on-ramp compared to field, potentially leading to inadequate acceleration. As illustrated in Fig. 18, there is minimal difference between field and simulation at the effectively merge end position. This suggests that simulation may require a longer lane change distance for merging, which could account for the extended duration observed. In contrast, field vehicles merging from a greater distance from the on-ramp, allowing for adequate acceleration, thereby enabling smoother integration into the main road. For diverging manoeuvres, the lane change duration distribution from simulation and field exhibited significant differences. The results of the one-sided alternative hypothesis test indicate that field lane change duration distribution was significantly lower than that of simulation. As depicted in Fig. 16, the field data showed a relatively uniform distribution. In contrast, simulation data was predominantly concentrated between 5 and 8 s. Additionally, as depicted in Fig. 17, the cumulative probability distribution of distance between the on-ramp and diverge effectively start position in field was significantly smaller than that in simulation. This suggests that field drivers tended to travel longer distances on the main road before changing lane, leading to more interactions with weaving traffic and consequently lower speeds compared to simulation (Figs. 12-13). This may explain for a longer lane change duration in field diverging scenarios.

3.4.2. *Accepted target lane lead and lag gap*

3.4.2.1. AG_TLead. The results in Table 9 show that AG_TLead distribution for discretionary lane changes differs significantly between field data and simulation. The one-sided alternative hypothesis test confirms the field distribution was significantly higher. Field data shows that AG_TLead was mostly within 0 to 5 s, while simulations had a more even distribution. This suggests that during discretionary lane changes, field drivers preferred smaller gaps with the target lane's leading vehicle. For mandatory lane change, the AG_TLead distribution in field data and simulations was nearly identical, indicating similar lead gaps with the target lane's leading vehicle during the manoeuvre. As shown in Fig. 19, the differences between simulation data and field data for manoeuvres are relatively small, except in the case of discretionary lane changes. Discretionary lane changes represent spontaneous decisions made by drivers, whereas the other three manoeuvres are mandatory, allowing drivers limited flexibility in selecting accepted gaps. As a result,

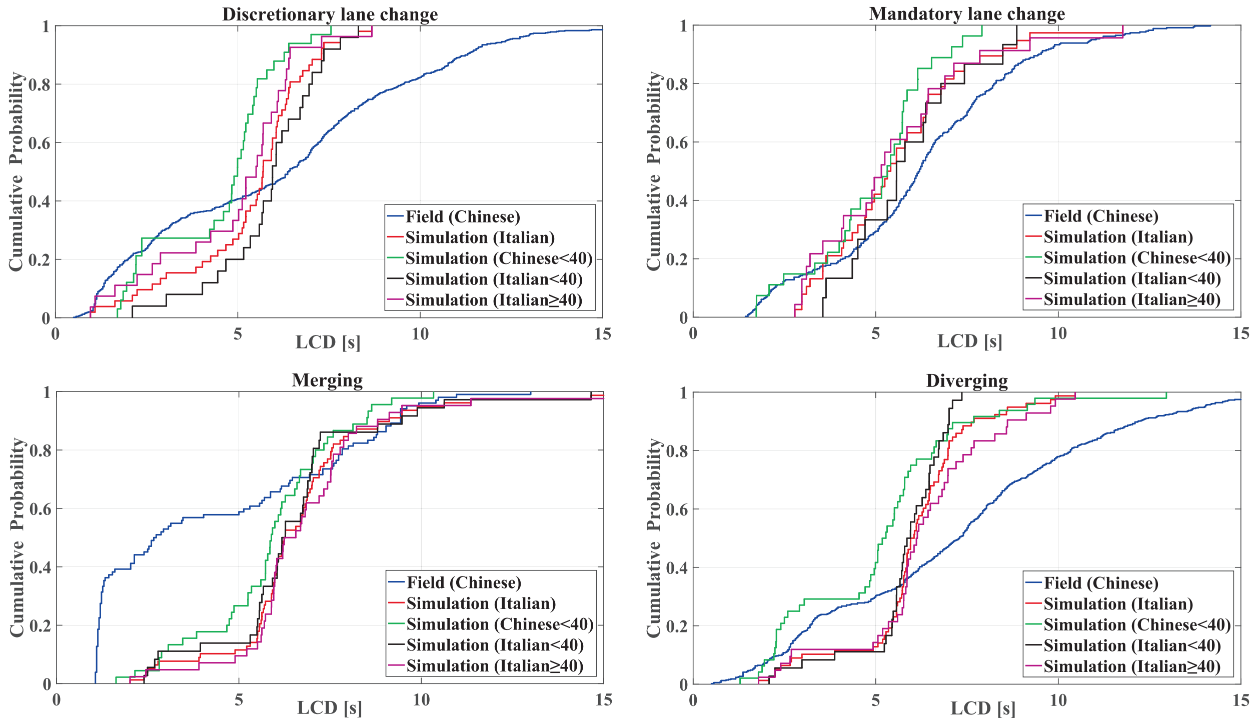


Fig. 16. Comparison of cumulative probability for lane change duration.

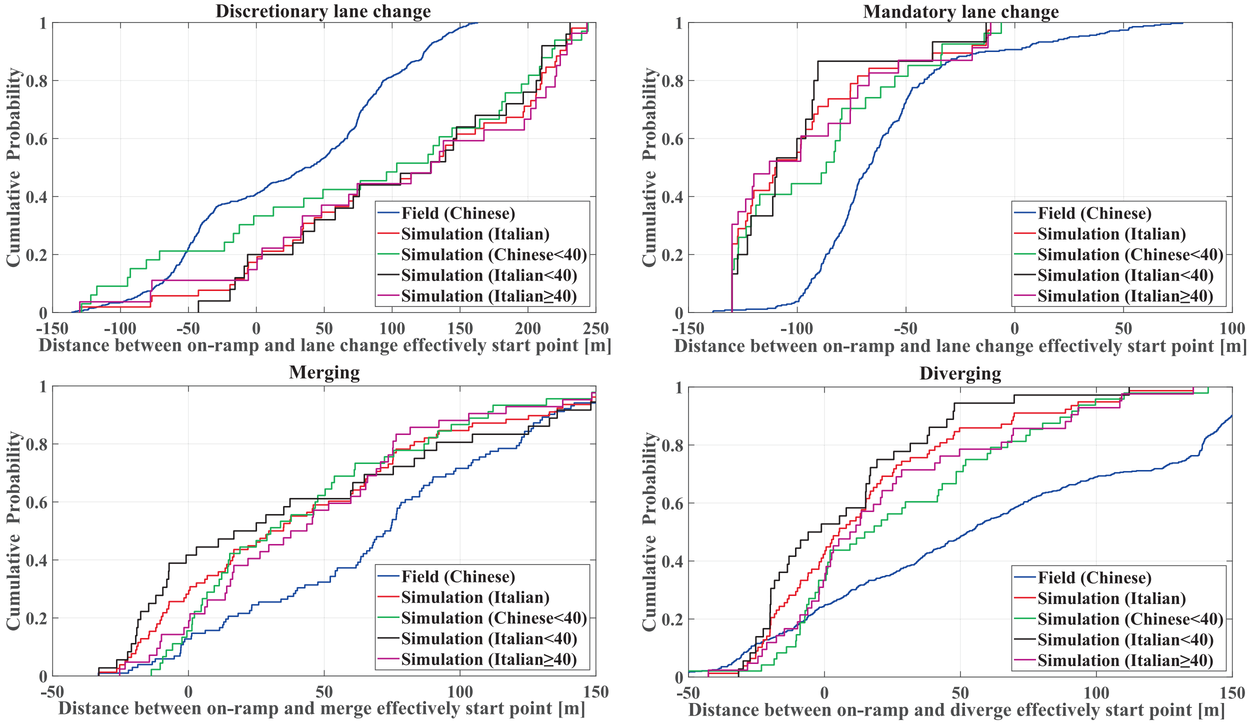


Fig. 17. Comparison of cumulative probability for distance between on-ramp and lane change effectively start position (positive values indicate lane change starts ahead the on-ramp; negative values indicate otherwise).

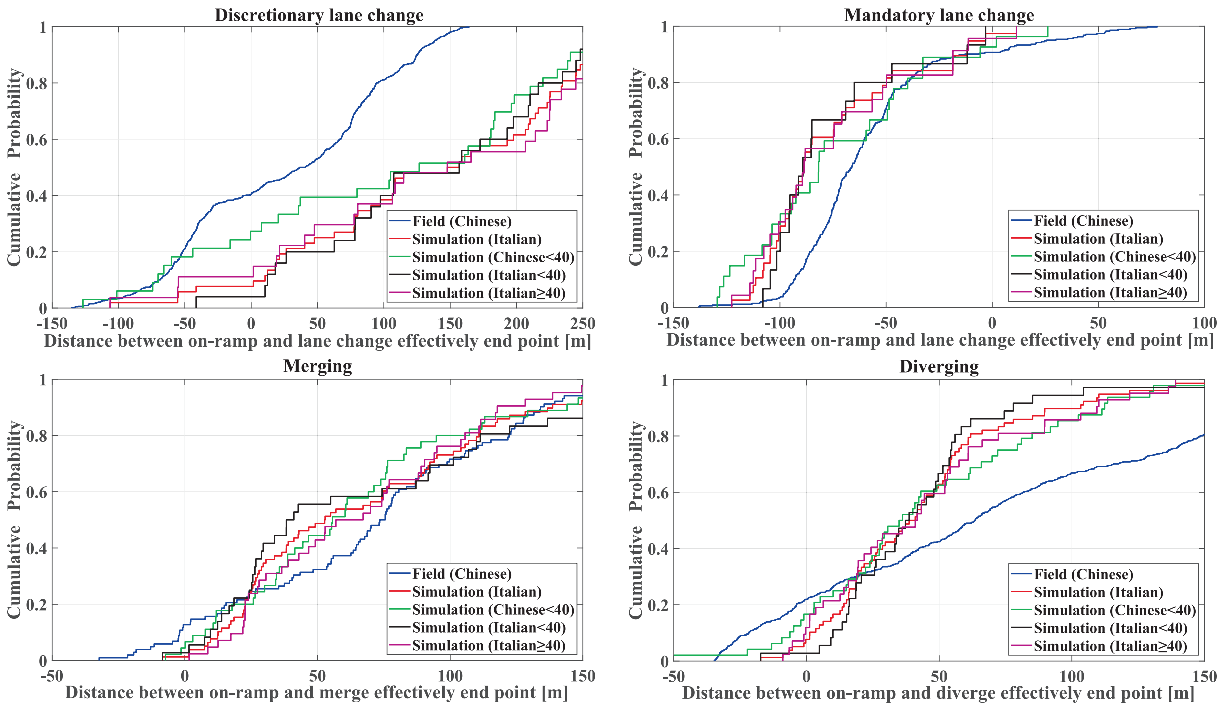


Fig. 18. Comparison of cumulative probability for distance between on-ramp and lane change effectively end position (positive values indicate lane change ends ahead the on-ramp; negative values indicate otherwise).

Table 9
The results of two-sided K-S test for AG_TLead.

LC[a] Type	Comparison Group Pairs			
	Field (Chinese) vs. Simulation (Italian)	Field (Chinese) vs. Simulation (Chinese < 40)	Field (Chinese) vs. Simulation (Italian < 40)	Field (Chinese) vs. Simulation (Italian ≥ 40)
DLC[b]	$D = 0.5265$ $p < .0001$	$D = 0.5064$ $p < .0001$	$D = 0.5094$ $p < .0001$	$D = 0.5483$ $p < .0001$
MLC[c]	$D = 0.1613$ $p = .3648$	$D = 0.2182$ $p = .2565$	$D = 0.1452$ $p = .9117$	$D = 0.1819$ $p = .4736$
Merging	$D = 0.2070$ $p = .0809$	$D = 0.4430$ $p < .0001$	$D = 0.2991$ $p = .0350$	$D = 0.2606$ $p = .0490$
Diverging	$D = 0.3096$ $p < .0001$	$D = 0.3068$ $p = .0039$	$D = 0.3540$ $p < .0001$	$D = 0.2820$ $p = .0079$

Note: [a]: lane change; [b]: discretionary lane change; [c]: mandatory lane change.

in the simulation experiments, the accepted gaps for mandatory manoeuvres were significantly smaller than those observed for discretionary lane changes. Regarding merging manoeuvres, there was no significant difference in the AG_TLead distribution among Italians, young Italians, old Italians, and field. Nevertheless, significant differences were observed between field and Chinese. The results of one-sided alternative hypothesis test indicated that the AG_TLead distribution for Chinese was significantly smaller compared to field, suggesting that Chinese accepted a larger gap with the leading vehicle of target lane during merging manoeuvres in simulation environment. A comparison of AG_TLead distribution in diverging manoeuvres shows that simulation significantly differed from field. The results of one-sided alternative hypothesis test show that, the AG_TLead distribution in field was significantly higher than that of simulation, suggesting that simulation drivers accepted a larger lead gap.

3.4.2.2. AG_TLag. As depicted in Table 10, in discretionary lane change, only Chinese and young Italians exhibited no significant differences in AG_TLag distribution compared to field, while Italians and old Italians showed significant differences. The one-sided alternative hypothesis tests indicate that field distribution was significantly larger than that of two groups, suggesting that Italian and older Italians accepted a larger lag gap. For mandatory lane change, the one-sided alternative hypothesis test results show that field distribution was significantly larger, indicating that the AG_TLag values in simulation tended to be larger than those in field. Additionally, it is observed that the proportion of AG_TLag values exceeding 15 s in simulation was higher compared to other manoeuvres. For the merging scenario, the results of one-sided alternative hypothesis test indicate that the AG_TLag distribution of field was significantly lower than that of Italian and old Italians, but significantly higher than that of Chinese and young Italians. This indicates that during the merging manoeuvre, the accepted lag gaps in ascending order were: Italian and old Italians, followed by those based on field data, and finally Chinese and young Italians. Given that all Chinese participants in simulation were under 40, this implies that young drivers may accept larger lag gaps when merging. Regarding diverging, the two-sided K-S test results indicate that the AG_TLag distribution of simulation significantly differed from that of field data. The significant differences are primarily concentrated on the right side, as illustrated in Fig. 20.

Specifically, AG_TLag values for field predominantly fell within the 0 to 5 s, whereas simulation exhibited a notable proportion of larger AG_TLag values.

3.4.3. GTTC

The two-sided K-S test results in Table 11 indicate no significant difference in *GTTC* between field and simulation during discretionary lane change (see Fig. 21). Notably, the difference between Chinese participants and field data was minimal ($D = 0.0925$, $p = .7952$). In contrast, during mandatory lane change, one-sided alternative hypothesis test results show that Italians and old Italians had greater *GTTC* values than field. Still, there were no significant differences in *GTTC* distribution between Chinese and young Italians compared to field, with the difference between Chinese and field data being particularly small ($D = 0.0672$, $p = .9783$). Given that all Chinese participants were under 40, this suggests that younger drivers in simulation may exhibit smaller *GTTC* values more closely resembling those in field when performing mandatory lane change. Regarding merging manoeuvres, there was no significant difference in *GTTC* distribution between Chinese and field. Conversely, the other three groups (Italians, young Italians, and old Italians) exhibited significant differences from field. One-sided alternative hypothesis test results indicate that these groups had larger *GTTC* values than field. For the diverging scenarios, the two-sided K-S test results showed no significant difference in *GTTC* distribution between the Chinese and field data ($D = 0.1862, p = .0413$). The one-sided alternative hypothesis test results demonstrate that, similar to merging scenarios, the three sets of simulation data exhibited higher *GTTC* values than the field data during diverging.

3.5. Relationship among GTTC, lane change duration and accepted gap

From the previous analysis, it can be inferred that *GTTC* may correlate with observed variables during lane change manoeuvres. Mwesige et al. (2016) and Karimi et al. (2021) demonstrated that TTC decreases as passing duration increases and effective accepted gap decreases. They also established a significant relationship between TTC and the ratio of lane change duration to the effective accepted gap. To investigate the potential relationships between observed variables of this study and *GTTC*, further data processing was conducted. Specifically, the minimum *GTTC* between the ego vehicle and the leading and following vehicles of target lane during

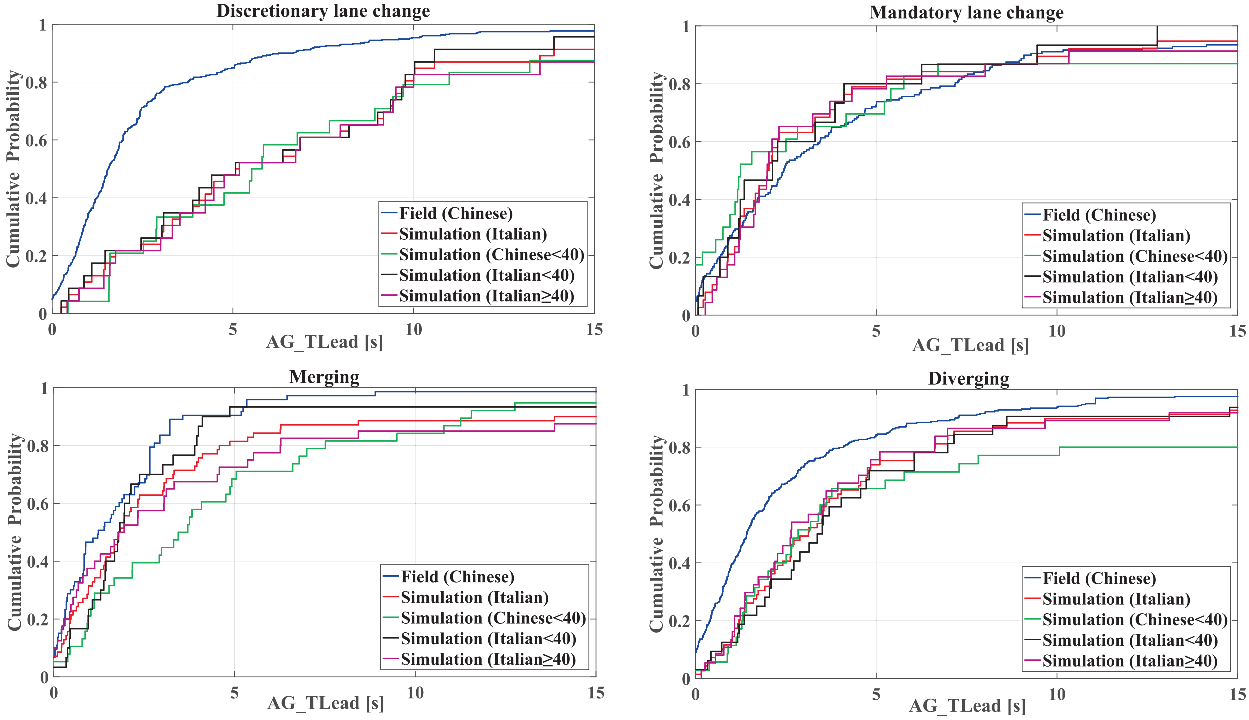


Fig. 19. Comparison of the cumulative probability of AG_TLead.

Table 10
The results of two-sided K-S test for AG_TLag.

LC[a] Type	Comparison Group Pairs Field (Chinese) vs. Simulation (Italian)	Field (Chinese) vs. Simulation (Chinese $<$ 40)	Field (Chinese) vs. Simulation (Italian $<$ 40)	Field (Chinese) vs. Simulation (Italian $\geq$ 40)
DLC[b]	$D = 0.4000$	$D = 0.2900$	$D = 0.4818$	$D = 0.4118$
	$p < .0001$	$p = .1685$	$p = .0340$	$p = .0051$
MLC[c]	$D = 0.8904$	$D = 0.7600$	$D = 0.9898$	$D = 0.8416$
	$p < .0001$	$p < .0001$	$p < .0001$	$p < .0001$
Merging	$D = 0.5345$	$D = 0.5568$	$D = 0.5822$	$D = 0.5064$
	$p < .0001$	$p < .0001$	$p < .0001$	$p < .0001$
Diverging	$D = 0.4596$	$D = 0.5402$	$D = 0.4045$	$D = 0.5419$
	$p < .0001$	$p = .0015$	$p < .0001$	$p < .0001$

Note: [a]: lane change; [b]: discretionary lane change; [c]: mandatory lane change.

each lane change is denoted as *minGTTC*, lane change duration is denoted as *LCD*, and the accepted gap of target lane at the initiation of lane change is represented by *AG*. The scatterplot illustrating the relationship between lane change duration to accepted gap ratio and *minGTTC* is presented in Fig. 22. The distribution of data points between field and simulation datasets is largely consistent. Most data points from both sources are concentrated in the region where the LCD/AG ratio ranges from 0 to 1. Specifically, when this ratio is within 0 to 1, a significantly higher number of data points exhibit *minGTTC* greater than 5 s. Conversely, when the ratio exceeds 1, the likelihood of *minGTTC* being less than 5 s increases. Field data shows a higher frequency of data points with larger LCD/AG ratios, possibly due to smaller accepted gaps observed in field compared to simulation as depicted in Fig. 20.

Notably, even within the 0 to 1 ratio range, numerous data points still have *minGTTC* less than 5 s. This discrepancy may be attributed to the dynamic nature of lane changes, where the observed variables represent only a single time segment and could be influenced by unobserved heterogeneity. Therefore, further statistical analysis is recommended.

4. DISCUSSIONS

In this study, driving behaviour in a weaving section derived from field data was compared with that obtained from a driving simulator experiment. Key metrics including speed, headway, lane change duration, accepted gap, and *GTTC* were employed for comparing the drivers' behaviour between the field and simulation.

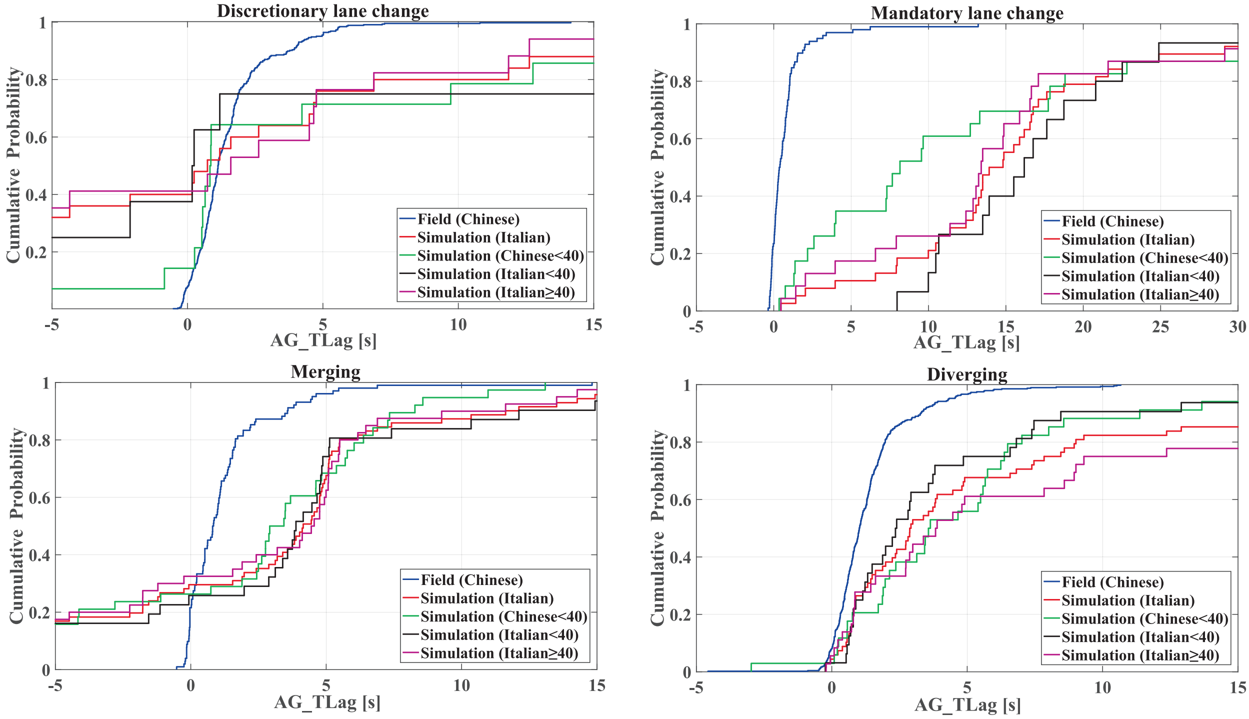


Fig. 20. Comparison of the cumulative probability of AG_TLag.

Table 11
The results of two-sided K-S test for *GTTC*.

LC[a] Type	Comparison Group Pairs			
	Field (Chinese) vs. Simulation (Italian)	Field (Chinese) vs. Simulation (Chinese < 40)	Field (Chinese) vs. Simulation (Italian < 40)	Field (Chinese) vs. Simulation (Italian ≥ 40)
DLC[b]	$D = 0.1652$ $p = .0286$	$D = 0.0925$ $p = .7952$	$D = 0.1605$ $p = .1854$	$D = 0.2150$ $p = .0323$
MLC[c]	$D = 0.2000$ $p = .0041$	$D = 0.0672$ $p = .9783$	$D = 0.1669$ $p = .1527$	$D = 0.2577$ $p = .0044$
Merging	$D = 0.3401$ $p < .0001$	$D = 0.1975$ $p = .0388$	$D = 0.3783$ $p < .0001$	$D = 0.3172$ $p < .0001$
Diverging	$D = 0.2771$ $p < .0001$	$D = 0.1862$ $p = .0413$	$D = 0.2678$ $p = .0037$	$D = 0.3152$ $p < .0001$

Note: [a]: lane change; [b]: discretionary lane change; [c]: mandatory lane change.

4.1. *Speed*

The results of two-sample K-S test presented in Table 6 indicate that only the speed distribution of young Italians field was significantly lower than that of field during through manoeuvre. Although no significant difference was found in speed distribution between field and simulation for merging manoeuvres, Fig. 13 reveals differences in the spatial distribution of average speed within the weaving section. Specifically, compared to field data, the simulation data showed higher average speeds at the on-ramp, exceeding the speed limit by up to 40 km/h, and lower speeds after merging into the main road. For diverging and ramp-to-ramp manoeuvres, the findings in Table 6 and Fig. 12 demonstrate that the speed in field was significantly lower than simulation in some groups (i.e., Italians and old Italians). Furthermore, Fig. 13 indicates that the average speed in the simulation data at the ramp section is significantly higher than that observed in the field data. This discrepancy may be attributed to the fact that the simulation environment could affect participants' speed perception and control, especially on curved sections like ramps (Godley et al., 2002; Chen et al., 2021). In real-world driving scenarios, the propensity to accelerate is significantly influenced by the driver's perceived risk, which escalates with increasing speed, particularly during ramp manoeuvres. Nevertheless, this effect is less pronounced in simulator environments, which often results in higher driving speeds (Fuller, 2005; Bella, 2008). This discrepancy is consistent with previous studies indicating that driving simulator speeds tend to be slightly higher than real-world data (Godley et al., 2002; Hussain et al., 2019). Additionally, a study conducted using the same simulator further validates these findings (Bassani et al., 2018).

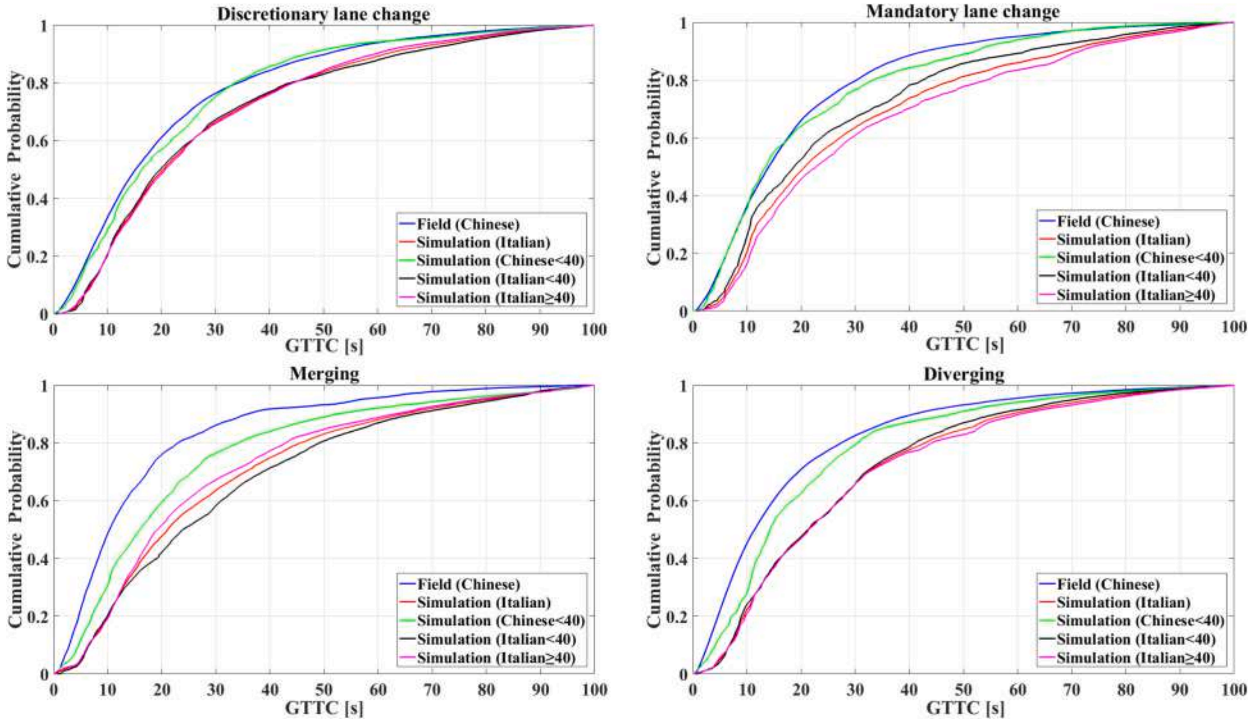


Fig. 21. Comparison of the cumulative probability of *GTTC*.

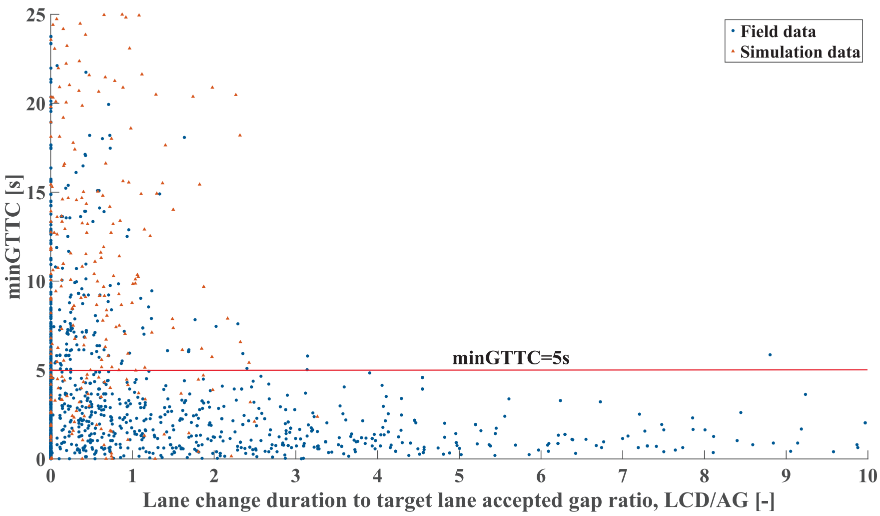


Fig. 22. Relationship between LCD/AG ratio and *minGTTC*.

4.2. Headway

During the through manoeuvre, Chinese exhibit a lower headway distribution compared to field, whereas the other three groups demonstrate similar headway distribution compared to field. For the diverging manoeuvre, no significant difference in headway distribution is observed between field and simulation data. In the merging manoeuvre, both Italians and old Italians display notably lower headways compared to field. For the ramp-to-ramp manoeuvre, Italians and young Italians show significantly reduced headways with field. Conversely, Fig. 15 reveals differences in the spatial distribution of headways during merging and ramp-to-ramp manoeuvres between simulation and field across all three sections. This discrepancy may be attributed to stop-and-go waves and frequent lane-changing behaviour in real-life scenarios, leading to traffic congestion propagation in the weaving section and affecting capacity at the tail of the weaving section (Oh & Yeo, 2015). In contrast, the simulation is constrained by general simulation software's limitations and struggles to accurately model the vehicle interactions and complex congestion dynamics at the tail of the weaving section (Wan et al., 2014). Additionally, the average headway maintained by Chinese during the four manoeuvres was significantly higher compared to other groups. This phenomenon could be attributed to the driving behaviour of a larger proportion of novice drivers within the Chinese participant group, who tended to keep a greater headway (Yang & Wu, 2017; Yang et al., 2021).

4.3. Lane change duration

The K-S test results presented in Table 8 indicate that simulation data exhibited a lower probability of discretionary lane changes taking a longer time. Discretionary lane change typically occurs when drivers seek better driving conditions, such as higher speeds or larger headways. Previous studies have demonstrated that as the complexity and risk of lane changes increase, the required duration for completing lane changes also increases (Toledo & Zohar, 2007; Aghabayk et al., 2011; Yang et al., 2019). In the simulation experiment, participants were not required to make a discretionary lane change, so they were more likely to abandon it if it took too long. For mandatory lane change, Chinese had a shorter lane change duration, which may be attributed to the operational characteristics of novice drivers. Research has demonstrated that novice drivers spend considerably less time checking the surrounding traffic compared to experienced drivers during lane changes (Yang et al., 2006).

Regarding the merging manoeuvre, the results presented in Figs. 17 and 18 indicate that simulation participants tended to initiate merges closer to the on-ramp, necessitating a longer lane-changing distance and time. In contrast, field drivers began merging at a position farther from the on-ramp, which allowed for sufficient acceleration and resulted in both a shorter lane-changing distance and time, thereby facilitating a smoother transition onto the main road. These findings align with prior research. Kou & Machemehl (1997) observed that some drivers prefer to remain in the acceleration lane for a longer duration, merging with the freeway traffic at a higher speed. Daamen et al. (2010) similarly noted that vehicles merging late during congestion tend to have a higher average speed.

For the diverging manoeuvre, field data exhibited a more uniform lane change duration distribution, while simulation data show a higher concentration between 5 and 10 s. Fig. 18 indicates that field drivers needed a greater distance before initiating diverging, leading to increased interactions with weaving traffic and consequently lower speeds, as shown in Figs. 12-13. According to Hu et al. (2019), vehicles approaching the diverging nose encountered progressively more complex and hazardous conditions, which led to prolonged lane change times. Additionally, when executing merging and diverging manoeuvres in both simulation and field data, the initiation points for these manoeuvres varied. This variation may stem from differences in depth perception and distance estimation between driving simulators and real-world environments (Kemeny & Panerai, 2003; Baumberger et al., 2005). Furthermore, depth perception and distance estimation can differ across simulators depending on their varying levels of fidelity (Himmels et al., 2024).

4.4. Accepted target lane lead and lag gap

During discretionary lane change simulation groups tended to accept larger AG_TLead values compared to field. Additionally, no significant difference was observed in AG_TLag distribution among Chinese, young Italians, and field data. Conversely, old Italians demonstrated greater AG_TLag characteristics. These results are consistent with previous findings suggesting that when discretionary lane change requires longer time, the simulation groups opt not to perform discretionary lane change due to accepting larger accepted gaps. Several studies have also demonstrated that accepted gaps are critical factors influencing a driver's decision to make a discretionary lane change (Xie et al., 2019; Alshehri & Abdul Aziz, 2022). For mandatory lane change, the distribution characteristics of AG_TLead between field and simulation showed no significant difference, supporting the conclusion that the lane change duration distribution in this scenario were similar. In contrast, one-sided K-S test results indicate that simulation groups exhibited significantly larger AG_TLag values. As illustrated in Fig. 17, simulation drivers changed lanes to the right earlier to create more favourable conditions for diverging. Consequently, this resulted in a notably larger lag gap. Regarding merging, no significant difference in the AG_TLead distribution characteristics was found between Italians, young Italians, old Italians, and field data, whereas Chinese exhibited larger AG_TLead values than field data. Additionally, the AG_TLag values were ranked in ascending order: Italians, old Italians, followed by field data, and finally Chinese and young Italians. This suggest that younger participants accepted larger lag gaps during merging. This observation contrasts with prior research (Chityala, 2017; Chityala et al., 2020). The discrepancy may be attributed to the fact that most young drivers in this study were students from universities and novice drivers, who generally display more conservative driving behaviours. Additionally, as shown in Fig. 20, certain simulation data during the merging process exhibited a negative AG_TLag, indicative of forced lane changes. This phenomenon can be attributed to the proximity of the simulation drivers' merge position to the curve, as illustrated in Fig. 17, which could obstruct the drivers' rear view of vehicles. The analysis reveals that simulation data consistently exhibited significantly higher AG_TLead and AG_TLag values compared to field data during diverging manoeuvres. It is evident that the AG_TLag distribution between simulation and field shows more discrepancies than the AG_TLead distribution. A plausible explanation is that drivers relied on mirrors to estimate the lag gap, which may result in less accurate perception than when estimating the lead gap (Yang et al., 2019). Furthermore, most drivers turned their heads to check the blind spot when changing lanes (Chovan et al., 1994; Barat & Das, 2024), but the field of view offered by the simulator utilized in this study is somewhat restricted. Consequently, simulation participants demonstrated reduced capability in detecting vehicles within the blind spot, resulting in more frequent forced lane changes and less confident lane changes that required larger lag gaps. Another possible cause of this phenomenon is the truncation of headway in the experimental design. Therefore, some relatively smaller headways were excluded from the simulation environment, leading to larger accepted gap being observed in the simulated vehicle data compared to those in the field.

4.5. GTTC

The results indicate that during the discretionary lane change scenarios, there was no significant difference in *GTTC* distribution between field and simulation. During mandatory lane change, Chinese and young Italians exhibited *GTTC* distributions that closely resemble those from field data, whereas Italians and old Italians displayed significantly higher *GTTC* values compared to field. This suggests that younger drivers had lower *GTTC* values during mandatory lane change, which aligns with previous findings (Montgomery et al., 2014; Ali et al., 2019). Regarding the merging and diverging scenarios, apart from Chinese data and field data showing no significant difference in *GTTC* distribution, the other three sets of simulation data' *GTTC* values were significantly higher than those of field data. This phenomenon may be attributed to the substantially smaller AG_TLag observed in field data (Fig. 20). Previous studies have demonstrated that a lower accepted gap correlates with a lower TTC value (Mwesige et al., 2016; Karimi et al., 2021). By analysing the scatter plot of LCD/AG ratio versus *minGTTC*, it is observed that the distribution of data points in the field data and simulation data was largely consistent, though some discrepancies remain. Additionally, field data exhibited a higher frequency of data points with larger LCD/AG ratios, likely attributable to the smaller accepted gaps observed in real-world scenarios. This indicates that simulation data may not fully capture the extremely hazardous conditions present in field data. Considering that lane changes entail a multifaceted process, risk assessment is inherently multidimensional and may be influenced by unobserved heterogeneity, thus warranting further investigation.

The results of the statistical tests on the measurements and corresponding interpretations are summarized in Table 12.

Table 12
Summary of statistical test results and corresponding interpretations.

Measurements	Statistical similarity	Statistical difference	Explanations and Implications
Speed	**Italians**: through, merging; **Chinese**: through, merging, diverging, ramp to ramp.	Italians showed higher speeds, particularly in ramps.	Simulation could affect speed perception and control, especially on curved sections like ramps (Godley et al., 2002; Chen et al., 2021).
Headway	**Italians**: through, diverging; **Chinese**: merging, diverging, ramp to ramp.	Differences found in the spatial distribution during merging and ramp-to-ramp manoeuvres.	Simulation struggles to accurately model the vehicle interactions and complex congestion dynamics at the tail of the weaving section (Wan et al., 2014).
		Chinese maintained larger headways.	Novice drivers tended to keep greater headways (Yang & Wu, 2017; Yang et al., 2021).
Lane change duration	**Italians**: mandatory lane change.	Longer discretionary lane change durations in field.	Fewer discretionary lane changes were observed in simulation.
		Chinese had shorter mandatory lane change durations.	Novice drivers spend less time checking the traffic compared to experienced drivers during lane changes (Yang et al., 2006).
		Simulation drivers required more time to merge compared to field drivers.	Field drivers merged from a position farther from the on-ramp, allowing enough acceleration and a smoother entry onto the main road (Kou & Machemehl, 1997; Daamen et al., 2010).
		Field drivers needed longer time to diverge.	Field drivers needed more distance to diverge, leading to increased interactions with weaving traffic and consequently lower speeds (Hu et al., 2019).
Accepted gap	**AG_Lead** (**Italians**: mandatory lane change, merging; **Chinese**: mandatory lane change) ; **AG_Lag** (**Chinese:** discretionary lane change).	Simulation groups accepted larger gaps during discretionary lane changes.	Simulation groups opt not to perform discretionary lane change due to accepting larger accepted gaps (Xie et al., 2019; Alshehri & Abdul Aziz, 2022).
		Simulation groups exhibited larger lag gaps during mandatory lane change.	Simulation drivers changed lanes to the right earlier to create more favourable conditions for diverging.
		Younger drivers accepted larger gaps during merging and diverging.	This observation contrasts with prior research (Chityala, 2017; Chityala et al., 2020), which may be due to the fact that most young drivers in this study were university students and novice drivers, who tend to drive more conservatively.
		Forced lane changes occurred in the simulation.	Drivers relied on mirrors to estimate the lag gap, which may result in less accurate perception than estimating the lead gap (Yang et al., 2019). Simulation drivers exhibited a diminished ability to detect vehicles within the blind spots (Chovan et al., 1994; Barat & Das, 2024).
GTTC	**Italians:** discretionary lane change; **Chinese:** discretionary lane change, mandatory lane change, merging, diverging.	Young drivers appeared to have lower GTTC values compared to other groups.	Though young drivers accepted larger gaps, they may not control speed well, leading to smaller GTTC values (Montgomery et al., 2014; Ali et al., 2019).
		Field drivers exhibited smaller GTTC values.	Substantially smaller lag gaps observed in the field. Previous studies have demonstrated that a lower accepted gap correlates with a lower TTC value (Mwesige et al., 2016; Karimi et al., 2021).

5. CONCLUSIONS

This study examined drivers' performance during lane change manoeuvres in weaving sections within both field and driving simulator. The objective was to validate the driving simulator as a reliable tool and enhance its utility for behavioural studies specifically targeting lane change manoeuvres in weaving sections. The field environment, specifically an expressway weaving section, was reconstructed into a 3D virtual environment. Comparable traffic conditions were replicated in the simulator scenarios. The same variables were defined and measured in both field and simulation to facilitate direct comparison. To enhance the generalizability of the findings, participants from both China and Italy were included in the driving simulation study.

In summary, as the results depicted in Table 12, for lane change behaviour in the weaving section, the method employed in this study achieves both absolute and relative validation for variables related to traffic efficiency (i.e., speed and headway) between driving simulator data and field data. In addition, relative validation is achieved for lane change related variables (i.e., lane change duration, accepted gap, *GTTC*), as the relative size relationship between simulation data and field data remains largely consistent. This can be attributed to the fact that simulating variables related to traffic efficiency is relatively straightforward, whereas lane-changing behaviour involves a complex process, making its simulation considerably more challenging. These results suggest that when researchers and practitioners use driving simulators to investigate driving behaviours in the weaving section in future studies, the driving behaviour variables associated with efficiency are largely consistent with those observed in real-world scenarios. In contrast, variables related to lane-changing behaviour require careful attention regarding their correspondence with real-world values.

In order to improve and promote the results of future research, several recommendations are proposed. (i) The provision of a high-fidelity immersive simulation environment is recommended to improve the driver's perception of speed and distance, particularly at ramp curves. Future research should focus on overcoming the limitations of simulation software to accurately replicate continuous traffic flow under complex road conditions such as weaving sections. (ii) The positions of merging and diverging initiations in weaving sections have a significant impact on the time taken to change the lane. Therefore, it is crucial to provide participants with navigation signals at appropriate locations during driving simulation experiments. (iii) When investigating lane change operations in a driving simulation environment, attention must be paid to simulating the cockpit mirror environment to provide the driver with an optimal field of view for observing the position and movement of surrounding vehicles. Virtual reality could be beneficial as it can provide a full 360-degree view. Furthermore, it would be advisable to carry out a validation study to assess its effectiveness.

Nevertheless, it must be acknowledged that there are certain limitations to this study. First, due to experimental limitations, the age and gender distributions of field drivers and simulator participants could not be perfectly matched. The current research relies on video trajectory data obtained in a city in China—Changsha. To broaden the applicability and robustness of these findings, the crucial next step is to collect and analyse driving behaviour data from various regions. Second, due to the limitations of the simulation software, it was difficult to precisely control vehicle operations within the traffic flow, and it was not possible to ensure that each driver interacted with vehicles in front and behind during each lane change, potentially affecting the overall distribution of variables. Third, interactions between human drivers play a crucial role during lane change manoeuvres in real driving environments, but this study could not fully capture these interactions due to experimental limitations. Last, the inherent complexity of traffic flow in weaving section poses challenges to accurately simulate micro-driving behaviours. Future research should focus on more precise data collection techniques and advanced simulation modelling methods to overcome the limitation. Therefore, caution should be exercised when applying the results of this study to other types of driving simulators for lane change behaviour studies in weaving sections, and each simulator should be validated individually.

CRediT authorship contribution statement

Suyi Mao: Writing – original draft, Methodology, Funding acquisition, Conceptualization. **Jaeyoung Jay Lee:** Writing – review & editing, Validation, Supervision, Funding acquisition. **Arastoo Karimi:** Writing – review & editing, Validation, Funding acquisition. **Alessandra Lioi:** Writing – review & editing, Validation. **Marco Bassani:** Writing – review & editing, Validation, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was funded by:

(1) the Natural Science Foundation of China under Grant No. 72471249;

(2) the Natural Science Foundation of Changsha under Grant No. kq2402221;

(3) the Hunan Provincial Innovation Foundation For Postgraduate under Grant No. CX20210235;

(4) the Fundamental Research Funds for the Central Universities of Central South University under Grant No. 2021zzts0166;

(5) the Changsha Major Science and Technology Project (No. kh2401002);

(6) the Compagnia di Sanpaolo Foundation, with the project: "DIRECTIONS - driver road interaction on future road infrastructures" (No. 2021.2146).

Data availability

Data will be made available on request.

References

Aghabayk, K., Moridpour, S., Young, W., Sarvi, M., & Wang, Y. B. (2011). Comparing heavy vehicle and passenger car lane-changing maneuvers on arterial roads and freeways. *Transportation research record, 2260*(1), 94–101. https://doi.org/10.3141/2260-11

Ali, Y., Zheng, Z., & Haque, M. M. (2018). Connectivity's impact on mandatory lane-changing behaviour: Evidences from a driving simulator study. *Transportation Research Part C: Emerging Technologies, 93*, 292–309. https://doi.org/10.1016/j.trc.2018.06.008

Ali, Y., Haque, M. M., Zheng, Z., Washington, S., & Yildirimoglu, M. (2019). A hazard-based duration model to quantify the impact of connected driving environment on safety during mandatory lane-changing. *Transportation research part C: emerging technologies, 106*, 113–131. https://doi.org/10.1016/j.trc.2019.07.015

Ali, Y., Zheng, Z., Haque, M. M., Yildirimoglu, M., & Washington, S. (2020). Detecting, analysing, and modelling failed lane-changing attempts in traditional and connected environments. *Analytic methods in accident research, 28*, Article 100138. https://doi.org/10.1016/j.amar.2020.100138

Alshehri, A., & Abdul Aziz, H. M. (2022). Analysis of Factors Affecting Discretionary Lane Change. In *International Conference on Transportation and Development 2022* (pp. 147-157).

Amini, E., Omidvar, A., & Elefteriadou, L. (2021). Optimizing operations at freeway weaves with connected and automated vehicles. *Transportation Research Part C: Emerging Technologies, 126*, Article 103072. https://doi.org/10.1016/j.trc.2021.103072

Association, W.M., 2024. WMA Declaration of Helsinki–Ethical Principles for Medical Research Involving Human Subjects. https://www.wma.net/policies-post/wma-declaration-of-helsinki/ [accessed on 10 March 2024].

Bham, G. (2011). A simple lane change model for microscopic traffic flow simulation in weaving sections. *Transportation Letters, 3*(4), 231–251.

Barat, A., & Das, S. (2024). A systematic literature review of driver behaviour in blind spot: Assessing risk and influencing factors for enhanced ergonomics in vehicle design. *Ergonomics, 1–17*. https://doi.org/10.1080/00140139.2024.2350667

Bassani, M., Catani, L., Ignazzi, A., & Piras, M. (2018). Validation of a fixed-base driving simulator to assess behavioural effects of road geometrics. In *Proceedings of the DSC2018 Europe* (pp. 101-108). Andras Kemeny, Florent Colombet, Frédéric Mérienne, Stephan Espié.

Baumberger, B., Flückiger, M., Paquette, M., Bergeron, J., & Delorme, A. (2005). Perception of relative distance in a driving simulator 1, 2. *Japanese Psychological Research, 47*(3), 230–237. https://doi.org/10.1111/j.1468-5884.2005.00292.x

Bella, F. (2005). Validation of a driving simulator for work zone design. *Transportation Research Record, 1937*(1), 136–144. https://doi.org/10.1177/0361198105193700119

Bella, F. (2008). Driving simulator for speed research on two-lane rural roads. *Accident Analysis & Prevention, 40*(3), 1078–1087. https://doi.org/10.1016/j.aap.2007.10.015

Berger, V. W., & Zhou, Y. (2014). Kolmogorov–smirnov test: Overview. Wiley statsref: Statistics reference online. doi: 10.1002/9781118445112.stat06558.

Branzi, V., Domenichini, L., & La Torre, F. (2017). Drivers' speed behaviour in real and simulated urban roads–a validation study. *Transportation research part F: traffic psychology and behaviour, 49*, 1–17. https://doi.org/10.1016/j.trf.2017.06.001

Cai, Q., Saad, M., Abdel-Aty, M., Yuan, J., & Lee, J. (2018). Safety impact of weaving distance on freeway facilities with managed lanes using both microscopic traffic and driving simulations. *Transportation research record, 2672*(39), 130–141. https://doi.org/10.1177/0361198118780884

Cassidy, M. J., & Bertini, R. L. (1999). Some traffic features at freeway bottlenecks. *Transportation Research Part B: Methodological, 33*(1), 25–42. https://doi.org/10.1016/S0191-2615(98)00023-X

Catani, L., & Bassani, M. (2019). Anticipatory distance, curvature, and curvature change rate in compound curve negotiation: a comparison between real and simulated driving. In *98th Annual Meeting of the Transportation Research Board, Washington, DC*.

Chen, D., Ahn, S., Laval, J., & Zheng, Z. (2014). On the periodicity of traffic oscillations and capacity drop: The role of driver characteristics. *Transportation research part B: Methodological, 59*, 117–136. https://doi.org/10.1016/j.trb.2013.11.005

Chen, D., & Ahn, S. (2018). Capacity-drop at extended bottlenecks: Merge, diverge, and weave. *Transportation Research Part B: Methodological, 108*, 1–20. https://doi.org/10.1016/j.trb.2017.12.006

Chen, L., Fang, J., Li, J., & Xie, J. (2025). Research on the Effectiveness of Driving simulation Systems in Risky Traffic Environments. *Systems, 13*(5), 329. https://doi.org/10.3390/systems13050329

Chen, L., Xie, J., Wu, S., Guo, F., Chen, Z., & Tan, W. (2021). Validation of vehicle driving simulator from perspective of velocity and trajectory based driving behavior under curve conditions. *Energies, 14*(24), 8429. https://doi.org/10.3390/en14248429

Chityala, S. (2017). Driver Behavior in Mixed Connected-Automated and Conventional Vehicle Traffic at a Freeway Merge.

Chityala, S., Sobanjo, J. O., Erman Ozguven, E., Sando, T., & Twumasi-Boakye, R. (2020). Driver behavior at a freeway merge to mixed traffic of conventional and connected autonomous vehicles. *Transportation research record, 2674*(11), 867–874. https://doi.org/10.1177/0361198120950721

Chovan, J. D., Tijerina, L., Alexander, G., & Hendricks, D. L. (1994). *Examination of lane change crashes and potential IVHS countermeasures (No*. United States. Joint Program Office for Intelligent Transportation Systems.

Chung, K., Rudjanakanoknad, J., & Cassidy, M. J. (2007). Relation between traffic density and capacity drop at three freeway bottlenecks. *Transportation Research Part B: Methodological, 41*(1), 82–95. https://doi.org/10.1016/j.trb.2006.02.011

Cowan, R. J. (1975). Useful headway models. *Transportation Research, 9*(6), 371–375.

Daamen, W., Loot, M., & Hoogendoorn, S. P. (2010). Empirical analysis of merging behavior at freeway on-ramp. *Transportation Research Record, 2188*(1), 108–118. https://doi.org/10.3141/2188-12

De Blasiis, M. R., Ferrante, C., Santilli, A., & Veraldi, V. (2017). Driving behavior in weaving maneuver: A driving simulator study. In *Advances in Human Aspects of Transportation: Proceedings of the AHFE 2016 International Conference on Human Factors in Transportation, July 27-31, 2016, Walt Disney World®, Florida, USA* (pp. 313-325). Springer International Publishing.

Drezner, Z., Turel, O., & Zerom, D. (2010). A modified Kolmogorov–Smirnov test for normality. *Communications in Statistics—Simulation and Computation®, 39*(4), 693–704. https://doi.org/10.1080/03610911003615816

Faschina, S., Stieglitz, R. D., Muri, R., Strohbeck-Kühner, P., Graf, M., Mager, R., & Pflueger, M. O. (2021). Driving errors, estimated performance and individual characteristics under simulated and real road traffic conditions–a validation study. *Transportation research part F: traffic psychology and behaviour, 82*, 221–237. https://doi.org/10.1016/j.trf.2021.07.018

Fuller, R. (2005). Towards a general theory of driver behaviour. *Accident analysis & prevention, 37*(3), 461–472. https://doi.org/10.1016/j.aap.2004.11.003

Godley, S. T., Triggs, T. J., & Fildes, B. N. (2002). Driving simulator validation for speed research. *Accident analysis & prevention, 34*(5), 589–600. https://doi.org/10.1016/S0001-4575(01)00056-2

Golob, T. F., Recker, W. W., & Alvarez, V. M. (2004). Safety aspects of freeway weaving sections. *Transportation Research Part A: Policy and Practice, 38*(1), 35–51. https://doi.org/10.1016/j.tra.2003.08.001

Gu, X., Abdel-Aty, M., Xiang, Q., Cai, Q., & Yuan, J. (2019). Utilizing UAV video data for in-depth analysis of drivers' crash risk at interchange merging areas. *Accident Analysis & Prevention, 123*, 159–169. https://doi.org/10.1016/j.aap.2018.11.010

Hall, F. L., & Agyemang-Duah, K. (1991). Freeway capacity drop and the definition of capacity. *Transportation research record*, (1320).

Hao, W., Zhang, Z., Gao, Z., Yi, K., Liu, L., & Wang, J. (2020). Research on mandatory lane-changing behavior in highway weaving sections. *Journal of advanced transportation, 2020*(1), Article 3754062. https://doi.org/10.1155/2020/3754062

HCM. (2000). Highway capacity manual. Washington, DC, 2(1), 1.

Himmels, C., Venrooij, J., Parduzi, A., Peller, M., & Riener, A. (2024). The bigger the better? investigating the effects of driving simulator fidelity on driving behavior and perception. *Transportation research part F: traffic psychology and behaviour, 101*, 250–266. https://doi.org/10.1016/j.trf.2024.01.007
Hu, M., Xiang, Q., Li, Y., Zheng, Z., & Zhang, X. (2019). Correlation analysis of traffic conflicts and driving behaviors at interchange diverging areas. In *CICTP 2019* (pp. 435-447).
Hussain, Q., Alhajyaseen, W. K., Pirdavani, A., Reinolsmann, N., Brijs, K., & Brijs, T. (2019). Speed perception and actual speed in a driving simulator and real-world: A validation study. *Transportation research part F: Traffic psychology and behaviour, 62*, 637–650. https://doi.org/10.1016/j.trf.2019.02.019
Karimi, A., Bassani, M., Boroujerdian, A. M., & Catani, L. (2020). Investigation into passing behavior at passing zones to validate and extend the use of driving simulators in two-lane roads safety analysis. *Accident Analysis & Prevention, 139*, Article 105487. https://doi.org/10.1016/j.aap.2020.105487
Karimi, A., Boroujerdian, A. M., & Amini, I. (2021). Risk evaluation of multiple passing maneuvers on two-lane rural highways in Iran. *Accident Analysis & Prevention, 163*, Article 106472. https://doi.org/10.1016/j.aap.2021.106472
Kanji, G. K. (2006). 100 statistical tests.
Kemeny, A., & Panerai, F. (2003). Evaluating perception in driving simulation experiments. *Trends in cognitive sciences, 7*(1), 31–37. https://doi.org/10.1016/S1364-6613(02)00011-6
Kou, C. C., & Machemehl, R. B. (1997). Modeling vehicle acceleration deceleration behavior during merge maneuvers. *Canadian journal of civil engineering, 24*(3), 350–358. https://doi.org/10.1139/l96-115
Kusuma, A., Liu, R., Choudhury, C., & Montgomery, F. (2015, January). Lane-changing characteristics at weaving section. In *Transportation Research Board 94th Annual Meeting* (Vol. 94, pp. 49-55).
Kusuma, A., Liu, R., & Choudhury, C. (2020). Modelling lane-changing mechanisms on motorway weaving sections. *Transportmetrica B: transport dynamics, 8*(1), 1–21. https://doi.org/10.1080/21680566.2019.1703840
Li, H., Zhang, J., Zhang, Z., & Huang, Z. (2021). Active lane management for intelligent connected vehicles in weaving areas of urban expressway. *Journal of intelligent and connected vehicles, 4*(2), 52–67. https://doi.org/10.1108/JICV-08-2020-0009
Marczak, F., Daamen, W., & Buisson, C. (2014, January). Empirical analysis of lane changing behavior at a freeway weaving section. In *TRB 2014-Transportation Research Board 93rd Annual Meeting* (pp. 14-p).
Marczak, F., Daamen, W., & Buisson, C. (2016). Empirical analysis of lane changing behavior at a freeway weaving section. *Traffic management, 3*, 139–151. https://doi.org/10.1002/9781119307822.ch10
May, A. D. (1990). *Traffic flow fundamentals*. Upper Saddle River, NJ: Prentice Hall.
Meng, Q., & Qu, X. (2012). Estimation of rear-end vehicle crash frequencies in urban road tunnels. *Accident Analysis & Prevention, 48*, 254–263. https://doi.org/10.1016/j.aap.2012.01.025
Meuleners, L., & Fraser, M. (2015). A validation study of driving errors using a driving simulator. *Transportation research part F: traffic psychology and behaviour, 29*, 14–21. https://doi.org/10.1016/j.trf.2014.11.009
Montgomery, D. C., & Runger, G. C. (2020). *Applied statistics and probability for engineers*. John wiley & sons.
Montgomery, J., Kusano, K. D., & Gabler, H. C. (2014). Age and gender differences in time to collision at braking from the 100-car naturalistic driving study. *Traffic injury prevention, 15*(sup1), S15–S20. https://doi.org/10.1080/15389588.2014.928703
Mullen, N., Charlton, J., Devlin, A., & Bedard, M. (2011). Simulator validity: Behaviours observed on the simulator and on the road. In *Handbook of driving simulation for engineering, medicine and psychology* (pp. 1-18). CRC Press.
Mwesige, G., Farah, H., & Koutsopoulos, H. N. (2016). Risk appraisal of passing zones on two-lane rural highways and policy applications. *Accident Analysis & Prevention, 90*, 1–12. https://doi.org/10.1016/j.aap.2016.01.014
Nguyen, H. D. (2017, August). A Two-sample Kolmogorov-Smirnov-like test for big data. In Australasian Conference on Data Mining (pp. 89-106). Singapore: Springer Singapore. doi: 10.1007/978-981-13-0292-3_6.
Oh, S., & Yeo, H. (2015). Impact of stop-and-go waves and lane changes on discharge rate in recovery flow. *Transportation Research Part B: Methodological, 77*, 88–102. https://doi.org/10.1016/j.trb.2015.03.017
Olea, R. A., & Pawlowsky-Glahn, V. (2009). Kolmogorov–Smirnov test for spatially correlated data. *Stochastic Environmental Research and Risk Assessment, 23*, 749–757. https://doi.org/10.1007/s00477-008-0255-1
Olsson, N. (2023). A validation study comparing performance in a low-fidelity train-driving simulator with actual train driving performance. *Transportation research part F: traffic psychology and behaviour, 97*, 109–122. https://doi.org/10.1016/j.trf.2023.07.007
Ouyang, P., Liu, P., Guo, Y., & Chen, K. (2023). Effects of configuration elements and traffic flow conditions on lane-changing rates at the weaving segments. *Transportation research part A: policy and practice, 171*, Article 103652. https://doi.org/10.1016/j.tra.2023.103652
Rakha, H., & Zhang, Y. (2004). INTEGRATION 2.30 framework for modeling lane-changing behavior in weaving sections. *Transportation Research Record, 1883*(1), 140–149. https://doi.org/10.3141/1883-16
Saha, P., Roy, R., Sarkar, A. K., & Pal, M. (2019). Preferred time headway of drivers on two-lane highways with heterogeneous traffic. *Transportation letters, 11*(4), 200–207. https://doi.org/10.1080/19427867.2017.1312859
Sarvi, M., Ejtemai, O., & Zavabeti, A. (2011, September). Modelling freeway weaving manoeuvre. In *Australasian Transport Research Forum* (pp. 28-30).
Sayed, T., Zaki, M. H., & Autey, J. (2013). Automated safety diagnosis of vehicle–bicycle interactions using computer vision analysis. *Safety science, 59*, 163–172. https://doi.org/10.1016/j.ssci.2013.05.009
Sulejic, D., Jiang, R., Sabar, N. R., & Chung, E. (2017). Optimization of lane-changing distribution for a motorway weaving segment. *Transportation Research Procedia, 21*, 227–239. https://doi.org/10.1016/j.trpro.2017.03.092
Sulejic, D. (2018). *Optimisation of lane-changing distributions for a freeway weaving segment*. Queensland University of Technology). Doctoral dissertation.
Tilg, G., Yang, K., & Menendez, M. (2018). Evaluating the effects of automated vehicle technology on the capacity of freeway weaving sections. *Transportation Research Part C: Emerging Technologies, 96*, 3–21. https://doi.org/10.1016/j.trc.2018.09.014
Toledo, T., & Zohar, D. (2007). Modeling duration of lane changes. *Transportation Research Record, 1999*(1), 71–78. https://doi.org/10.3141/1999-08
Underwood, G., Crundall, D., & Chapman, P. (2011). Driving simulator validation with hazard perception. *Transportation research part F: traffic psychology and behaviour, 14*(6), 435–446. https://doi.org/10.1016/j.trf.2011.04.008
Wan, X., Jin, P. J., Yang, F., Zhang, J., & Ran, B. (2014). Modeling vehicle interactions during merge in congested weaving section of freeway ramp. *Transportation Research Record, 2421*(1), 82–92. https://doi.org/10.3141/2421-10
Wan, X., Jin, P. J., Gu, H., Chen, X., & Ran, B. (2017). Modeling freeway merging in a weaving section as a sequential decision-making process. *Journal of Transportation Engineering, Part A: Systems, 143*(5), Article 05017002. https://doi.org/10.1061/JTEPBS.0000048
Wang, C., Zeng, B., & Shao, J. (2011, June). Application of bootstrap method in Kolmogorov-Smirnov test. In *2011 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering* (pp. 287-291). IEEE. doi: 10.1109/ICQR2MSE.2011.5976614.
Wang, W., Cheng, Q., Li, C., André, D., & Jiang, X. (2019). A cross-cultural analysis of driving behavior under critical situations: A driving simulator study. *Transportation research part F: Traffic psychology and behaviour, 62*, 483–493. https://doi.org/10.1016/j.trf.2019.02.007
Ward, J. R., Agamennoni, G., Worrall, S., Bender, A., & Nebot, E. (2015). Extending time to collision for probabilistic reasoning in general traffic scenarios. *Transportation Research Part C: Emerging Technologies, 51*, 66–82. https://doi.org/10.1016/j.trc.2014.11.002
Wynne, R. A., Beanland, V., & Salmon, P. M. (2019). Systematic review of driving simulator validation studies. *Safety science, 117*, 138–151. https://doi.org/10.1016/j.ssci.2019.04.004
Xie, D. F., Fang, Z. Z., Jia, B., & He, Z. (2019). A data-driven lane-changing model based on deep learning. *Transportation research part C: emerging technologies, 106*, 41–60. https://doi.org/10.1016/j.trc.2019.07.002
Yang, C., Chen, X., Lin, X., & Li, M. (2022). Coordinated trajectory planning for lane-changing in the weaving areas of dedicated lanes for connected and automated vehicles. *Transportation Research Part C: Emerging Technologies, 144*, Article 103864. https://doi.org/10.1016/j.trc.2022.103864

Yang, C., Shao, C., & Liu, L. (2012). Study on capacity of urban expressway weaving segments. *Procedia-Social and Behavioral Sciences, 43*, 148–156. https://doi.org/10.1016/j.sbspro.2012.04.087

Yang, G., Jaeger, B., & Mourant, R. R. (2006, October). Driving performance of novice and experienced drivers in lane-change scenarios. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 50, No. 26, pp. 2712-2716). Sage CA: Los Angeles, CA: SAGE Publications. doi: 10.1177/154193120605002609.

Yang, H., & Wu, Z. (2017). Effect of novice driver's car-following characteristic on roadway segment capacity. *Advances in Mechanical Engineering, 9*(7). https://doi.org/10.1177/1687814017711853, 1687814017711853.

Yang, M., Wang, X., & Quddus, M. (2019). Examining lane change gap acceptance, duration and impact using naturalistic driving data. *Transportation research part C: emerging technologies, 104*, 317–331. https://doi.org/10.1016/j.trc.2019.05.024

Yang, J., Lee, J., Mao, S., & Hu, J. (2023). Dynamic safety estimation of airport pick-up area based on video trajectory data. *IEEE Transactions on Intelligent Transportation Systems*. https://doi.org/10.1109/TITS.2023.3316199

Yang, Z., Yu, Q., Zhang, W., & Shen, H. (2021). A comparison of experienced and novice drivers' rear-end collision avoidance maneuvers under urgent decelerating events. *Transportation research part F: traffic psychology and behaviour, 76*, 353–368. https://doi.org/10.1016/j.trf.2020.12.009

Yuan, J., Abdel-Aty, M., Cai, Q., & Lee, J. (2019). Investigating drivers' mandatory lane change behavior on the weaving section of freeway with managed lanes: A driving simulator study. *Transportation research part F: Traffic psychology and behaviour, 62*, 11–32. https://doi.org/10.1016/j.trf.2018.12.007

Zhang, J., Lee, J., Abdel-Aty, M., Zheng, O., & Xiao, G. (2023). Enhanced index of risk assessment of lane change on expressway weaving segments: A case study of an expressway in China. *Accident Analysis & Prevention, 180*, Article 106909. https://doi.org/10.1016/j.aap.2022.106909

Zhang, S., Zhao, C., Zhang, Z., & Lv, Y. (2024). Driving Simulator Validation Studies: A Systematic Review. *Simulation Modelling Practice and Theory*. , Article 103020. https://doi.org/10.1016/j.simpat.2024.103020

Zhao, Y., Wang, Z., Wu, Y., & Ma, J. (2022). Trajectory-based characteristic analysis and decision modeling of the lane-changing process in intertunnel weaving sections. *PLoS One, 17*(4), Article e0266489. https://doi.org/10.1371/journal.pone.0266489

Zheng, O. (2019). Developing a Traffic Safety Diagnostics System for Unmanned AerialVehicles UsingDeep Learning Algorithms.

Zheng, Z., Ahn, S., Chen, D., & Laval, J. (2011). Applications of wavelet transform for analysis of freeway traffic: Bottlenecks, transient traffic, and traffic oscillations. *Transportation Research Part B: Methodological, 45*(2), 372–384. https://doi.org/10.1016/j.trb.2010.08.002

Zhou, Z., Li, L., Ran, B., & Qu, X. (2023). A Cooperative Lane-Changing strategy for Weaving Sections of Urban Expressway under the Connected Autonomous Vehicle Environment. *Journal of Advanced Transportation, 2023*(1), Article 3363057. https://doi.org/10.1155/2023/3363057