Article:
Xu, Z., Wang, Y., Liu, R. orcid.org/0000-0003-0627-3184 et al. (3 more authors) (2025) Impact of the spatial distribution of CAVs on stochastic traffic oscillations in mixed traffic flow. Physica A: Statistical Mechanics and its Applications, 677. 130940. ISSN: 0378-4371

https://doi.org/10.1016/j.physa.2025.130940

Impact of the spatial distribution of CAVs on stochastic traffic oscillations in mixed traffic flow

Zeqi Xu[1,2], Yi Wang[1], Ronghui Liu[2], Yunxia Wu[1], Yangsheng Jiang[1], Zhihong Yao[1*]

1. School of Transportation and Logistics, National Engineering Laboratory of Integrated Transportation Big Data Application Technology, National United Engineering Laboratory of Integrated and Intelligent Transportation, Southwest Jiaotong University, Chengdu, Sichuan 610031, China;

2. Institute for Transport Studies, University of Leeds, Leeds LS2 9JT, UK

Abstract

Although existing studies have demonstrated the potential of Connected and Automated Vehicles (CAVs) to optimise traffic flow and suppress disturbances, most current mixed traffic flow models adopt idealised deterministic approaches for modelling Human-Driven Vehicles (HDVs). Such methods are limited in their ability to accurately capture the evolution of oscillations observed in real-world traffic. Moreover, the influence of the spatial distribution of CAVs on system performance is often overlooked. To address these limitations, this study introduces a novel metric—Platoon intensity—to quantify the spatial clustering characteristics of CAVs within mixed traffic flow. This indicator enables a unified characterisation of CAV distribution patterns across various penetration rates, and theoretical bounds on pairwise vehicle probabilities under different traffic conditions are derived accordingly. A mixed traffic flow model is further developed, incorporating stochastic car-following behaviour of HDVs, behavioural degradation of CAVs, and a constraint on maximum platoon size. By introducing stochastic differential equations, the model successfully reproduces velocity fluctuations triggered by endogenous disturbances. Based on this framework, a series of systematic numerical experiments are conducted to comprehensively analyse traffic efficiency, stability, and energy consumption under varying CAV penetration rates and spatial distribution patterns. A quantitative relationship is established between platoon intensity and macroscopic traffic performance indicators. The main findings of this study are as follows: (1) The spatial distribution of vehicles significantly impacts macroscopic traffic performance, with maximum differences of 9.70%, 145.20%, and 7.58% observed in average speed, coefficient of variation of speed, and average energy consumption, respectively. (2) At a fixed CAV penetration rate, increasing platoon intensity enhances traffic efficiency and reduces average energy consumption, but exacerbates traffic instability. This research provides theoretical insights and practical implications for future CAV deployment strategies and traffic management measures.

Keywords: *Mixed traffic flow ; Connected and Automated Vehicles ; stochastic car-following model ; Spatial distribution of vehicles ; Platoon intensity*

1. Introduction

With the rapid advancement of intelligent connected and automated driving technologies, Connected and Automated Vehicles (CAVs) are widely regarded as one of the key driving forces behind the transformation of future transport systems. Leveraging high-precision sensing, vehicle-to-vehicle communication, and autonomous control technologies, CAVs possess capabilities such as cooperative driving, rapid response, and disturbance suppression [1], and are thus expected to significantly enhance traffic safety [2], [3] , efficiency[4], [5] , and sustainability [6], [7]. However, for the foreseeable mid-to-long term, road traffic systems will inevitably operate under a mixed traffic condition comprising both CAVs and Human-Driven Vehicles (HDVs) [8]. The study of such mixed traffic flow is of crucial importance for future traffic control and optimisation strategies [9]. During this transitional period, the behavioural heterogeneity between CAVs and HDVs will pose unprecedented challenges to traffic flow stability, evolution patterns, and control mechanisms [10], while also offering new research opportunities for traffic modelling and system optimisation.

As a typical socio-technical complex system, the dynamic behaviour of road traffic is profoundly influenced by the stochasticity inherent in HDV driving behaviours. Such randomness is manifested in car-following processes through acceleration fluctuations and uncertain gaps during lane changes [11]. The stochastic nature of HDV behaviour originates from heterogeneous factors such as driver perception differences, psychological and physiological characteristics, and environmental sensing errors [12]. Studies have shown that the standard deviation of acceleration during car-following can reach 0.2–0.4 m/s^2 [13]. These microscopic random perturbations are nonlinearly amplified through inter-vehicle interactions, leading to macroscopic instabilities such as traffic oscillations and reduced throughput [14]. Although CAVs naturally exhibit the potential to optimise flow and suppress disturbances, previous studies [7], [15] have indicated that a 20% CAV penetration rate can improve road capacity by 9%–18% and enhance fuel efficiency by approximately 8.14%. Nevertheless, many existing mixed traffic flow models continue to represent HDV behaviour using deterministic car-following models [16], [17]. While such idealised modelling facilitates theoretical analysis, it neglects the widespread metastable and fluctuating phenomena observed in real-world traffic [18], limiting the model's ability to capture the evolution of traffic oscillations and potentially underestimating the systemic benefits of CAVs. More importantly, current research often assumes random distribution of CAVs within traffic flow [3], [19], or considers only a limited number of typical configurations [20] (e.g., front/rear concentration or alternating arrangements), largely overlooking the profound impact of CAV spatial distribution patterns on macroscopic traffic performance.

The long-standing lack of systematic investigation into CAV spatial distribution stems from two main challenges: first, the absence of a unified physical interpretation of spatial distribution; second, the high diversity of CAV arrangements and the lack of an established mapping between these configurations and macroscopic traffic performance. To address these challenges, this study introduces the concept of platoon intensity to quantitatively describe the spatial distribution of CAVs. This metric reflects the degree of spatial clustering among CAVs by measuring the proportion of CAV–CAV follower pairs among all CAVs. It thereby offers a unified and comparable spatial distribution indicator across different penetration rates and arrangement scenarios. Building on this, a mixed traffic flow model is developed that simultaneously accounts for the stochastic car-following behaviour of HDVs and the spatial distribution characteristics of CAVs. At the microscopic level, stochastic differential equations are introduced to represent the random nature of HDV behaviour, enabling the reproduction of endogenous traffic oscillations not caused by lane changes or external disturbances. At the spatial level, the evolution of CAV clustering structures in mixed traffic is modelled using platoon intensity. This framework supports comparative analyses of traffic efficiency, stability, and energy consumption under varying CAV penetration levels and spatial distribution patterns.

The main contributions of this study are as follows:

1） An improved definition of platoon intensity is proposed to quantify the spatial distribution of CAVs

in mixed traffic flow. In addition, the theoretical proportions of four types of vehicle pairs under different traffic conditions are derived, along with the theoretical bounds of platoon intensity.

2) A mixed traffic flow model is constructed that incorporates HDV behavioural stochasticity, CAV degradation, and a maximum platoon size constraint, offering a more realistic reproduction of endogenous traffic oscillation generation and propagation.

3) A comprehensive simulation study is conducted covering all CAV spatial distribution scenarios, through which quantitative relationships between platoon intensity and macroscopic performance indicators are established, revealing the mechanisms by which CAV spatial distribution affects traffic efficiency, instability, and energy consumption.

The remainder of this paper is structured as follows: Section 2 provides a comprehensive review of the development of traffic flow modelling theories, with an emphasis on stochastic modelling approaches and progress in CAV spatial distribution research. Section 3 details the construction and validation of the proposed mixed traffic flow model, which incorporates both behavioural stochasticity and vehicle spatial distribution. Section 4 presents the design and analysis of the numerical simulation experiments. Section 5 concludes the paper and discusses the implications of the findings.

2. Literature review

2.1. Stochastic car-following behaviour and traffic oscillations

In traditional traffic flow studies, the formation of traffic oscillations has long been attributed to lane-changing manoeuvres, sudden deceleration, or disturbances at bottlenecks. However, Laval et al. [21], through theoretical modelling and numerical simulation, demonstrated that even in the absence of external disturbances or lane changes, oscillations can still arise purely from the inherent stochasticity in human driving behaviour. This phenomenon was initially proposed by Yeo et al. [22], although it was not validated at the modelling level. Such endogenous traffic oscillations reveal the deep-rooted influence of microscopic uncertainty in HDV driving behaviour on the macroscopic stability of traffic flow [23]. Jiang et al. [24] further validated the generation and propagation of such endogenous oscillations through full-scale vehicle experiments on open roads. This uncertainty stems from driver heterogeneity and the random nature of physiological responses and behavioural actions [13]. To simulate and investigate such uncertain behaviours, scholars have incorporated various forms of random disturbances into car-following models to characterise stochastic fluctuations in drivers' perception, decision-making, and execution processes. These disturbances represent deviations in reaction time, desired headway, acceleration judgement, and throttle or brake operations [25], becoming a crucial modelling approach for representing the stochastic nature of microscopic traffic flow.

Currently, mainstream stochastic car-following models can be broadly classified into three types: (1) Additive noise models: These introduce noise into deterministic car-following models. For instance, Liu et al. [26] incorporated a stochastic disturbance term into the acceleration function of the Optimised Velocity Model and designed a velocity-difference feedback control strategy to enhance flow stability. Laval et al. [21] introduced white noise into the Newell model, formulating a stochastic differential equation that revealed a concave growth pattern of velocity standard deviation along the vehicle platoon. Sharma et al. [27] proposed a Human driver Intelligent Driver Model with estimation errors, using a Wiener process to model time-varying estimation errors in spacing and relative speed. Bouadi et al. [28] proposed a general stochastic car-following equation with speed-dependent noise and, through linearisation, derived a state-space model that confirmed the destabilising effect of stochasticity on traffic flow. (2) Parameter-randomised models: These assume that parameters of deterministic models follow specific distributions. Jin et al. [29] modelled driver heterogeneity using a normally distributed desired time headway. Zheng et al. [20] treated HDVs' free-flow speed, minimum safety gap, and inverse reaction time as Beta-distributed random variables and employed the Newell nonlinear speed-spacing relation to capture car-following behaviour. Li et al. [30] introduced randomness by allowing the desired time gap in the IDM to fluctuate within an interval, capturing

behavioural variability. (3) Behavioural-cognitive models: These are based on human behavioural traits and cognitive processes. Wang et al. [31] classified HDV driving states into inattentive, alert, and close-following, using state-dependent reaction times and acceleration transitions to simulate stochastic following behaviour. Similarly, Tian et al. [32] proposed a mode-switching stochastic model that incorporated spacing-insensitive regions and adopted different car-following modes under varying headways. Yuan et al. [33] provided a behaviourally grounded explanation by modelling the mean and standard deviation of desired acceleration as linear functions of speed, and proposed a desired acceleration error framework to overcome the limitations of deterministic models. It is worth noting that not all stochastic car-following models successfully reproduce endogenous oscillations. Many studies introduced external disturbances, such as sudden deceleration or bottlenecks, into HDV traffic to trigger oscillations [11], [25], [27]. These oscillations are exogenous rather than induced by stochastic car-following behaviour. The key difference is that endogenous oscillations, while smaller in amplitude, are widely generated in HDV traffic [34]. Successfully reproducing spontaneous oscillations has become a key criterion for evaluating stochastic car-following models. Building on the successful reproduction of endogenous traffic oscillations, this study employs the indicator of "platoon intensity" to explore the spatial distribution of CAVs — an aspect rarely considered in existing literature [1], [13], [35], [36].

2.2. Spatial distribution of CAVs and platoon intensity

In mixed traffic flow systems, the introduction of CAVs brings not only new car-following behaviours but also transforms the spatiotemporal dynamics of traffic. In recent years, researchers have increasingly recognised that the spatial distribution of CAVs within traffic flow significantly influences system performance [37]. Even under the same CAV penetration rate, different spatial arrangements — such as clustering of CAVs, uniform distribution, or interleaving with HDVs — may lead to markedly different traffic outcomes [38]. As such, the macroscopic effects of CAV spatial distribution in mixed traffic have become an emerging research focus. Some scholars have explored this issue by constructing and simulating specific spatial distribution scenarios. Zhu et al. [39] examined six vehicle distribution scenarios and found that concentrating CAVs at the front of the traffic flow yielded better dynamic performance. Dong et al. [40] studied platoons of 2 to 20 vehicles, discovering that placing CAVs at both ends of the platoon optimised energy efficiency. Sharma et al. [27] defined three distribution scenarios — optimal, worst, and random — based on the number of degraded CAVs and investigated them under varying penetration rates. However, such studies are limited in scope and lack quantitative indicators to measure spatial distribution [41], making them insufficiently comprehensive and physically interpretable.

To quantify and characterise CAV spatial distribution, some studies have proposed indicators based on consecutive vehicle types, such as continuous CAV sequence length [42] or average CAV platoon size [43]. However, these indicators are difficult to generalise or normalise under high penetration rates or complex distributions. To address this limitation, researchers have introduced the concept of platoon intensity, a normalised metric for measuring CAV spatial clustering. Low platoon intensity indicates dispersed CAVs, whereas high platoon intensity suggests aggregation into platoons [15]. Ghiasi et al. [44] proposed a Markov chain-based platoon intensity definition suited to infinite traffic flows, though it lacks interpretability and is less applicable to finite vehicle numbers. Jiang et al. [45] defined platoon intensity as the proportion of CAVs forming cooperative platoons, but this measure has limited sensitivity to spatial variation. More recently, Zhao et al. [46] and He et al. [47] introduced new definitions based on adjacent vehicle relationships, linking platoon intensity to HDVs and making it suitable for studying endogenous oscillations triggered by HDV randomness. Zhao's definition, based on autocorrelation functions, is relatively complex and lacks clear physical interpretation. By contrast, He's definition — the ratio of CAV–CAV pairs to total CAVs — offers a unified and interpretable measure of clustering across different CAV penetration rates and is more suitable for the current study. Building upon He's simplified definition, this study expands the range of platoon intensity values and explores its relationship with CAV penetration rate.

Table 1 summarises recent studies on randomness in mixed traffic and vehicle spatial distribution.

Compared with existing literature, this work uniquely considers both HDV stochasticity and varying spatial distributions of CAVs, providing a comprehensive examination of how spatial distribution affects macroscopic traffic performance under stochastic HDV behaviour.

Table 1 Comparison of mixed traffic flow research

Author, Year	Stochasticity	Endogenous traffic oscillation	Spatial distribution	Platoon intensity	Max platoon size	Degeneration of CAVs
Mao et al., 2024[11]	√	√	√			
Li et al. 2024[13]	√	√				
Lu et al., 2024[25]	√		√			√
Sharma et al., 2021[27]	√		√			√
Jin et al., 2021[29]	√		√			√
Zheng et al., 2019[34]	√	√	√			
Yao et al., 2022[38]			√	√	√	√
Dong et al., 2025[40]	√	√	√			√
Liu et al., 2024[41]	√		√			√
Ghiasi et al., 2017[44]			√	√		
Zhao et al., 2024[46]			√	√		
Wang et al., 2024[48]					√	√
Jiang et al. 2021[49]					√	√
This paper	√	√	√	√	√	√

3. Methodology

3.1. Basic definitions and assumptions

The mixed traffic flow considered in this study consists of two types of vehicles: Human-Driven Vehicles (HDVs) and Connected and Automated Vehicles (CAVs). HDVs are vehicles that are fully controlled by human drivers, without any sensing devices, communication modules, or autonomous driving capabilities. In contrast, CAVs are equipped with advanced information and communication systems, sensors, and autonomous driving systems that enable vehicle-to-everything (V2X) communication and driverless operation. CAVs can be further classified based on their leading vehicle into: DCAV (Degraded CAV), FCAV (Following CAV), and LCAV (Leading CAV).

Given the research aims and the influence of various factors in complex road traffic systems, the following assumptions are made:

(1) Sensor failures and communication delays are not considered.

(2) The road is a single-lane with no lane-changing or on/off-ramps; the total number of vehicles remains constant.

(3) All vehicles are assumed to have the same length; variations in CAV models and performance are neglected.

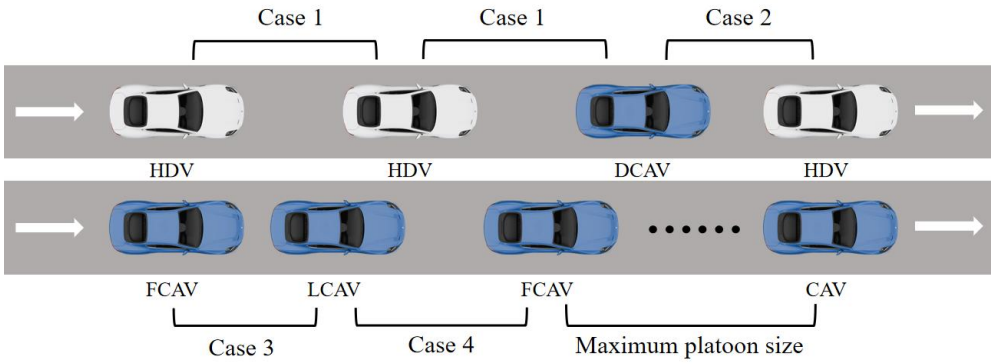


Fig. 1 Car-following scenarios in mixed traffic flow

Figure 1 illustrates the different car-following scenarios observed in mixed traffic flow. These scenarios are closely related to the car-following models, stochastic behaviour, and spatial distribution. Four types of car-following relationships are defined based on vehicle combinations.

Case 1：HDV Following Behaviour (HDV following HDV or CAV)

Here, the following vehicle is an HDV, which cannot communicate with the vehicle ahead. This corresponds to a traditional human-driving scenario, where the minimum headway depends on driver reaction time. Even if the preceding vehicle is a CAV, the HDV cannot perceive or respond to its data. When sudden changes occur in the lead vehicle's behaviour, the human driver must respond with a delay. Due to the high variability in driver behaviour and surrounding conditions, this scenario is highly stochastic and the cause of endogenous traffic oscillations.

Case 2：DCAV Following Behaviour (CAV following HDV)

When a CAV follows an HDV, it degrades into a DCAV because it can no longer use V2X communication. It must rely solely on sensors to perceive the HDV's behaviour. Since the HDV lacks communication capabilities, the CAV cannot anticipate changes and must react in real-time. DCAVs, as the CAVs closest to HDVs, play a critical role in suppressing stochastic oscillations. Unlike HDVs, DCAVs exhibit deterministic car-following behaviour.

Case 3：FCAV Following Behaviour (CAV following CAV within the same platoon)

When two CAVs are part of the same platoon, the following CAV (FCAV) can receive real-time data from the lead CAV, including acceleration and deceleration. This allows very short headways and coordinated driving. Whether the preceding vehicle is a DCAV or FCAV, the car-following mode remains identical. This scenario reflects fully automated platoon-following behaviour and is devoid of randomness.

Case 4：LCAV Following Behaviour (CAV following CAV outside the same platoon)

To avoid overly long platoons, a maximum platoon size is defined. When exceeded, new platoons are formed, and the first CAV in the new group becomes the LCAV. LCAVs do not coordinate with vehicles ahead. This car-following scenario is also fully deterministic.

3.2. Spatial distribution in mixed traffic flow

3.2.1 Definition of the platoon intensity index

The spatial distribution of CAVs refers to their positioning within the traffic flow, including the number of CAVs, platoon formations, and their locations relative to other vehicles. In this study, we quantify the degree of CAV clustering using the Platoon Intensity (PI), an indicator that reflects how densely CAVs are grouped. This metric helps examine how spatial distribution evolves under different penetration rates, thereby supporting theoretical analyses of mixed traffic scenarios. In the illustrative examples presented, we assume a fixed CAV penetration rate of 50%, meaning that in a 10-vehicle flow, there are 5 CAVs and 5 HDVs.

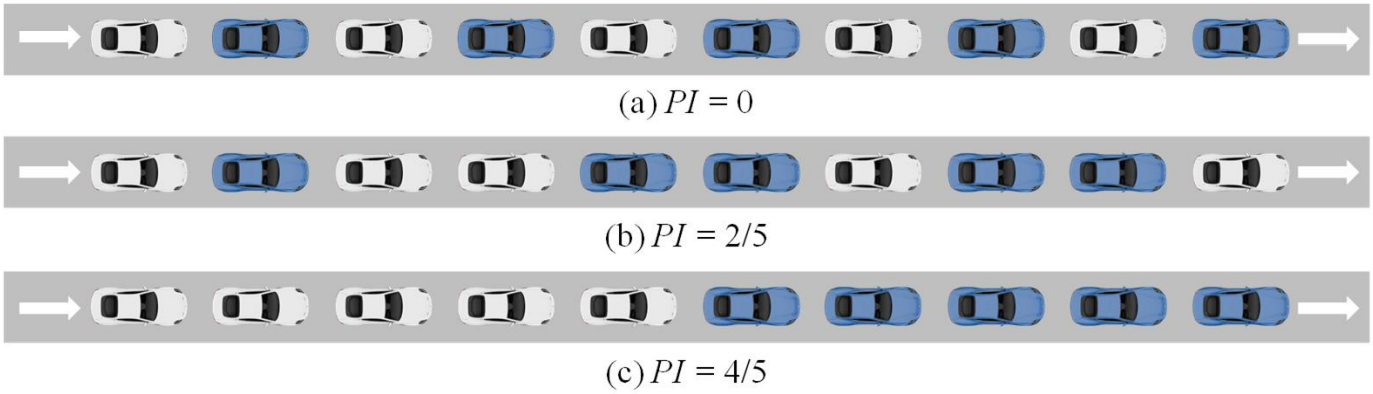
Due to the multitude of possible spatial arrangements, quantifying vehicle distributions with a clear physical interpretation presents a significant challenge. This study approaches the problem from the perspective of CAV clustering, adopting the concept of platoon intensity to explore the underlying mechanisms through which spatial distribution affects traffic flow. Platoon intensity is defined as the proportion of CAVs that are following another CAV within a platoon[47], as shown in Equation (1). This definition allows for an accurate representation of various spatial distribution patterns of CAVs. The primary reason for employing this metric lies in its clear physical relevance to traffic dynamics. The intensity of CAV clustering directly determines the car-following regimes in mixed traffic flow and can be further linked to the emergence of traffic oscillations induced by stochastic following behaviours.

$$PI = \frac{N_{CC}}{N_{CAV}}, \tag{1}$$

where N_{CC} is the number of CAV-CAV pairs, and N_{CAV} is the total number of CAVs. By definition, PI ranges from 0 (no platooning) to 1 (only achievable in a full-CAV scenario).

As illustrated in Figure 2, in scenario (a), all CAVs travel individually and are separated by HDVs,

without forming any platoons. In this case, the number of CAV-following-CAV pairs is zero, resulting in a platoon intensity of 0. In scenario (b), four out of the five CAVs operate in two platoons of size two. There are two CAV-following-CAV pairs, giving a platoon intensity of 2/5. Compared with scenario (a), the degree of CAV clustering within the mixed traffic flow increases. In scenario (c), all five CAVs travel as a single platoon, resulting in four CAV-following-CAV pairs and thus a platoon intensity of 4/5 — the highest possible value for this traffic setting, indicating the greatest level of CAV clustering. Therefore, the use of platoon intensity provides an intuitive and physically meaningful measure to characterise the spatial distribution of vehicles.



(a) $PI = 0$

(b) $PI = 2/5$

(c) $PI = 4/5$

Fig. 2 Use platoon intensity to quantify different spatial distribution of CAVs

3.2.2 Vehicle pairing probabilities

Platoon intensity, expressed as a proportion of vehicle counts, characterises the spatial distribution of CAVs within mixed traffic flow. Based on this definition, it is possible to derive the spatial arrangement of vehicles and the occurrence probabilities of different vehicle pairs under a given level of platoon intensity. This derivation facilitates a deeper understanding of the relationship between platoon intensity and spatial distribution, and supports the use of platoon intensity to represent specific mixed traffic configurations — thereby laying a foundation for future large-scale simulation experiments.

Lemma 1: The probability of four types of vehicle pairs appearing in mixed traffic flow is:

$$\begin{cases} P_{CC} = \rho \cdot PI \\ P_{HC} = \rho \cdot (1 - PI) \\ P_{CH} = (1 - PI) \cdot \rho \\ P_{HH} = (1 - \rho) - (1 - PI) \cdot \rho \end{cases}, \tag{2}$$

where $P_{CC}, P_{HC}, P_{CH}, P_{HH}$ representing the probabilities of a CAV following a CAV, an HDV following a CAV, a CAV following an HDV, and an HDV following an HDV, respectively. ρ denotes the CAV penetration rate, i.e. the proportion of CAVs in the traffic flow.

Proof:

Assuming a total of N vehicles, the numbers of CAVs and HDVs are ρN and $(1 - \rho)N$, respectively. The platoon intensity further categorises CAVs into two groups: those following another CAV and those following an HDV. Accordingly, the quantities of these two groups are ρNPI and $\rho N(1 - PI)$, respectively. To model the mixed traffic flow under the given definition of platoon intensity, a Markov chain framework is adopted. It is assumed that the vehicle sequence follows a Markov process, wherein the type of the following vehicle (i.e., the transition probability) depends solely on the type of the current vehicle. Let the state of the nth vehicle be denoted by $Z_n \in \{CAV, HDV\}$, where C represents a CAV and H represents an HDV. Under this formulation, four possible vehicle pair types exist in the mixed traffic flow: CC, CH, HH, and HC. The definition of platoon intensity can thus be reformulated in terms of the probabilities of these vehicle pair types as follows:

$$PI = \frac{N_{CC}}{N_{CAV}} = \frac{N_{CC}}{N_{CC} + N_{CH}}, \tag{3}$$

where N_{CC} is the number of vehicle pairs in which a CAV follows another CAV, N_{CH} is the number of

vehicle pairs where a CAV follows an HDV, N_{CAV} is the total number of CAVs.

The corresponding transition probability matrix can be expressed as follows:

$$T = \begin{bmatrix} t_{CC} & t_{HC} \\ t_{CH} & t_{HH} \end{bmatrix}. \tag{4}$$

The elements of this matrix represent the conditional probabilities of a particular type of vehicle following a given leading vehicle. For instance, t_{HC} denotes the probability that an HDV follows a CAV. For the sake of derivation, we assume $t_{CC} = q$, which represents the probability that a CAV is followed by another CAV. Given that only two vehicle types exist in the mixed traffic flow, it follows that $t_{HC} = 1 - q$, i.e., the probability that a CAV is followed by an HDV. Similarly, we define $t_{CH} = r$, and consequently $t_{HH} = 1 - r$.

At steady state, the Markov chain satisfies the detailed balance condition, meaning that the proportion of HDVs converted to CAVs is equal to the proportion of CAVs converted to HDVs. Based on the definitions of t_{CH} and t_{HC}, the following equilibrium condition is obtained:

$$r \cdot (1 - \rho) = (1 - q) \cdot \rho. \tag{5}$$

Since the platoon intensity reflects the proportion of CC vehicle pairs among all CC and CH vehicle pairs, Eq. (3) can be reformulated to yield Eq. (6).

$$PI = \frac{N_{CC}}{N_{CC} + N_{CH}} = \frac{q \cdot \rho}{q \cdot \rho + r \cdot (1 - \rho)}. \tag{6}$$

By substituting Eq. (5) into this expression and simplifying, the transition probabilities q and r can be explicitly derived in terms of PI and ρ, as shown below:

$$q = PI, \tag{7}$$

$$r = \frac{(1 - PI) \cdot \rho}{1 - \rho}. \tag{8}$$

Accordingly, based on the state transition probabilities, the probabilities of the four possible vehicle pair types in mixed traffic flow can be determined:

$$\begin{cases} P_{CC} = P_{CAV} \cdot t_{CC} = \rho \cdot q = \rho \cdot PI \\ P_{HC} = P_{CAV} \cdot t_{HC} = \rho \cdot (1 - q) = \rho \cdot (1 - PI) \\ P_{CH} = P_{HDV} \cdot t_{CH} = (1 - \rho) \cdot r = (1 - PI) \cdot \rho \\ P_{HH} = P_{HDV} \cdot t_{HH} = (1 - \rho) \cdot (1 - r) = (1 - \rho) - (1 - PI) \cdot \rho \end{cases}. \tag{9}$$

These conditional probabilities are denoted as P_{CC}, P_{HC}, P_{CH}, and P_{HH}, representing respectively: the probability that a CAV is followed by a CAV; the probability that a CAV is followed by an HDV; the probability that an HDV is followed by a CAV; and the probability that an HDV is followed by an HDV. According to the definitions and Eq. (9), it is straightforward to verify the validity of Eq. (10), which confirms that the sum of the probabilities of all vehicle pair types equals one. This consistency ensures the rationality of the derived vehicle pair probabilities in mixed traffic flow.

$$\begin{cases} P_{CC} + P_{CH} = P_{CAV} \\ P_{HH} + P_{HC} = P_{HDV} \\ P_{CC} + P_{CH} + P_{HH} + P_{HC} = 1 \end{cases}. \tag{10}$$

Lemma 1 proof ends.

To evaluate the effectiveness and accuracy of the proposed method for constructing specific mixed traffic flows based on vehicle pair probabilities, a series of computer-based simulation experiments were designed for systematic validation. Under a fixed CAV penetration rate of 50%, mixed traffic flows were generated with platoon intensities of 0.2, 0.4, 0.6, and 0.8, respectively. For each scenario, a traffic flow consisting of 1,000 vehicles was created, and each condition was repeated 10 times. The actual CAV penetration rates and platoon intensities of the generated flows were then computed for each run. Figure 3 presents the mean and standard deviation of each group of experiments, with error bars indicating the standard deviation across repeated trials. It can be observed that, across all specified platoon intensity levels, the actual traffic flow parameters closely align with their theoretical expectations, with minimal deviation. Furthermore, the method exhibits strong stability and reproducibility. These results confirm the feasibility and precision of

using the derived vehicle pair probabilities to generate mixed traffic flows with target spatial distributions of CAVs under various platoon intensity settings.

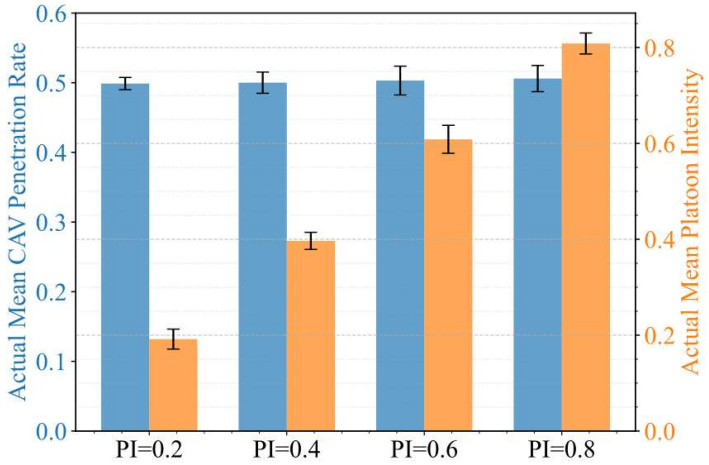


Fig.3 The average data and standard deviations of the simulated mixed traffic flow

3.2.3 Range of platoon intensity

In mixed traffic flow, platoon intensity serves as a key metric for quantifying the clustering degree of CAVs. Its value is not only determined by the spatial distribution patterns of HDVs and CAVs, but is also influenced by the total number of vehicles and the CAV penetration rate. To isolate the effect of spatial distribution from other traffic parameters and focus solely on its impact, it is essential to study the feasible range of platoon intensity under fixed traffic conditions. This allows for a clearer understanding of how platoon intensity characterises mixed traffic flow and how it evolves with changes in the traffic environment.

The theoretical range of platoon intensity in mixed traffic flow is given as follows:

$$PI = \begin{cases} 0, \ \rho = 0 \\ \left[0, 1 - \frac{1}{\rho N}\right], \ 0 < \rho \le 0.5 \\ \left[2 - \frac{1}{\rho}, 1 - \frac{1}{\rho N}\right], \ 0.5 \le \rho < 1 \\ 1, \ \rho = 1 \end{cases}, \tag{11}$$

where ρ denotes the CAV penetration rate and N is the total number of vehicles:

When $\rho = 0$, there are no CAVs present in the traffic flow, and hence the platoon intensity is 0. When $\rho = 1$, the traffic flow consists entirely of CAVs, and the platoon intensity reaches its maximum value of 1. For intermediate values of ρ, we discuss the minimum and maximum platoon intensity values separately. Intuitively, the platoon intensity reaches its maximum when all CAVs are arranged consecutively, forming a single contiguous group. In this case, the first CAV follows an HDV, and the number of CC vehicle pairs is $\rho N - 1$. Therefore, $PI_{max} = \frac{\rho N - 1}{\rho N} = 1 - \frac{1}{\rho N}$. Determining the minimum platoon intensity requires case-based discussion. When $0 < \rho \le 0.5$, the number of CAVs is smaller than that of HDVs, allowing all CAVs to be fully separated by HDVs. In this situation, no CC vehicle pairs exist, and hence $PI_{min} = 0$.

Lemma 2: *When* $0.5 \le \rho < 1$, *the minimum value of platoon intensity in mixed traffic flow is* $2 - \frac{1}{\rho}$.

Proof:

To minimise platoon intensity, the spatial distribution of vehicles should be arranged such that HDVs maximally separate the CAVs, thereby reducing the number of CC vehicle pairs to the lowest possible level. We begin by introducing the concept of HDV intervals, where each HDV acts as a separator between CAVs, partitioning them into multiple groups. The number of such intervals is $m = (1 - \rho)N$. To achieve the minimum platoon intensity, it is necessary for each HDV interval to contain at least one CAV. Let x_i denote the number of CAVs within the ith HDV interval. In each such interval, the number of CC vehicle pairs is $x_i - 1$, assuming the CAVs are arranged consecutively.

Therefore, the total number of CC vehicle pairs in the entire mixed traffic flow is given by:

$$N_{CC}^{min} = \sum_{i=1}^{m} (x_i - 1) = \sum_{i=1}^{m} x_i - m = \rho N - m, \tag{12}$$

Substituting $m = (1-\rho)N$ into the above expression yields:

$$N_{CC}^{min} = N(2\rho - 1), \tag{13}$$

Therefore,

$$PI_{min} = \frac{N_{CC}^{min}}{N_{CAV}} = \frac{N(2\rho - 1)}{\rho N} = 2 - \frac{1}{\rho}, \tag{14}$$

Based on the non-negativity constraint of the four vehicle pair probabilities derived in Eq. (9), the relationship $PI \geq 2 - \frac{1}{\rho}$ can be verified.

Lemma 2 proof ends.

To investigate the distributional characteristics of platoon intensity under fixed traffic conditions, we employed an exhaustive permutation approach using computer simulations. For a mixed traffic flow consisting of 10 vehicles, all possible spatial arrangements of CAVs were systematically generated. The corresponding platoon intensity for each arrangement was then calculated. Figure 4 presents the full distribution of platoon intensity values and their occurrence probabilities under nine different CAV penetration rates. Firstly, the observed ranges of platoon intensity values in all scenarios are consistent with the theoretical bounds defined in Eq. (11), confirming the applicability of the proposed bounds in small-scale mixed traffic systems. Additionally, as the CAV penetration rate increases, the probability of observing a platoon intensity of zero—i.e., where no CAVs are adjacent—declines significantly. For instance, when the penetration rate is 0.1, this probability is 100%; however, it progressively decreases and approaches zero as the penetration rate rises. Moreover, the overall distribution exhibits a rightward shift, indicating a tendency towards higher platoon intensities. In moderate penetration scenarios, the distribution becomes approximately normal. It is important to note that the range and distribution of platoon intensity are highly sensitive to the total number of vehicles. Different traffic scales yield distinct distribution patterns; however, the general trend of increasing platoon intensity with rising CAV penetration remains consistent. Therefore, despite variations in traffic size, platoon intensity serves as a robust and generalisable metric for quantifying spatial distribution characteristics in mixed traffic flow.

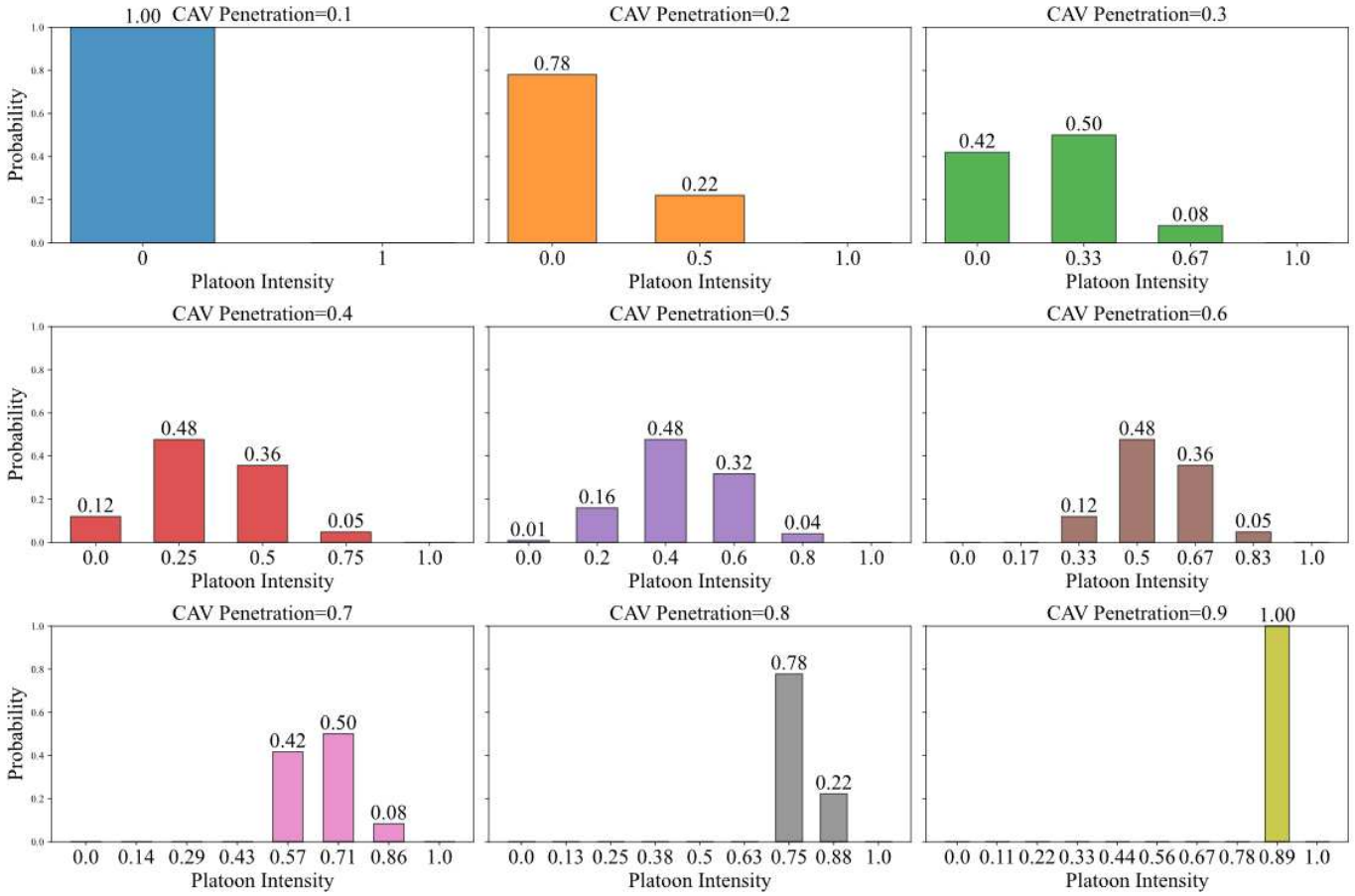


Fig. 4 The probability distribution of platoon intensity under different CAV penetration rates

3.3. Stochastic car-following model in mixed traffic flow

3.3.1 Stochastic car-following model of HDVs

In this study, the stochastic optimal velocity model (SOVM) is employed to represent the car-following behaviour of HDVs. The original optimal velocity model (OVM) calculates a vehicle's desired speed at the next time step based on the current headway and velocity. Its high sensitivity to inter-vehicle spacing enables it to perform effectively in the study of traffic congestion and oscillations [50]. The SOVM extends this by incorporating a stochastic differential equation, adding a Wiener process—dependent on headway—to the acceleration term, thereby capturing the random nature of HDV car-following behaviour [11]. The car-following model of HDVs is given as:

$$\begin{cases} \dfrac{dv_n(t)}{dt} = \beta \cdot \left[V_{op}\left(s_n(t)\right) - v_n(t)\right] + \mu \cdot \sqrt{s_n(t)} \cdot \dfrac{dW(t)}{dt} \\ V_{op}\left(s_n(t)\right) = \dfrac{1}{2} v_f \left[\tanh\left(\dfrac{s_n(t)}{s_0} - \gamma\right) + \tanh(\gamma)\right] \\ s_n(t) = p_{n-1}(t) - p_n(t) - l \\ a_n(t+\Delta t) = \dfrac{v_n(t+\Delta t) - v_n(t)}{\Delta t} \end{cases}, \tag{15}$$

where $v_n(t)$ denotes the velocity of vehicle n at time t, and β is the velocity sensitivity coefficient. V_{op} is the optimal velocity function, v_f is the free-flow speed, s_0 is the safety spacing, and γ is a dimensionless parameter. The spacing between vehicle n and its leader $n-1$ at time t is denoted by $s_n(t)$. $p_n(t)$ is the position of vehicle n, and l is the vehicle length (set to 5 m). μ represents the stochastic term coefficient, and $W(t)$ is a standard Wiener process. Δt is the time step, and $a_n(t+\Delta t)$ is the acceleration at the next time step. The parameter values used in this study are adopted from the calibration of the SOVM using the NGSIM dataset, as proposed by Mao et al. [11]: β=0.93, μ=0.2, v_f=30.63 m/s, s_0=12.14 m, γ=1.91.

A numerical approach is subsequently adopted to solve the stochastic differential equation (SDE)

governing the SOVM in order to obtain the vehicle velocity. The acceleration SDE includes both a deterministic term, $\beta \cdot [V_{op}(s_n(t)) - v_n(t)]$, and a stochastic term, $\mu \cdot \sqrt{s_n(t)} \cdot \frac{dW(t)}{dt}$. To facilitate discretisation, the equation is reformulated as follows:

$$dv_n(t) = \beta \cdot [V_{op}(s_n(t)) - v_n(t)] \cdot dt + \mu \cdot \sqrt{s_n(t)} \cdot dW(t). \tag{16}$$

By integrating the equation from time t to $t+1$, the left side $dv(t)$ becomes:

$$\int_t^{t+1} dv(t) = v(t + \Delta t) - v(t), \tag{17}$$

while the deterministic term is approximated using the left-endpoint value at time t multiplied by the time step Δt, yielding:

$$\int_t^{t+1} \beta \cdot [V_{op}(s_n(t)) - v_n(t)] \cdot dt \approx \beta \cdot [V_{op}(s_n(t)) - v_n(t)] \cdot \Delta t. \tag{18}$$

For the stochastic term:

$$\int_t^{t+1} \mu \cdot \sqrt{s_n(t)} \cdot dW(t) \approx \mu \cdot \sqrt{s_n(t)} \cdot [w(t+1) - w(t)], \tag{19}$$

where the increment $w(t+1) - w(t) \sim N(0, \Delta t)$, can be expressed as $\sqrt{\Delta t} \cdot w(t)$ $(w(t) \sim N(0,1))$, hence:

$$\int_t^{t+1} \mu \cdot \sqrt{s_n(t)} \cdot dW(t) \approx \mu \cdot \sqrt{s_n(t)} \sqrt{\Delta t} \cdot w(t). \tag{20}$$

By combining these components, the approximate discrete-time solution of the acceleration SDE is given by:

$$v_n(t + \Delta t) = v_n(t) + \beta \cdot [V_{op}(s_n(t)) - v_n(t)] \cdot \Delta t + \mu \cdot \sqrt{s_n(t)} \sqrt{\Delta t} \cdot w(t). \tag{21}$$

The optimal velocity function V_{op} used in this study takes the form of a sigmoidal hyperbolic tangent function, originally proposed by Bando et al. [50]. The free-flow speed v_f sets the upper limit of the function, while the parameter s_0 controls the steepness of the transition, and γ determines the inflection point—i.e., where the speed starts to rise significantly as the vehicle spacing increases. As illustrated in Figure 5, the red curve shows the relationship between vehicle spacing and the optimal velocity for the parameter settings adopted in this study. A decrease in the value of s_0 results in a steeper curve. When the value of γ is increased, the inflection point of the curve lags.

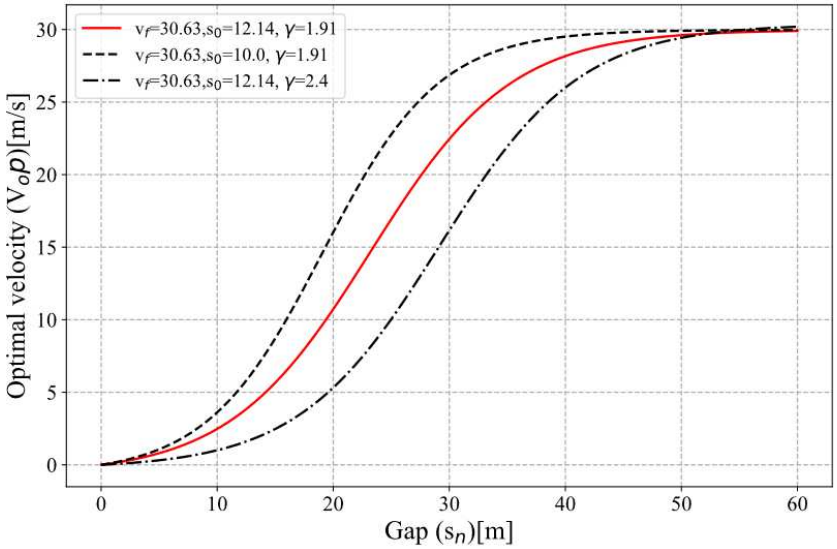


Fig. 5 Curves of speed function V_{op} under different parameters

3.3.2 Car-following model of FCAV

For FCAVs in mixed traffic flow—i.e., CAVs that follow other CAVs within a platoon—their car-following behaviour is modelled using the Cooperative Adaptive Cruise Control (CACC) model. The

behaviour of FCAVs is deterministic. The CACC model, developed by the PATH laboratory in the United States [51], adjusts a CAV's following speed based on vehicle-to-vehicle (V2V) communication and spacing error. The car-following behaviour of an FCAV is described as follows:

$$\begin{cases} v_n(t+\Delta t) = v_n(t) + k_p e_n(t) + k_d \dot{e_n}(t) \\ e_n(t) = p_{n-1}(t) - p_n(t) - l - d - T_{CACC} v_n(t) \\ a_n(t) = \dfrac{k_p\left(x_{n-1}(t) - x_n(t) - l - d - T_{CACC} v_n(t)\right) + k_d \Delta v_n(t)}{\Delta t + k_d T_{CACC}} \end{cases}, \tag{22}$$

where $v_n(t+\Delta t)$ denotes the speed of vehicle n at time $t+\Delta t$, and Δt is the simulation time step. The parameters k_p and k_d are the model parameters of the CACC controller, set to 0.45 s^{-2} and 0.25 s^{-1}, respectively. $e_n(t)$ represents the spacing error at time t, defined as the difference between the desired spacing and the actual spacing. $p_n(t)$ denotes the position of vehicle n at time t, l is the vehicle length, set to 5 m, and d is the standstill safety distance, set to 2 m. The time headway T_{CACC} for FCAVs using the CACC mode is set to 0.8 s. $a_n(t)$ is the acceleration of vehicle n at time t.

3.3.3 Car-following model of DCAV and LCAV

The Intelligent Driver Model (IDM) [52] is employed to describe the car-following behaviour of degraded CAVs (DCAVs)—i.e., CAVs following HDVs—and leading CAVs (LCAVs) that initiate a new platoon after reaching the maximum platoon size. These vehicles exhibit deterministic behaviour. The IDM is a car-following model derived from the perspective of statistical physics, characterised by a small number of parameters with clear physical meanings. It is particularly well-suited to capturing the car-following dynamics of CAVs. The car-following behaviour of DCAVs and LCAVs is described as follows:

$$\begin{cases} a_n(t) = A\left[1 - \left(\dfrac{v_n(t)}{v_f}\right)^{\delta} - \left(\dfrac{s_n^*(t)}{s_n(t)}\right)^2\right] \\ s_n^*(t) = d + v_n(t) T_{DCAV} + \dfrac{v_n(\mathrm{t})[v_n(t) - v_{n-1}(t)]}{2\sqrt{Ab}} \end{cases} \tag{23}$$

where $a_n(t)$ denotes the desired acceleration of vehicle n at time t, A is the maximum acceleration (set to 2 m/s^2), and b is the comfortable deceleration (set to –1 m/s^2). $v_n(t)$ is the speed of vehicle n at time t, and v_f is the free-flow speed, set to 33.3 m/s. The exponent δ, which controls the acceleration behaviour, is set to 4. $s_n^*(t)$ is the desired spacing of vehicle n at time t, $s_n(t)$ is the actual bumper-to-bumper spacing between vehicle n and its predecessor. d is the minimum standstill spacing, set to 2 m, and T_{DCAV} is the desired time headway for DCAVs and LCAVs, set to 1.1 s.

3.4. Mixed traffic flow performance indicators

3.4.1 Traffic efficiency

Traffic efficiency is one of the key indicators used to evaluate the operational state of a transport system, and it is typically measured by the average vehicle speed. In mixed traffic environments, the behaviour of CAVs and HDVs significantly influences overall flow performance. Average speed reflects not only the degree of road capacity utilisation but also provides an indirect indication of congestion levels. Generally, due to their faster reaction times and shorter headways, CAVs can alleviate congestion and improve flow rates to some extent. However, in mixed traffic flows with stochastic HDV behaviour, the impact of different spatial distributions of vehicles on traffic efficiency remains insufficiently studied. By comparing average speeds across various CAV penetration rates and spatial configurations, the operational efficiency of the mixed traffic system can be effectively evaluated. The average speed is defined as:

$$\bar{v} = \frac{1}{NK} \sum_{i=1}^{N} \sum_{k=1}^{K} v_{i,k}\,, \tag{24}$$

where $\bar{v}$ denotes the average speed of the mixed traffic flow over the measurement period, N is the total number of vehicles, K is the total number of time steps during the measurement, and $v_{i,k}$ represents the

instantaneous speed of vehicle i at time step k.

3.4.2 Instability

Due to the stochastic nature of HDV car-following behaviour, backward-propagating traffic oscillations may emerge in the flow. CAVs are expected to suppress such oscillations through deterministic car-following and precise control mechanisms. To assess the damping effects of different spatial configurations of CAVs on traffic oscillations, this study employs the coefficient of variation of vehicle speed as a measure of traffic flow instability. This metric reflects the degree of speed fluctuation within the traffic flow—larger fluctuations indicate greater instability. By comparing the coefficient of variation under different conditions, we can gain deeper insights into the mechanisms by which CAV spatial distribution influences traffic oscillations. Compared to the standard deviation of speed, the coefficient of variation eliminates the influence of measurement scale and better captures the instability of traffic flow. The coefficient of variation of vehicle speed is expressed as:

$$\bar{c_v} = \frac{\sqrt{\frac{1}{NK-1}\sum_{i=1}^{N}\sum_{k=1}^{K}\left(v_{i,k}-\bar{v}\right)^2}}{\bar{v}}, \tag{25}$$

where $\bar{c_v}$ denotes the coefficient of variation of vehicle speed, $v_{i,k}$ is the instantaneous speed of vehicle i at time step k, and $\bar{v}$ is the average speed of the traffic flow.

3.4.3 Energy consumption

Energy consumption is a key parameter for evaluating the sustainability and environmental impact of transport systems. In this study, the Normalised Fuel Consumption Rate (NFR) is employed as the metric for measuring energy consumption [53]. The NFR is calculated based on instantaneous vehicle acceleration, speed, and model parameters, offering an accurate representation of energy use under specific traffic conditions. In mixed traffic environments characterised by traffic oscillations, CAVs—owing to their superior driving strategies and cooperative control capabilities—have the potential to reduce overall system energy consumption. However, the relationship between platoon intensity and energy use in oscillatory flows caused by the stochastic behaviour of HDVs remains unclear. By comparing the NFR across different levels of platoon intensity, the effectiveness of various spatial distributions of CAVs in reducing energy consumption can be revealed. The NFR is derived from the Vehicle Specific Power (VSP), which is defined as:

$$VSP_{i,k} = v_{i,k}\cdot\left(1.1\cdot a_{i,k}+0.132\right)+0.000302\cdot v_{i,k}^3, \tag{26}$$

where $VSP_{i,k}$ is the Vehicle Specific Power of vehicle i at time step k, measured in kW/ton, and $a_{i,k}$ is the instantaneous acceleration of vehicle i at time step k.

The NFR is then calculated using the following piecewise function:

$$NFR_{i,k} = \begin{cases} 1.0\,, & VSP_{i,k} \le 0 \\ 1.71\cdot VSP_{i,k}^{0.42}\,, & VSP_{i,k} > 0 \end{cases}, \tag{27}$$

where $NFR_{i,k}$ denotes the Normalised Fuel Consumption Rate of vehicle i at time step k.

The average NFR of the traffic flow over the measurement period is given by:

$$\overline{NFR} = \frac{1000}{\bar{v}} \times \frac{1}{NK}\sum_{i=1}^{N}\sum_{k=1}^{K} NFR_{i,k}\,. \tag{28}$$

In particular, to eliminate the influence of average speed, the result is divided by $\bar{v}$, yielding the average per-kilometre fuel consumption, expressed in grams per kilometre (g/km).

4. Numerical experiment

4.1. Experimental setting

Due to the current immaturity of autonomous driving technologies in the commercial vehicle sector, large-scale field studies on mixed traffic flow are difficult to implement in real-world scenarios in the short

term. Meanwhile, conducting physical experiments in closed test facilities imposes stringent requirements on environmental control and hardware resources, resulting in high implementation costs and operational complexity. As a result, most contemporary studies on CAVs predominantly adopt simulation-based approaches to obtain stable, controllable, and representative experimental data. In this study, a mixed traffic flow simulation platform was developed using Python. A single-lane ring road was constructed to emulate the operation of traffic on an infinitely long road. Owing to its closed boundary conditions and simple structure, the ring road has been widely used in studies on traffic oscillation propagation and traffic flow dynamics, making it particularly well-suited for the present analysis of traffic instability. The simulation setup follows the classic experimental configurations used by Stern et al. [35] and Sugiyama et al. [54] in their real-vehicle experiments on inducing traffic oscillations. A total of 15 vehicles were evenly distributed along the ring road, with an initial inter-vehicle spacing of 15 m and an initial speed of 20 m/s. The total simulation duration was set to 450 seconds with a time step of 0.1 seconds. To eliminate the influence of initial disturbances, the first 50 seconds of the simulation were designated as a warm-up phase, and all performance indicators were measured over the interval from 50 to 450 seconds. The instantaneous acceleration of each vehicle was dynamically updated based on the car-following models described in the previous section, with parameter values listed in Table 2.

Table 2 Parameter settings

Parameter	Value	Definition	Parameter	Value	Definition
l	5 m	Uniform vehicle length	A	2 m/s²	Maximum acceleration of CAVs
β	0.93	Speed sensitivity coefficient for HDVs	b	-1 m/s²	Comfortable deceleration of CAVs
μ	0.2	Coefficient of Wiener process for HDVs	v_{max}	33.3 m/s	Free-flow speed of CAVs
v_f	30.63 m/s	Free-flow speed of HDVs	δ	4	Acceleration exponent for CAVs
s_0	12.14 m	Desired safety gap for HDVs	T_{DCAV}	1.1 s	Desired time headway for DCAVs
γ	1.91	Optimisation speed function constant for HDVs	N	15	Total number of vehicles
k_p	0.45	Proportional control gain for FCAVs	K	4000	Total number of time steps during measurement
k_d	0.25	Derivative control gain for FCAVs	Δt	0.1 s	Simulation time step
d	2 m	Standstill gap for CAVs	v_0	20 m/s	Initial vehicle speed in simulation
T_{CACC}	0.8 s	Desired time headway for FCAVs	S_{max}	4	Maximum platoon size for CAVs

4.2. Reproduction of traffic oscillations in human-driven traffic flow

To verify the effectiveness of the stochastic car-following model in reproducing traffic oscillations induced by randomness, a simulation was first conducted under a purely human-driven traffic environment in which all 15 vehicles were set as HDVs. Using the position, time, and speed data of the 15 HDVs on the ring road, a spatiotemporal trajectory diagram and a speed heatmap were generated, as shown in Figure 6. Evident oscillations in both vehicle trajectories and speeds can be observed from the diagrams. These oscillations continuously propagate upstream along the traffic flow. Under conditions with no lane changing or external disturbances, vehicles exhibit repeated stop-and-go behaviour throughout the simulation. This result indicates that the stochastic car-following model adopted in this study can effectively capture the endogenous oscillatory characteristics of traffic flow, demonstrating strong physical plausibility and modelling capability. It lays a solid foundation for subsequent investigations into the effects of vehicle spatial distribution on the performance of mixed traffic flow within this modelling framework.

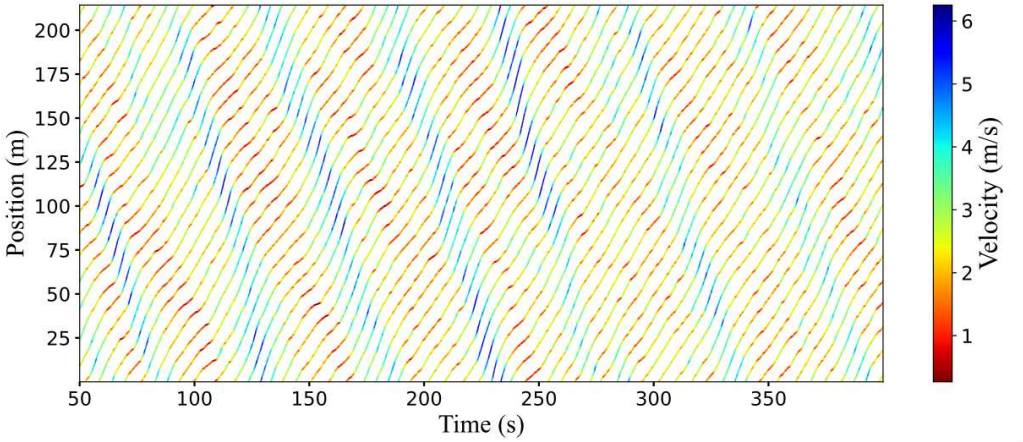
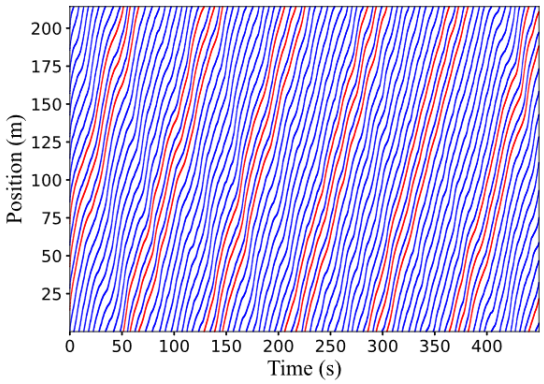


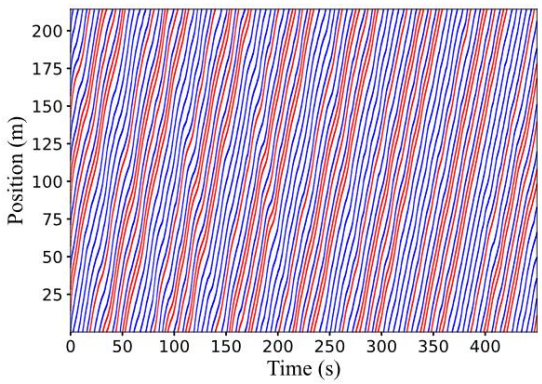
Fig. 6 Traffic oscillations caused by stochasticity of HDVs

4.3. Mixed traffic flow experiments

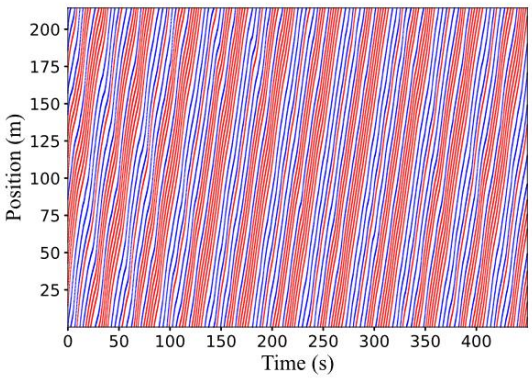
To investigate the impact of CAVs and their spatial distribution on the overall performance of mixed traffic flow under the influence of HDV stochasticity, a large number of simulation experiments were conducted in this study. Considering different CAV penetration rates, a total of 10,922 possible vehicle distribution patterns were generated for a 15-vehicle scenario. Each of these configurations was simulated independently on a single-lane ring road. During the simulations, instantaneous position, speed, and acceleration data for all vehicles were recorded. Based on these data, key performance indicators including traffic efficiency, instability, and energy consumption were calculated to evaluate how different spatial distributions of CAVs affect mixed traffic performance. Figure 7 presents spatiotemporal trajectory diagrams under representative CAV penetration levels. It can be observed that the blue trajectories, representing HDVs, still exhibit strong randomness in their car-following behaviour. These results demonstrate that endogenous traffic oscillations persist across various CAV penetration scenarios, highlighting the continued influence of HDV-induced stochasticity in mixed traffic environments.



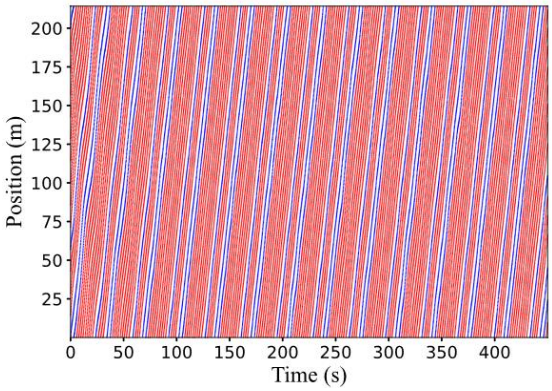
(a)20%



(b)40%



(c)60%



(d)80%

Fig. 7 Mixed traffic flow trajectory diagram under different CAV penetration

4.3.1 Impact of vehicle spatial distribution on traffic efficiency

To investigate traffic efficiency in mixed traffic flow, the average speed is employed as the performance metric. Figure 8 presents a scatter plot of platoon intensity versus average speed across all vehicle distribution scenarios, with different colours representing different CAV penetration rates. A clear positive correlation is observed between platoon intensity and average speed, with a Pearson correlation coefficient of $r = 0.747$. This suggests that an increase in platoon intensity generally leads to improved average speed. Specifically, when platoon intensity is 0, the average speed falls below 2.5 m/s, whereas when platoon intensity reaches 0.8, the average speed exceeds 2 m/s. This reflects the effect of spatial aggregation: when CAVs are more closely grouped, they are able to travel cooperatively in the form of tightly connected platoons, thereby enhancing overall traffic efficiency. The coloured scatter points indicate different levels of CAV penetration, and the convex hull of data points with the same penetration rate represents the overall range of observations. It is evident that the range of platoon intensity shifts to the right as the CAV penetration rate increases, which aligns with the theoretical derivation of platoon intensity ranges previously discussed. Moreover, a general upward trend in average speed is observed with increasing CAV penetration rates.

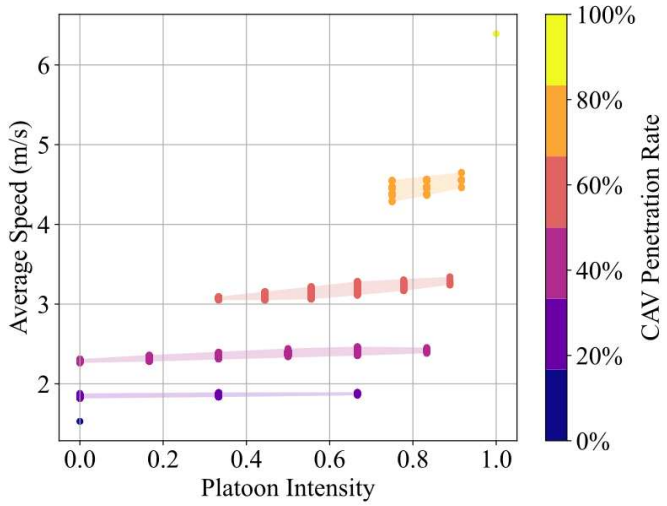
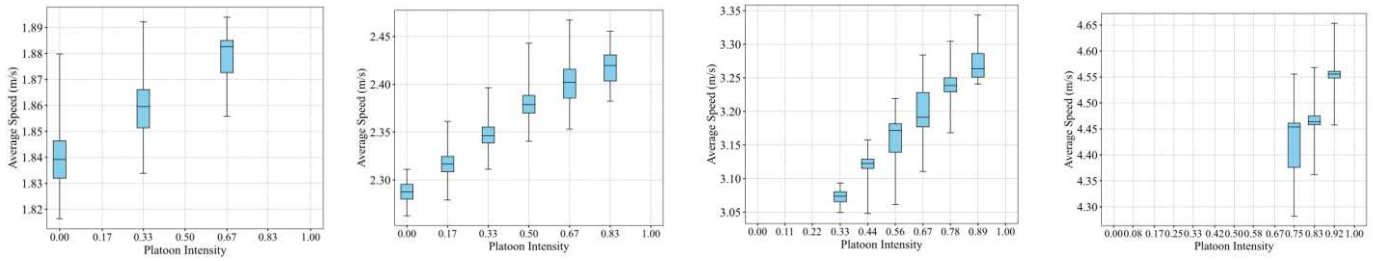


Fig. 8 Heatmap-style scatter plot of average speed under different platoon intensities

Given that platoon intensity is highly dependent on CAV penetration rate, we isolate data under identical penetration levels to decouple the effect of penetration from that of spatial distribution. Boxplots are drawn to illustrate this relationship more clearly. These boxplots show the maximum, minimum, interquartile range, and median of average speeds under various platoon intensity at fixed CAV penetration levels. As shown in Figure 9, the boxplots for four different CAV penetration rates (20%, 40%, 60%, and 80%) exhibit the same pattern: average speed tends to increase with higher platoon intensity. To clearly display this trend, the x-axis of each subplot in Figure 9 is fixed to the full platoon intensity range (0–1), while the y-axis is scaled independently for each penetration level. Although overlap in average speed values exists across different platoon intensity within the same penetration level, the upward trend remains evident. This confirms that, under the same CAV penetration rate, more clustered CAV distributions are beneficial to traffic efficiency. The underlying mechanism is that when CAVs are more spatially concentrated, they can form cooperative platoons using V2V communication technology. Vehicles within such platoons share synchronised motion data and are able to travel with smaller headways, thus increasing the average speed of the traffic flow.



(a)20% (b)40% (c)60% (d)80%

Fig. 9 Boxplots of average speed against platoon intensity under four CAV penetration levels

Table 3 presents the comparison of the maximum and minimum average speeds under different CAV penetration levels. In all four penetration scenarios, the highest average speeds are associated with higher platoon intensity (i.e., more clustered CAV distributions), while the lowest average speeds occur in low platoon intensity scenarios (i.e., more dispersed CAVs). However, for the 40% penetration case, the highest average speed is observed at a platoon intensity of 0.67, rather than the maximum possible value of 0.83; similarly, for 60% penetration, the lowest average speed occurs at a platoon intensity of 0.44, not the minimum of 0.33. This discrepancy may be attributed to the stochastic behaviour of HDVs. When the difference in platoon intensity between two configurations is relatively small, the influence of vehicle distribution on traffic efficiency may be masked by the randomised following behaviour of HDVs. In terms of magnitude, the difference between the highest and lowest average speeds at the same penetration level can be as much as 0.3712 m/s, with the relative difference reaching up to 9.70%. This highlights the critical role that spatial distribution plays in determining traffic efficiency in mixed traffic flow.

Table 3 Comparison of optimal and worst cases for average speed

CAV penetration	Case	Platoon intensity	Vehicles spatial distribution	Value	Difference	Percentage of difference
20%	Optimal	0.67	(0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)	1.894	0.0776	4.27%
	Worst	0	(0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0)	1.8164		
40%	Optimal	0.67	(0, 0, 0, 1, 1, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0)	2.4672	0.2046	9.04%
	Worst	0	(1, 0, 1, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0)	2.2626		
60%	Optimal	0.89	(1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1)	3.3436	0.2956	9.70%
	Worst	0.44	(1, 1, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 1)	3.048		
80%	Optimal	0.92	(1, 1, 1, 1, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1)	4.6535	0.3712	8.67%
	Worst	0.75	(1, 0, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1)	4.2823		

4.3.2 Impact of vehicle spatial distribution on the instability

The coefficient of variation of vehicle speed is adopted to evaluate the instability of mixed traffic flow. Figure 10 presents the scatter plot of platoon intensity versus the speed variation coefficient for all vehicle distribution scenarios, where different colours represent different CAV penetration rates. Overall, there is no significant linear relationship between platoon intensity and the speed variation coefficient, with a Pearson correlation coefficient of $r = -0.481$. A noticeable reduction in the minimum value of the speed variation coefficient only appears when the platoon intensity approaches 1. Further examination of the convex hulls under various CAV penetration rates reveals that at 80% CAV penetration, the distribution of the speed variation coefficient is more reduced and convergent, indicating that a higher CAV presence contributes to enhanced control over traffic flow instability. However, since CAV penetration and platoon intensity change simultaneously in Figure 10, and are inherently coupled, it is necessary to further investigate the influence of platoon intensity under fixed CAV penetration conditions.

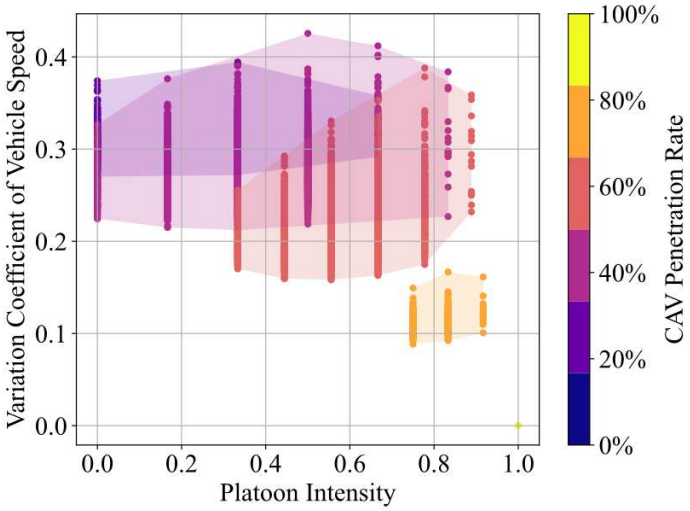
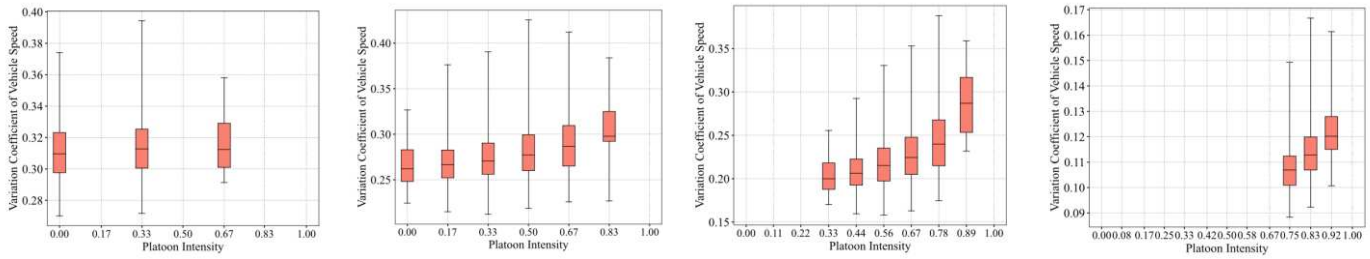


Fig. 10 Heatmap-style scatter plot of speed variation coefficient under different platoon intensities

Figure 11 illustrates the boxplots of the speed variation coefficient under four typical CAV penetration rates (20%, 40%, 60%, and 80%) across various levels of platoon intensity. The results indicate that, under fixed penetration rates, the speed variation coefficient tends to increase as platoon intensity rises—suggesting that the more clustered the CAVs are, the higher the instability of the mixed traffic flow. Conversely, a more dispersed CAV distribution corresponds to weaker speed fluctuations and enhanced stability. This phenomenon arises primarily from the interaction between the deterministic car-following behaviour of CAVs and the stochastic driving behaviour of HDVs. When CAVs are more dispersed, a larger number follow behind HDVs. Their shorter reaction time and reduced headway allow them to absorb or dampen HDV-induced acceleration and deceleration disturbances more effectively, thereby mitigating the propagation of traffic oscillations and improving the overall flow stability. In contrast, when CAVs are densely clustered, fewer follow HDVs, weakening this stabilising effect.



(a)20% (b)40% (c)60% (d)80%

Fig. 11 Boxplots of speed variation coefficient against platoon intensity under four CAV penetration levels

Additionally, Table 4 compares the maximum and minimum values of the speed variation coefficient across different CAV penetration levels. In all cases, the lowest instability corresponds to lower platoon intensities (i.e., more dispersed CAV distributions), while the highest instability appears under higher platoon intensities (i.e., more aggregated CAV distributions). It is worth noting, however, that the maximum speed variation coefficient does not always occur at the absolute highest platoon intensity, suggesting that the random behaviour of HDVs can still exert interference at extreme ends, occasionally masking the effect of spatial distribution. Numerically, the difference between the maximum and minimum values of the speed variation coefficient reaches up to 0.2297, with a maximum relative difference of 145.20%, further confirming that vehicle spatial distribution has a significant impact on the instability of mixed traffic flow. This also implies that, under a given CAV penetration rate, optimising the spatial layout of CAVs can effectively enhance traffic flow stability.

Table 4 Comparison of optimal and worst cases for the speed variation coefficient

CAV penetration	Case	Platoon intensity	Vehicles spatial distribution	Value	Difference	Percentage of difference
20%	Optimal	0	(1, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0)	0.2701	0.1243	46.02%
	Worst	0.33	(1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1)	0.3944		
40%	Optimal	0.3	(1, 1, 1, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0)	0.2121	0.2134	100.61%
	Worst	0.50	(0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 1, 1, 1, 0, 0)	0.4255		
60%	Optimal	0.56	(1, 0, 1, 1, 1, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1)	0.1582	0.2297	145.20%
	Worst	0.78	(0, 0, 0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1)	0.3879		
80%	Optimal	0.75	(1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 0, 1)	0.0884	0.0783	88.57%
	Worst	0.83	(1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1)	0.1667		

4.3.3 Impact of vehicle spatial distribution on energy consumption

Figure 12 presents the scatter plot of platoon intensity versus average energy consumption under all vehicle distribution conditions, with colour indicating the CAV penetration rate. Overall, a significant negative correlation is observed between platoon intensity and average energy consumption, with a Pearson correlation coefficient of $r = -0.764$. This indicates that as platoon intensity increases, the average energy consumption tends to decrease. When platoon intensity equals 0, the average energy consumption exceeds 700 g/km in all cases, whereas when platoon intensity exceeds 0.8, it consistently falls below 800 g/km. This suggests that when CAVs are more spatially clustered in the traffic flow, the resulting cooperative CAV platoons play a positive role in reducing energy consumption. Furthermore, an overall decreasing trend in energy consumption is observed as the CAV penetration rate increases.

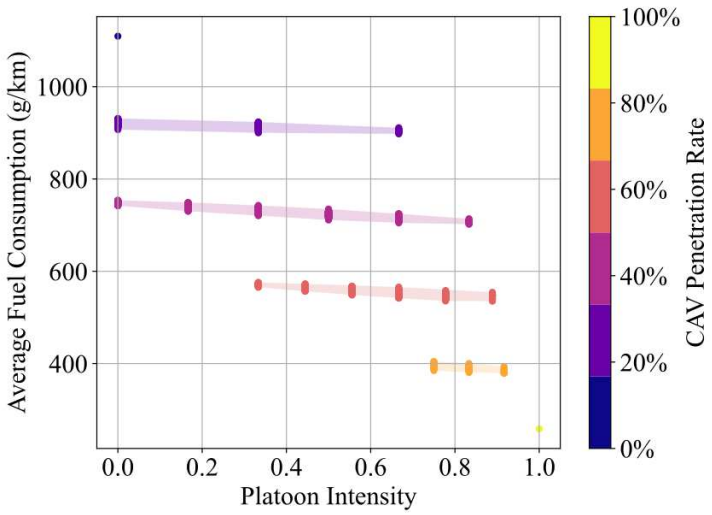
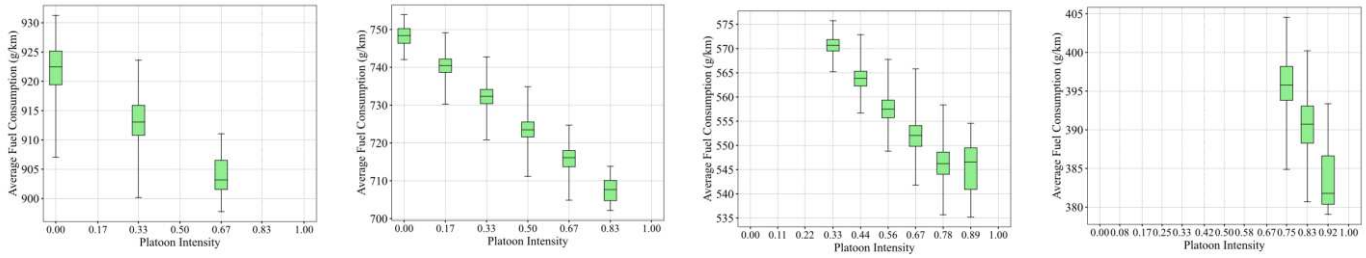


Fig. 12 Heatmap-style scatter plot of average energy consumption under different platoon intensities.

Figure 13 shows the boxplots of average energy consumption across four CAV penetration rates (20%, 40%, 60%, and 80%), under varying platoon intensities**.** A consistent trend can be observed: as platoon intensity increases—i.e. as CAVs become more spatially concentrated—the average energy consumption decreases. This implies that grouping CAVs into platoons is beneficial for reducing energy consumption in mixed traffic conditions. It is worth noting that, unlike average speed or speed variability, average energy consumption is not a direct output of the simulation. Rather, it is a composite indicator derived from each vehicle's instantaneous speed and acceleration, used to compute the VSP, and then indirectly estimated based on the relationship between power and NFR. Thus, average energy consumption reflects the combined influence of both traffic efficiency and driving stability. Mechanistically, when CAVs form platoons, they adopt cooperative control strategies that enhance overall traffic speed while maintaining well-synchronised acceleration behaviour within the platoon. This reduces the frequent acceleration and deceleration caused by uncertainty in the behaviour of preceding vehicles. As a result, although slight increases in speed variability may occur, the fluctuation in acceleration does not necessarily increase, and may in fact decrease, thereby

contributing to lower energy consumption. Therefore, although higher platoon intensity may lead to increased local speed variation, the associated gains in average speed, smoother acceleration profiles, and improved drivetrain efficiency collectively result in reduced energy use. This phenomenon highlights once again the profound impact of vehicle spatial distribution on the performance of mixed traffic flow. In particular, with respect to energy efficiency, the rational aggregation of CAVs into platoons offers substantial advantages.



(a)20% (b)40% (c)60% (d)80%

Fig. 13 Boxplots of average energy consumption against platoon intensity under four CAV penetration levels

Table 5 compares the maximum and minimum values of average energy consumption under different penetration rates. At all four penetration levels, the highest energy consumption corresponds to the lowest platoon intensity (dispersed CAVs), while the lowest energy consumption is observed under the highest platoon intensity (concentrated CAVs). Numerically, the largest absolute difference between the maximum and minimum values reaches 51.7809 g/km, with the highest relative difference being 7.58%. This further confirms the significant impact of vehicle spatial distribution on the average energy consumption of mixed traffic flow.

Table 5 Comparison of optimal and worst cases for average energy consumption.

CAV penetration	Case	Platoon intensity	Vehicles spatial distribution	Value	Difference	Percentage of difference
20%	Optimal	0.67	(1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1)	897.7493	33.521	3.73%
	Worst	0	(0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0)	931.2703		
40%	Optimal	0.83	(0, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0)	702.1213	51.7809	7.37%
	Worst	0	(1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0, 1, 0, 0)	753.9022		
60%	Optimal	0.89	(1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1)	535.2118	40.5498	7.58%
	Worst	0.33	(1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1)	575.7616		
80%	Optimal	0.92	(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1)	379.0531	25.5008	6.73%
	Worst	0.75	(1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1)	404.5539		

4.3.4 Influence of CAV penetration rate on mixed traffic flow

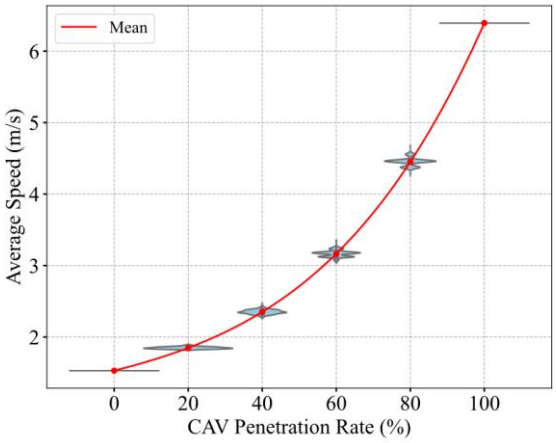
Although previous subsections have examined the impact of varying CAV penetration rates on the performance of mixed traffic flow, it is important to note that CAV penetration and platoon intensity (i.e. the degree of spatial concentration among vehicles) are inherently coupled. Analysing their influence from a single perspective may obscure key details, particularly when comparing how different proportions of CAVs affect traffic stability and efficiency. To further investigate the interaction mechanisms between these factors, Figure 14 plots violin diagrams and mean trend lines for three indicators—average speed, speed variability coefficient, and average energy consumption—across the full set of simulation results, using CAV penetration rate as the horizontal axis. The violin plots illustrate the distribution of each indicator at each given penetration level.

Figures 14(a) and 14(c) show the trends in average speed and average energy consumption, respectively, as the CAV penetration rate increases. It is evident that average speed rises steadily while average energy consumption declines, indicating that the introduction of CAVs has a generally positive effect on improving traffic efficiency and reducing energy usage. Furthermore, the distribution ranges of these two indicators

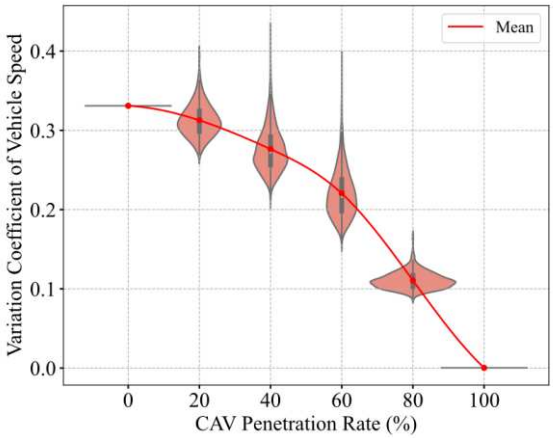
remain relatively narrow under different spatial distributions, as reflected by the smaller areas of the violin plots. This suggests that the influence of CAV penetration rate on these aspects of traffic performance is considerably stronger than that of spatial distribution.

However, the relationship between speed variability coefficient and CAV penetration rate, depicted in Figure 14(b), exhibits a more complex pattern. Although the mean trend demonstrates an overall decline in speed variability with increasing CAV penetration—confirming the stabilising effect of CAVs in suppressing speed fluctuations—the violin plot area is markedly larger, and there is substantial overlap between distributions at different penetration levels. This reveals that spatial distribution of vehicles exerts a stronger disturbance and variation effect on the instability metric. Of particular note, at penetration rates of 20%, 40%, and 60%, while the majority of experiments show significantly reduced speed variability compared to the pure HDV (0% penetration) scenario, a small number of cases still exhibit higher levels of speed fluctuation. This phenomenon can be attributed to the interplay of several factors. Firstly, at low to moderate penetration levels, the number of CAVs may be insufficient to establish a global shock-absorbing chain. If these limited CAVs are also poorly distributed, their presence may disrupt the existing stable car-following structure among HDVs. Secondly, CAVs and HDVs differ fundamentally in their control mechanisms—CAVs typically feature shorter reaction times and more responsive acceleration. In the absence of coordinated platooning control, these characteristics can induce abrupt local speed changes, which are rapidly propagated to following HDVs, thereby exacerbating local instability. Moreover, HDV driving behaviour is inherently stochastic; even with identical vehicle orderings on a ring road, differences in individual driver behaviour alone can result in significant variations in speed standard deviation, masking some of the structural effects.

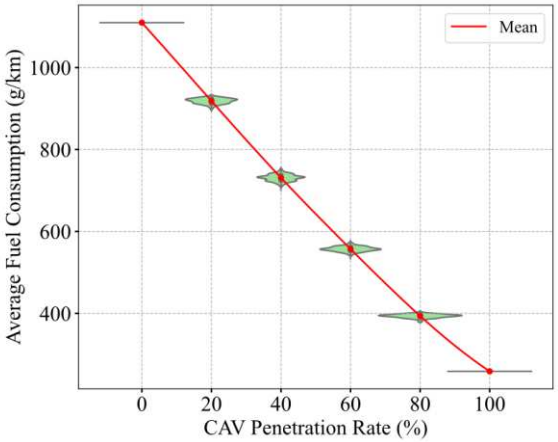
In summary, while increasing the CAV penetration rate is broadly beneficial for improving traffic efficiency and reducing energy consumption, its effectiveness in mitigating traffic instability is more contingent upon the specific spatial distribution of vehicles and the nature of disturbances introduced by HDVs. Only at high penetration rates do CAVs consistently enhance traffic stability. This underscores the complex interplay of multiple interacting factors within mixed traffic systems.



(a) Average speed



(b) Speed variability coefficient



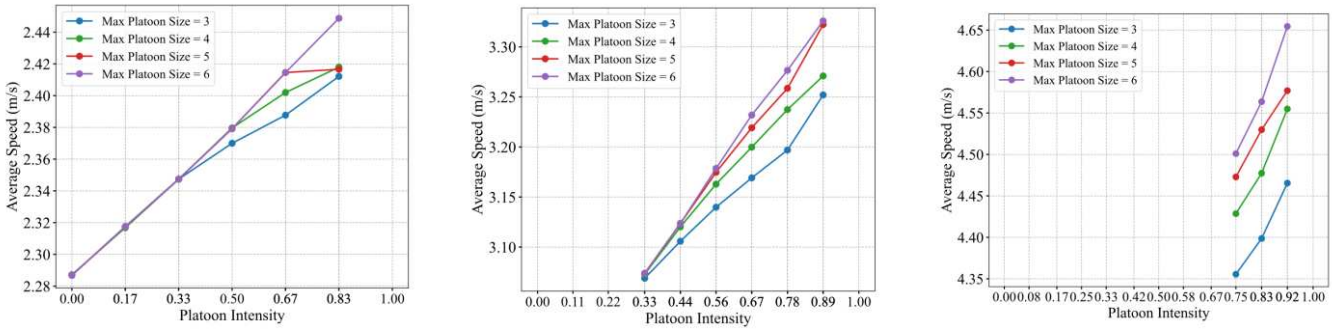
(c) Average energy consumption

Fig. 14 Influence of CAV penetration rate on mixed traffic flow under three indicators

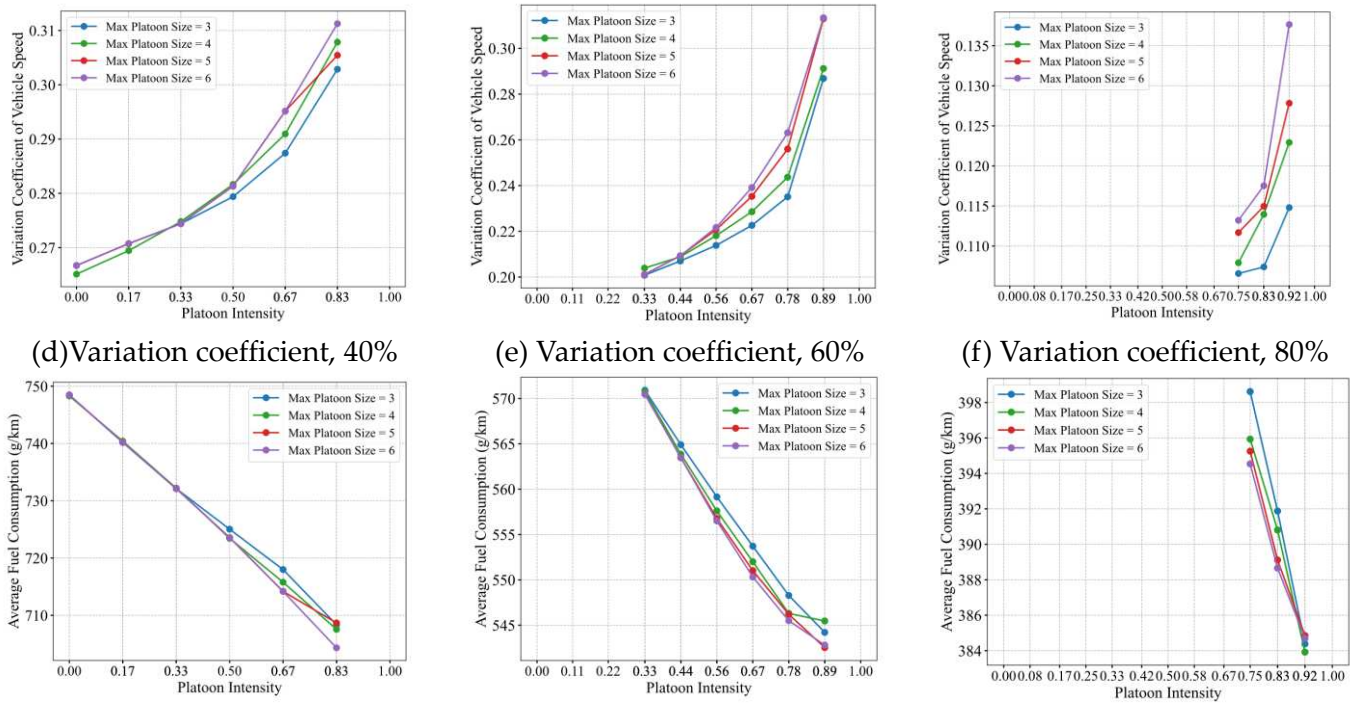
4.4. Sensitivity analysis of maximum platoon size

In practice, inter-vehicle communication among CAVs is constrained by factors such as communication range, signal errors, and system stability. As a result, research on CAV platoons often introduces an artificially defined maximum platoon size. Once the number of CAVs within a platoon reaches this upper limit, any subsequent CAVs are prevented from joining the existing platoon and must instead form a new one. In the context of this study, the setting of the maximum platoon size does not affect the value or distribution range of the platoon intensity metric. This is because platoon intensity, which quantifies the degree of spatial clustering among vehicles, is inherently defined by the number of CAVs following other CAVs, regardless of whether they are formally classified within the same CAV platoon. Therefore, even if the imposed limit on platoon size causes some otherwise continuous sequences of CAVs to be split into multiple platoons, their contributions to platoon intensity remain intact as long as the CAV-following-CAV relationship is preserved. Consequently, for a given spatial distribution of vehicles, the calculated platoon intensity remains consistent regardless of the maximum platoon size, demonstrating the robustness and validity of this metric in capturing spatial structure.

Nevertheless, conducting a sensitivity analysis on the maximum platoon size is still of significant importance. While the value of platoon intensity is unaffected by the maximum platoon size, the performance of mixed traffic flow under different spatial configurations may still vary when the maximum platoon size is altered. To investigate this, a series of simulations were conducted under consistent experimental conditions but with varying maximum platoon sizes (3, 4, 5, and 6 vehicles), to examine their impact on key macroscopic traffic flow indicators. Figure 15 illustrates the variation in average speed, speed variability coefficient, and average energy consumption with respect to platoon intensity under different maximum platoon size settings. For each level of platoon intensity, data points represent the mean values across all simulation runs. This sensitivity analysis captures how traffic efficiency, instability, and energy performance respond to changes in maximum platoon size and spatial distribution. It is important to note that, as the experimental setup in this study involves a 15-vehicle mixed traffic flow, scenarios with a CAV penetration rate below 20% contain a maximum of only 3 CAVs. Given that the maximum platoon size parameter ranges from 3 to 6, the system performance in such cases is unaffected by changes in this parameter. Therefore, the sensitivity analysis focuses exclusively on CAV penetration rates of 40%, 60%, and 80%. Moreover, under a 40% penetration rate, platoons of length ≥3 can only emerge when the platoon intensity exceeds 0.33. As a result, data points within the 0–0.33 platoon intensity range in Figures 15(a), 15(d), and 15(g) overlap heavily across different maximum platoon size settings.



(a)Average speed, 40% (b) Average speed, 60% (c) Average speed, 80%



(d)Variation coefficient, 40% (e) Variation coefficient, 60% (f) Variation coefficient, 80%

(g)Energy consumption, 40% (h) Energy consumption, 60% (i) Energy consumption, 80%

Fig. 15 Line chart of sensitivity analysis of maximum platoon size

Figures 15(a–c) present the average speed in mixed traffic flow for the three selected CAV penetration rates. A clear pattern emerges: at multiple levels of platoon intensity, the average speed increases with the maximum platoon size, indicating a consistent effect of maximum platoon size across different spatial distributions. This improvement in speed can be attributed to the greater proportion of CAVs travelling cooperatively; tighter inter-vehicle communication and coordination among CAVs enhance overall traffic flow efficiency. Figures 15(d–f) display the trends in the speed variability coefficient. Similar to average speed, this coefficient also increases with greater maximum platoon size. However, this trend indicates a negative traffic implication: increased instability. This is because longer platoons result in tightly coordinated and highly responsive groups of CAVs. When this intense coordination interacts with surrounding HDVs—which have inherently slower reaction times—speed fluctuations are more readily amplified and propagated among HDVs, leading to the emergence of local oscillations. In other words, while the internal operation of the CAV platoons remains stable, the disturbances they generate may be transmitted more abruptly to adjacent HDVs, thereby reducing the overall system stability. Figures 15(g–i) show the trends in average energy consumption. In most cases, energy consumption decreases with increasing maximum platoon size, particularly at moderate levels of platoon intensity. This indicates that longer CAV platoons can reduce unnecessary acceleration and deceleration through smoother car-following and acceleration control, thereby enhancing energy efficiency. However, when platoon intensity approaches 1—indicating full aggregation of all CAVs—the sensitivity of energy consumption to maximum platoon size diminishes. This may be due to a saturation effect in coordinated control: when all CAVs are already operating in close coordination, the marginal benefit of further increasing the maximum platoon size is reduced.

Overall, the maximum platoon size has a substantial influence on the macroscopic performance of mixed traffic flow. A larger maximum platoon size generally contributes to higher average speeds and lower energy consumption, but it may also introduce increased risk of instability within the system.

5. Conclusions and discussion

This study proposed a mixed traffic flow modelling framework that explicitly considers both the stochastic behaviour of HDVs and the spatial distribution of CAVs. Firstly, the concept of platoon intensity was introduced, and the vehicle adjacency probability and the theoretical bounds of spatial distribution were derived, thereby establishing a theoretical foundation for investigating the spatial structure of mixed traffic flows. Secondly, a stochastic car-following model based on stochastic differential equations was employed

to reproduce endogenous traffic oscillations on a ring road, which arise without lane-changing or abrupt deceleration behaviours. In the experimental section, this study systematically investigates the impact of CAV spatial distribution on traffic performance. Through numerical simulations, the relationships between three key indicators—average speed, coefficient of variation of speed, and average energy consumption—are analysed under varying levels of CAV penetration and platoon intensity. The results reveal a coupling effect between CAV penetration rate and spatial distribution, leading to the following key findings:

（1）Average speed exhibits a strong positive correlation with platoon intensity (Pearson correlation coefficient $r = 0.747$). When CAVs are more spatially concentrated, it is easier to form stable CAV platoons that maintain short headways and synchronised longitudinal behaviour. This coordination substantially improves traffic efficiency; under the same penetration rate, clustered CAV distributions can increase average speed by up to approximately 9.70% compared to more dispersed distributions.

（2）Regarding the coefficient of variation of speed, a general increasing trend is observed with rising platoon intensity under the same penetration level. When CAVs are more aggregated, their interactions with HDVs become less frequent, weakening the stabilising effect on traffic oscillations. Across different CAV penetration rates, similar levels of speed variation can be observed under different spatial distributions, indicating that both CAV penetration and platoon intensity significantly influence traffic flow stability—though the effect is less pronounced at low to medium penetration levels.

（3）Average energy consumption shows a strong negative correlation with platoon intensity ($r = -0.764$). The formation of CAV platoons through cooperative driving has a clear positive effect on reducing fuel consumption. Under the same penetration level, the difference in average energy consumption between the best and worst spatial distributions reaches 51.7809 g/km, representing a relative reduction of 7.58%.

These findings offer important implications for the management of mixed traffic systems in real-world contexts. Merely increasing the proportion of CAVs is insufficient to maximise the systemic benefits of autonomous driving. More critically, CAVs must be guided to form rational spatial configurations. Therefore, future efforts in CAV scheduling and route planning should focus on the development of platoon coordination strategies and formation mechanisms. From a traffic management and infrastructure perspective, dynamic control strategies—such as cooperative car-following zones or dynamic lane allocation systems—may be introduced to fully exploit the advantages of CAVs in mixed traffic conditions.

This study adopts a ring road scenario, which is advantageous for eliminating boundary effects and isolating vehicle interaction dynamics. However, such a simplified environment does not reflect real-world complexities such as signalised intersections, multi-lane geometries, or fluctuating traffic demand, which limits the generalisability of the findings. Future work will explore optimal spatial distribution patterns in more realistic traffic scenarios. Moreover, this study assumes ideal communication conditions among CAVs, excluding communication delays, data loss, or system failures. Future models should incorporate more realistic communication constraints to enhance model validity. Finally, although platoon intensity is adopted here as the primary metric for quantifying spatial distribution, future research could explore additional quantitative indicators to capture the spatial structure of CAVs in mixed traffic more comprehensively.

Acknowledgments

The paper received research funding support from the National Natural Science Foundation of China (72471200), the Sichuan Science and Technology Program (2024NSFSC0179), and the Fundamental Research Funds for the Central Universities (2682025GH023).

References

[1] Y. Wang, Y. Jiang, Y. Wu, and Z. Yao, "Mitigating traffic oscillation through control of connected automated vehicles: A cellular automata simulation," *Expert Syst. Appl.*, vol. 235, p. 121275, Jan. 2024, doi: 10.1016/j.eswa.2023.121275.

[2] J. Shen, J. Zhao, Z. Yu, S. Zheng, and R. Jiang, "The elimination and absorption mechanism of oscillatory motion wave based on jam-absorption driving for mixed traffic flow in intelligent connected environment,"

Phys. Stat. Mech. Its Appl., vol. 664, p. 130485, Apr. 2025, doi: 10.1016/j.physa.2025.130485.
[3] Y. Jiang, T. Ren, Y. Ma, Y. Wu, and Z. Yao, "Traffic safety evaluation of mixed traffic flow considering the maximum platoon size of connected automated vehicles," *Phys. Stat. Mech. Its Appl.*, vol. 612, p. 128452, Feb. 2023, doi: 10.1016/j.physa.2023.128452.
[4] M. W. Levin and S. D. Boyles, "A multiclass cell transmission model for shared human and autonomous vehicle roads," *Transp. Res. Part C Emerg. Technol.*, vol. 62, pp. 103–116, Jan. 2016, doi: 10.1016/j.trc.2015.10.005.
[5] D. Chen, S. Ahn, M. Chitturi, and D. A. Noyce, "Towards vehicle automation: Roadway capacity formulation for traffic mixed with regular and automated vehicles," *Transp. Res. Part B Methodol.*, vol. 100, pp. 196–221, Jun. 2017, doi: 10.1016/j.trb.2017.01.017.
[6] Z. Yao, Y. Wang, B. Liu, B. Zhao, and Y. Jiang, "Fuel consumption and transportation emissions evaluation of mixed traffic flow with connected automated vehicles and human-driven vehicles on expressway," *Energy*, vol. 230, p. 120766, Sep. 2021, doi: 10.1016/j.energy.2021.120766.
[7] B. Zhao, Y. Lin, H. Hao, and Z. Yao, "Fuel Consumption and Traffic Emissions Evaluation of Mixed Traffic Flow with Connected Automated Vehicles at Multiple Traffic Scenarios," *J. Adv. Transp.*, vol. 2022, pp. 1–14, Jan. 2022, doi: 10.1155/2022/6345404.
[8] Y. Jiang, S. Sun, F. Zhu, Y. Wu, and Z. Yao, "A mixed capacity analysis and lane management model considering platoon size and intensity of CAVs," *Phys. Stat. Mech. Its Appl.*, vol. 615, p. 128557, Apr. 2023, doi: 10.1016/j.physa.2023.128557.
[9] J. Zhang, Q. Liu, S. Li, and L. Li, "Unleashing the Power of Connected and Automated Vehicles: A Dedicated Link Strategy for Efficient Management of Mixed Traffic," *IEEE Trans. Intell. Transp. Syst.*, pp. 1–18, 2024, doi: 10.1109/TITS.2024.3367943.
[10] C. Liu, F. Zheng, H. X. Liu, and X. Liu, "Optimizing Mixed Traffic Flow: Longitudinal Control of Connected and Automated Vehicles to Mitigate Traffic Oscillations," *IEEE Trans. Intell. Transp. Syst.*, vol. 26, no. 3, pp. 3482–3498, Mar. 2025, doi: 10.1109/TITS.2024.3522002.
[11] P. Mao, X. Ji, S. Li, X. Qu, and B. Ran, "An internal stochastic car-following model: Stochasticity analysis of mixed traffic environment," *Phys. Stat. Mech. Its Appl.*, vol. 653, p. 130051, Nov. 2024, doi: 10.1016/j.physa.2024.130051.
[12] Q. Zeng, S. Hao, N. Zhao, and R. Liu, "Modeling and Analysis of Mixed Traffic Flow Considering Driver Stochasticity and CAV Connectivity Uncertainty," *Sensors*, vol. 25, no. 9, p. 2806, Apr. 2025, doi: 10.3390/s25092806.
[13] L. Li, S. Li, J. Gan, X. Qu, and B. Ran, "Revealing the impact of stochastic driving characteristics on car-following behavior with locally collected vehicle trajectory data," *Transp. B Transp. Dyn.*, vol. 12, no. 1, p. 2299993, Dec. 2024, doi: 10.1080/21680566.2023.2299993.
[14] Y. Zhang, A. Kouvelas, and M. A. Makridis, "A time-varying shockwave speed model for reconstructing trajectories on freeways using Lagrangian and Eulerian observations," *Expert Syst. Appl.*, vol. 253, p. 124298, Nov. 2024, doi: 10.1016/j.eswa.2024.124298.
[15] Z. Yao, Y. Wu, Y. Wang, B. Zhao, and Y. Jiang, "Analysis of the impact of maximum platoon size of CAVs on mixed traffic flow: An analytical and simulation method," *Transp. Res. Part C Emerg. Technol.*, vol. 147, p. 103989, Feb. 2023, doi: 10.1016/j.trc.2022.103989.
[16] J. Dong, J. Wang, and D. Luo, "Impact of connected and autonomous vehicles on traffic safety of mixed traffic flow: from the perspective of connectivity and spatial distribution," *Transp. Saf. Environ.*, vol. 4, no. 3, p. tdac021, Aug. 2022, doi: 10.1093/tse/tdac021.
[17] M. Sala and F. Soriguera, "Capacity of a freeway lane with platoons of autonomous vehicles mixed with regular traffic," *Transp. Res. Part B Methodol.*, vol. 147, pp. 116–131, May 2021, doi: 10.1016/j.trb.2021.03.010.
[18] J. Tian, R. Jiang, B. Jia, Z. Gao, and S. Ma, "Empirical analysis and simulation of the concave growth pattern of traffic oscillations," *Transp. Res. Part B Methodol.*, vol. 93, pp. 338–354, Nov. 2016, doi: 10.1016/j.trb.2016.08.001.
[19] X. Chang, H. Li, J. Rong, X. Zhao, and A. Li, "Analysis on traffic stability and capacity for mixed traffic flow with platoons of intelligent connected vehicles," *Phys. Stat. Mech. Its Appl.*, vol. 557, p. 124829, Nov. 2020,

doi: 10.1016/j.physa.2020.124829.
[20] F. Zheng, C. Liu, X. Liu, S. E. Jabari, and L. Lu, "Analyzing the impact of automated vehicles on uncertainty and stability of the mixed traffic flow," *Transp. Res. Part C Emerg. Technol.*, vol. 112, pp. 203–219, Mar. 2020, doi: 10.1016/j.trc.2020.01.017.
[21] J. A. Laval, C. S. Toth, and Y. Zhou, "A parsimonious model for the formation of oscillations in car-following models," *Transp. Res. Part B Methodol.*, vol. 70, pp. 228–238, Dec. 2014, doi: 10.1016/j.trb.2014.09.004.
[22] H. Yeo and A. Skabardonis, "Understanding Stop-and-go Traffic in View of Asymmetric Traffic Theory," in *Transportation and Traffic Theory 2009: Golden Jubilee*, W. H. K. Lam, S. C. Wong, and H. K. Lo, Eds., Boston, MA: Springer US, 2009, pp. 99–115. doi: 10.1007/978-1-4419-0820-9_6.
[23] Y.-X. Huang *et al.*, "Experimental study and modeling of car-following behavior under high speed situation," *Transp. Res. Part C Emerg. Technol.*, vol. 97, pp. 194–215, Dec. 2018, doi: 10.1016/j.trc.2018.10.022.
[24] R. Jiang *et al.*, "Traffic Experiment Reveals the Nature of Car-Following," *PLoS ONE*, vol. 9, no. 4, p. e94351, Apr. 2014, doi: 10.1371/journal.pone.0094351.
[25] L. Lu, F. Zheng, and X. Liu, "Uncertainty, Efficiency, and Stability of Mixed Traffic Flow: Stochastic Model-Based Analyses," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2678, no. 8, pp. 336–357, Aug. 2024, doi: 10.1177/03611981231215338.
[26] Z. Liu and Q. Kong, "An extended stochastic car-following model and its feedback control," *Math. Methods Appl. Sci.*, vol. 47, no. 18, pp. 14402–14416, Dec. 2024, doi: 10.1002/mma.10280.
[27] A. Sharma, Z. Zheng, J. Kim, A. Bhaskar, and Md. Mazharul Haque, "Assessing traffic disturbance, efficiency, and safety of the mixed traffic flow of connected vehicles and traditional vehicles by considering human factors," *Transp. Res. Part C Emerg. Technol.*, vol. 124, p. 102934, Mar. 2021, doi: 10.1016/j.trc.2020.102934.
[28] M. Bouadi, B. Jia, R. Jiang, X. Li, and Z.-Y. Gao, "Stochastic factors and string stability of traffic flow: Analytical investigation and numerical study based on car-following models," *Transp. Res. Part B Methodol.*, vol. 165, pp. 96–122, Nov. 2022, doi: 10.1016/j.trb.2022.09.007.
[29] S. Jin, D.-H. Sun, and Z. Liu, "The Impact of Spatial Distribution of Heterogeneous Vehicles on Performance of Mixed Platoon: A Cyber-Physical Perspective," *KSCE J. Civ. Eng.*, vol. 25, no. 1, pp. 303–315, Jan. 2021, doi: 10.1007/s12205-020-2363-5.
[30] Z.-H. Li, S.-T. Zheng, R. Jiang, J.-F. Tian, K.-X. Zhu, and R. Di Pace, "Empirical and simulation study on traffic oscillation characteristic using floating car data," *Phys. Stat. Mech. Its Appl.*, vol. 605, p. 127973, Nov. 2022, doi: 10.1016/j.physa.2022.127973.
[31] J. Wang, R. Liu, and F. Montgomery, "Car-Following Model for Motorway Traffic," *Transp. Res. Rec.*, 2005.
[32] J. Tian, H. M. Zhang, M. Treiber, R. Jiang, Z.-Y. Gao, and B. Jia, "On the role of speed adaptation and spacing indifference in traffic instability: Evidence from car-following experiments and its stochastic model," *Transp. Res. Part B Methodol.*, vol. 129, pp. 334–350, Nov. 2019, doi: 10.1016/j.trb.2019.09.014.
[33] K. Yuan, J. Laval, V. L. Knoop, R. Jiang, and S. P. Hoogendoorn, "A geometric Brownian motion car-following model: towards a better understanding of capacity drop," *Transp. B Transp. Dyn.*, vol. 7, no. 1, pp. 915–927, Dec. 2019, doi: 10.1080/21680566.2018.1518169.
[34] F. Zheng, L. Lu, R. Li, X. Liu, and Y. Tang, "Traffic Oscillation using Stochastic Lagrangian Dynamics: Simulation and Mitigation via Control of Autonomous Vehicles," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2673, no. 7, pp. 1–11, Jul. 2019, doi: 10.1177/0361198119844455.
[35] R. E. Stern *et al.*, "Dissipation of stop-and-go waves via control of autonomous vehicles: Field experiments," *Transp. Res. Part C Emerg. Technol.*, vol. 89, pp. 205–221, Apr. 2018, doi: 10.1016/j.trc.2018.02.005.
[36] Z. He, L. Zheng, L. Song, and N. Zhu, "A Jam-Absorption Driving Strategy for Mitigating Traffic Oscillations," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 4, pp. 802–813, Apr. 2017, doi: 10.1109/TITS.2016.2587699.
[37] Y. Jiang, F. Zhu, Q. Gu, Y. Wu, X. Wen, and Z. Yao, "Influence of CAVs platoon characteristics on fundamental diagram of mixed traffic flow," *Phys. Stat. Mech. Its Appl.*, vol. 624, p. 128906, Aug. 2023, doi:

10.1016/j.physa.2023.128906.
[38] Z. Yao, Q. Gu, Y. Jiang, and B. Ran, "Fundamental diagram and stability of mixed traffic flow considering platoon size and intensity of connected automated vehicles," *Phys. Stat. Mech. Its Appl.*, vol. 604, p. 127857, Oct. 2022, doi: 10.1016/j.physa.2022.127857.
[39] Y. Zhu, Y. Li, H. Zhao, X. Huang, Y. Ou, and S. Peeta, "The effects of communication topology on mixed traffic flow: modelling and analysis," *Transp. B Transp. Dyn.*, vol. 12, no. 1, p. 2422373, Dec. 2024, doi: 10.1080/21680566.2024.2422373.
[40] H. Dong, J. Shi, W. Zhuang, Z. Li, and Z. Song, "Analyzing the impact of mixed vehicle platoon formations on vehicle energy and traffic efficiencies," *Appl. Energy*, vol. 377, p. 124448, Jan. 2025, doi: 10.1016/j.apenergy.2024.124448.
[41] H. Liu, Z. Mu, Z. Wang, and B. B. Park, "Mixed Traffic Flow Performance Evaluation Considering Spatial Distribution of Mixed-Automated Vehicles," in *2024 IEEE 27th International Conference on Intelligent Transportation Systems (ITSC)*, Edmonton, AB, Canada: IEEE, Sep. 2024, pp. 3351–3358. doi: 10.1109/ITSC58415.2024.10920006.
[42] E. Fauchet, P.-A. Laharotte, K. Bhattacharyya, and N.-E. E. Faouzi, "Does the Spatial Distribution of Connected Vehicles Affect Traffic Control Efficiency?," in *2023 8th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, Nice, France: IEEE, Jun. 2023, pp. 1–6. doi: 10.1109/MT-ITS56129.2023.10241646.
[43] G. Hu, F. Wang, W. Lu, T. A. Kwembe, and R. W. Whalin, "Cooperative bypassing algorithm for connected and autonomous vehicles in mixed traffic," *IET Intell. Transp. Syst.*, vol. 14, no. 8, pp. 915–923, Aug. 2020, doi: 10.1049/iet-its.2019.0707.
[44] A. Ghiasi, O. Hussain, Z. (Sean) Qian, and X. Li, "A mixed traffic capacity analysis and lane management model for connected automated vehicles: A Markov chain method," *Transp. Res. Part B Methodol.*, vol. 106, pp. 266–292, Dec. 2017, doi: 10.1016/j.trb.2017.09.022.
[45] Y. Jiang, F. Zhu, Z. Yao, Q. Gu, and B. Ran, "Platoon Intensity of Connected Automated Vehicles: Definition, Formulas, Examples, and Applications," *J. Adv. Transp.*, vol. 2023, pp. 1–17, Apr. 2023, doi: 10.1155/2023/3325530.
[46] P. Zhao, Y. D. Wong, and F. Zhu, "Modeling the clustering strength of connected autonomous vehicles and its impact on mixed traffic capacity," *Commun. Transp. Res.*, vol. 4, p. 100151, Dec. 2024, doi: 10.1016/j.commtr.2024.100151.
[47] S. He, F. Ding, C. Lu, and Y. Qi, "Impact of connected and autonomous vehicle dedicated lane on the freeway traffic efficiency," *Eur. Transp. Res. Rev.*, vol. 14, no. 1, p. 12, Dec. 2022, doi: 10.1186/s12544-022-00535-4.
[48] Z. Wang, "A cellular automaton model for mixed traffic flow considering the size of CAV platoon," *Phys. A*, 2024.
[49] Y. Jiang, S. Wang, Z. Yao, B. Zhao, and Y. Wang, "A cellular automata model for mixed traffic flow considering the driving behavior of connected automated vehicle platoons," *Phys. Stat. Mech. Its Appl.*, vol. 582, p. 126262, Nov. 2021, doi: 10.1016/j.physa.2021.126262.
[50] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama, "Dynamical model of traffic congestion and numerical simulation," *Phys. Rev. E*, vol. 51, no. 2, pp. 1035–1042, Feb. 1995, doi: 10.1103/PhysRevE.51.1035.
[51] V. Milanes, S. E. Shladover, J. Spring, C. Nowakowski, H. Kawazoe, and M. Nakamura, "Cooperative Adaptive Cruise Control in Real Traffic Situations," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 1, pp. 296–305, Feb. 2014, doi: 10.1109/TITS.2013.2278494.
[52] M. Treiber, A. Hennecke, and D. Helbing, "Congested traffic states in empirical observations and microscopic simulations," *Phys. Rev. E*, vol. 62, no. 2, pp. 1805–1824, Aug. 2000, doi: 10.1103/PhysRevE.62.1805.
[53] G. Song and L. Yu, "Estimation of Fuel Efficiency of Road Traffic by Characterization of Vehicle-Specific Power and Speed Based on Floating Car Data," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2139, no. 1, pp. 11–

20, Jan. 2009, doi: 10.3141/2139-02.
[54] Y. Sugiyama *et al.*, "Traffic jams without bottlenecks—experimental evidence for the physical mechanism of the formation of a jam," *New J. Phys.*, vol. 10, no. 3, p. 033001, Mar. 2008, doi: 10.1088/1367-2630/10/3/033001.