Modeling multilane traffic flow in Lagrangian coordinates
Formulation and implementation

Lu, Liang; Zheng, Fangfang; Liu, Xiaobo; Liu, Henry X.; Yuan, Yufei

DOI
10.1016/j.trb.2025.103384

Publication date
2025

Document Version
Final published version

Published in
Transportation Research Part B: Methodological

Citation (APA)
Lu, L., Zheng, F., Liu, X., Liu, H. X., & Yuan, Y. (2025). Modeling multilane traffic flow in Lagrangian coordinates: Formulation and implementation. *Transportation Research Part B: Methodological*, *204*, Article 103384. https://doi.org/10.1016/j.trb.2025.103384

Transportation Research Part B 204 (2026) 103384

Contents lists available at ScienceDirect

Transportation Research Part B

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

Modeling multilane traffic flow in Lagrangian coordinates: Formulation and implementation

Liang Lu [a], Fangfang Zheng [a,*], Xiaobo Liu [a], Henry X. Liu [b], Yufei Yuan [c]

[a] School of Transportation and Logistics, Southwest Jiaotong University, Western Hi-tech Zone Chengdu, Sichuan 611756, China
[b] Department of Civil and Environmental Engineering, University of Michigan, 2350 Hayward Street, Ann Arbor, MI 48109, USA
[c] Department of Transport and planning, Delft university of technology, Delft, the Netherlands

ARTICLE INFO

Keywords:
Multilane traffic flow
Lagrangian coordinates
Lane-changing number estimation
Spacing dynamics

ABSTRACT

This paper proposes a multilane traffic flow model based on the notions of conservation laws in Lagrangian coordinates. Both continuous formulation and discretization of the model are derived explicitly considering lane-changing characteristics. For model discretization, a lane-changing number estimation model is developed to calculate the net lane-changing number for each vehicle group considering the relative position between vehicle groups in the current and adjacent lanes. With model discretization, the spacing of vehicle groups for each lane can be dynamically calculated. In addition, the boundary conditions for both the continuous Lagrangian model and its discretization are also derived. A numerical implementation of the model in the case of a three-lane highway section with a lane-drop is discussed, and results indicate that the proposed Lagrangian model can well simulate traffic dynamics, including the generation and propagation of congestion, and the perturbation caused by lane-changing behaviors. The lane-changing characteristics in terms of a cumulative net number of lane-changing vehicles for each lane, and the spacing dynamics of vehicle groups can be estimated as well. We further validate the proposed model using real-world data observed from a two-lane freeway section in Japan. The results show that the proposed multilane Lagrangian model can well capture traffic dynamic properties and could provide a relatively accurate estimation in terms of lane volume dynamics, vehicle spacing dynamics, and the cumulative net number of lane-changing vehicles. Comparisons with the Eulerian multilane model indicate that the Lagrangian model offers superior performance in predicting vehicle counts and spacing, especially under congested conditions. This improved performance can be attributed to the Lagrangian model's capability to track individual vehicle groups, resulting in a more precise representation of traffic dynamics.

1. Introduction

The Lighthill-Whitham-Richards (LWR) model (Lighthill and Whitham, 1955a, 1955b; Richards, 1956), commonly referred to as the kinematic wave model, is a vital tool for understanding traffic flow dynamics. Over the past few decades, the LWR model and its extensions (Newell, 1993; Daganzo, 1994, 1995, 2002a, 2002b; Wong and Wong, 2002; Chanut and Buisson, 2003; Laval and Daganzo, 2006; Levin and Boyles, 2016) have played a significant role in traffic estimation and prediction. The model originally operates on the Eulerian coordinate system (t,x), which describes traffic state variables in terms of time (t) and space (x), making it especially suitable

* Corresponding author.
E-mail address: fzheng@swjtu.cn (F. Zheng).

https://doi.org/10.1016/j.trb.2025.103384
Received 19 September 2023; Received in revised form 4 September 2024; Accepted 19 December 2025
Available online 26 December 2025

for analyzing traffic data obtained from fixed-location measurements. Common types of traffic data include location-based information (Eulerian) and vehicle-based information (Lagrangian). Historically, traffic measurements have predominately depended on Eulerian information, acquired through inductive loop detectors or cameras positioned at fixed locations (Coifman, 2002, 2003; Mehran et al., 2012). However, with technological advancements, Lagrangian information's accessibility has increased considerably. Such data include vehicle trajectory data from travel service companies, vehicles equipped with GPS navigation systems, and the growing number of connected/autonomous vehicles in the future. Notably, the emergence of connected and autonomous technologies has highlighted the dual potential of connected/automated vehicles. These vehicles possess the ability to improve current traffic situations while simultaneously serving as mobile sensors for traffic data collection (Molnár et al., 2021; Ma and Qian, 2019; Orosz et al., 2017). This rising prevalence of Lagrangian data poses a challenge to the traditional Eulerian models.

To address this challenge, two primary methods are typically pursued. The first involves employing complex data transformations to adapt Lagrangian data to a fixed-location measurement format. These transformation procedures typically involve temporal and spatial aggregation, as well as conversions between different types of measurements utilizing flow-density-speed relationships. The second method integrates measurement theory into the traffic flow model. This can be achieved by incorporating techniques such as Kalman filtering (Work et al., 2010, 2008), Newton relaxation (Herrera et al., 2010), and the integration of car-following models and traffic flow models (Delle Monache et al., 2019). While the first method involves extracting the initial and boundary conditions required by the Eulerian model from Lagrangian data, the complex transformation procedures can be cumbersome and error-prone. On the other hand, the second method aims to unite measurements with traffic flow models in a theoretical sense. However, depending on the complexity of the problem, modeling and managing integrated models may also be cumbersome. While Eulerian models can certainly make use of Lagrangian data through these operations, employing a Lagrangian model itself provides a more natural and straightforward approach for leveraging Lagrangian data.

The Lagrangian coordinate system, denoted as (t, n), offers a straightforward approach to handling the vehicle trajectory data, where t represents time and n represents the vehicle index, allowing the system to move seamlessly along vehicle paths. This approach was first introduced into the traffic flow by Leclercq et al. (2007). They proposed an alternative formulation of the LWR model based on Lagrangian coordinates by combining the Lagrangian conservation law and the speed-spacing relationship. More recently, Laval and Leclercq (2013) applied the Hamilton-Jacobi formula to establish a first-order traffic flow model framework in three different coordinate systems, including the traditional Eulerian coordinate system and two Lagrangian coordinate systems. Their models leverage variational techniques and the Godunov scheme with a triangular fundamental diagram. Notably, the Godunov scheme reduces to the upwind format in Lagrangian coordinates and does not require a transformation between the upwind and downwind formats, which leads to a simplification of the solution process. Simulations have demonstrated that this Lagrangian model is more accurate-with less numerical diffusion-and computationally efficient (Van Wageningen-Kessels et al., 2010). These models have attracted considerable attention due to their reduced nonlinearity and efficient data assimilation capabilities with Lagrangian-type data. Consequently, they have been widely applied in traffic state estimation and control, particularly in conjunction with data assimilation techniques utilizing vehicle trajectory data (Yuan et al., 2012; Duret and Yuan, 2017; Zheng et al., 2018). Nevertheless, the existing Lagrangian models assume a single pipe or lane of traffic and do not explicitly account for lane-changing. While van Wageningen-Kessels et al. (2010, 2013) explored multiclass traffic flow and discontinuities like merges and diverges in Lagrangian coordinates, lane-changing remains underexplored. Laval et al. (2016) investigate the effect of source terms on traffic flow using variational theory, where these terms represent the net lateral inflow rate, typically involving lane changes, merges, and diverges. Their findings suggest that analytic solutions in variational theory may not be applicable to source terms in Lagrangian coordinates, whether they are exogenous or endogenous. As a result, Lagrangian models that incorporate lane-changing characteristics must rely on numerical solutions. However, there is no definitive method for numerically solving these models, as most research has focused on models without source terms (Leclercq et al., 2007; Van Wageningen-Kessels et al., 2010; Yuan et al., 2012; Han et al., 2021) or on those addressing specific scenarios like ramp merging and diverging (van Wageningen-Kessels et al., 2013).

To address the limitations of existing literature, we propose a novel multilane traffic flow model formulated using the Lagrangain coordinate system. The main contributions of this study are as follows:

1. **Development of Lagrangian Multilane Model:** Our model is the first to explicitly incorporate lane-changing behaviors within a Lagrangian framework. This innovation allows for a more detailed analysis of traffic dynamics in individual lanes. Additionally, we provide meticulous boundary conditions that enable the simulation to consider a wider range of real-world scenarios.
2. **Model Discretization implementation:** In the numerical example, we have developed a model discretization process that facilitates the numerical implementation of our proposed model. We introduce a lane-changing number estimation model that calculates the net lane-changing number for groups of vehicles in each lane, thereby enhancing the model's ability to predict lane-changing characteristics.
3. **Extensive Validation:** We have rigorously validated the model's effectiveness in replicating various characteristics of traffic flow, both spatially and temporally. This includes the model's ability to capture phenomena such as congestion propagation due to lane-drops and lane-changing perturbations. The validation process utilizes real-world data, which confirms the model's ability to accurately reflect lane volume, vehicle spacing, and lane-changing dynamics, while also demonstrating superior performance compared to the Eulerian multilane model.

The rest of this paper is organized as follows. Section 2 provides a detailed formulation of the multilane traffic flow model in the Lagrangian coordinates. Then, a discretization of the proposed model is discussed in Section 3. Section 4 presents a model to estimate the critical variable for discretization, i.e., the net lane-changing number. Section 5 provides a numerical example to demonstrate how the proposed model works. A validation of the proposed model using real-world data (the ZTD data) and a comparison with the Eulerian multilane model are provided in Section 6. Section 7 concludes the paper with some discussions.

2. Lagrangian formulation of multilane LWR model

The multilane Lighthill-Whitham-Richards (LWR) model in traditional Eulerian coordinates was developed by Laval and Daganzo (2006). Let's briefly revisit the Eulerian conservation law for multilane traffic.

$$\frac{\partial \rho_l}{\partial t} + \frac{\partial q_l}{\partial x_l} = g_l(t, x_l), l = 1, 2, \cdots, N_{TL} \tag{1}$$

where an equilibrium relation between flow q_l and density ρ_l is given by a fundamental diagram (FD): $q_l = Q_l(\rho_l) = k_l V_l(\rho_l)$. N_{TL} represents the total number of lanes and $g_l(t, x_l)$ represents the net lane-changing rate in lane l over time and space, measured in vehicles per time–distance. It can be expressed as:

$$g_l(t, x_l) = F_l(t, x_l) H_l(x_l) \tag{2}$$

where $F_l(t, x_l)$ is the net lane-changing flow in lane l, measured in vehicles per time. $H_l(x_l)$ is a function that satisfies $\int_{-\infty}^{\infty} H_l(x_l) dx_l = 1$, measured in 1/distance, and describes the distribution of lane-changing locations in lane l. This concept was simplified to specific points for merging and diverging flows at ramps in van Wageningen-Kessels et al. (2013), but here, we extend it to a road segment for better representation of lane-changing behavior. For example, if lane-changing occurs uniformly between points a and b in space. $H_l(x_l)$ can be represented by the following block function:

$$H_l(x_l) = \begin{cases} \frac{1}{b-a}, a < x_l < b \\ 0, \text{otherwise} \end{cases}, \int_{-\infty}^{\infty} H_l(x_l) \mathrm{d}x_l = 1 \tag{3}$$

Moving forward, we will derive the Lagrangian formulation of this multilane LWR model and explore the boundary conditions related to inflow and outflow restrictions at both ends of the road section.

2.1. Mathematical derivation

The multilane LWR model in the Eulerian coordinate (t, x_l) can be reformulated in a new coordinate system called the Lagrangian coordinate (t, n_l). To enable the transformation from x_l to n_l, we assume that vehicle index n_l increases in the same direction as x_l decreases. In other words, a larger value of x_l corresponds to a smaller value of n_l. Therefore, the transformation can be expressed as:

$$n_l(t, x_l) = n_l(t^0, x_l^0) + \int_{t^0}^{t} q_l(t', x_l) dt' - \int_{x_l^0}^{x_l} \rho_l(t^0, x') dx' \tag{4}$$

where $n(t^0, x_l^0)$ is the vehicle index located at position x_l^0 at time t^0 in lane l, $\int_{t^0}^{t} q_l(t', x_l) dt'$ is the number of vehicles passing through position x_l from t^0 to t, and $\int_{x_l^0}^{x_l} \rho_l(t^0, x') dx'$ is the number of vehicles between positions x_l^0 and x_l at time t^0. Since we have $\rho_l(t, x_l) = 1 / s_l(t, n_l(t, x_l))$ and spacing s_l is the distance between the tails of two adjacent vehicles in lane l, we can derive:

$$q_l(t, x_l) = \rho_l(t, x_l) V_l(\rho_l(t, x_l)) = \frac{V_l^*(s_l(t, n_l(t, x_l)))}{s_l(t, n_l(t, x_l))} \tag{5}$$

where $V_l(\rho_l)$ and $V_l^*(s_l)$ are the equilibrium relationships (fundamental diagrams) between speed and density, as well as speed and spacing, respectively. There is a relationship between them, which is $v_l = V_l(\rho_l) = V_l\left(\frac{1}{s_l}\right) = V_l^*(s_l)$.

Thus, Eq. (1) can be reformulated as:

$$\begin{aligned} g_l(t, x_l(t, n_l)) &= \frac{\partial}{\partial t}\left(\frac{1}{s_l(t, n_l(t, x_l))}\right) + \frac{\partial}{\partial x_l}\left(\frac{V_l^*(s_l(t, n_l(t, x_l)))}{s_l(t, n_l(t, x_l))}\right) \\ &= -\frac{1}{s_l^2}\left(\frac{\partial s_l}{\partial t} + \frac{\partial s_l}{\partial n_l}\frac{\partial n_l}{\partial t}\right) + \frac{1}{s_l^2}\left(s_l \frac{dV_l^*}{ds_l}\frac{\partial s_l}{\partial n_l}\frac{\partial n_l}{\partial x_l} - V_l^* \frac{\partial s_l}{\partial n_l}\frac{\partial n_l}{\partial x_l}\right) \\ &= -\frac{1}{s_l^2}\left(\frac{\partial s_l}{\partial t} + \frac{\partial s_l}{\partial n_l}\frac{\partial n_l}{\partial t} - s_l \frac{\partial V_l^*}{\partial n_l}\frac{\partial n_l}{\partial x_l} + V_l^* \frac{\partial s_l}{\partial n_l}\frac{\partial n_l}{\partial x_l}\right) \end{aligned} \tag{6}$$

By applying the relations given in Eq. (7):

$$\frac{\partial n_l}{\partial t} = q_l = \frac{V_l^*(s_l)}{s_l}, \ \frac{\partial n_l}{\partial x_l} = -\rho_l = -\frac{1}{s_l} \tag{7}$$

we can derive the following equation:

$$\begin{aligned} g_l(t, x_l(t, n_l)) &= -\frac{1}{{s_l}^2}\left(\frac{\partial s_l}{\partial t} + \frac{\partial s_l}{\partial n_l}q_l - s_l\frac{\partial V_l^*}{\partial n_l}\left(-\frac{1}{s_l}\right) + V_l^*\frac{\partial s_l}{\partial n_l}\left(-\frac{1}{s_l}\right)\right) \\ &= -\frac{1}{{s_l}^2}\left(\frac{\partial s_l}{\partial t} + \frac{\partial V_l^*}{\partial n_l} + \frac{\partial s_l}{\partial n_l}q_l - \frac{\partial s_l}{\partial n_l}q_l\right) \\ &= -\frac{1}{{s_l}^2}\left(\frac{\partial s_l}{\partial t} + \frac{\partial V_l^*}{\partial n_l}\right) \end{aligned} \tag{8}$$

By rearranging and multiplying $-s_l^2$ on both sides of Eq. (8), we obtain the conservation law of multilane traffic in Lagrangian coordinates for a specific lane l as follows:

$$\frac{\partial s_l(t, n_l)}{\partial t} + \frac{\partial V_l^*(s_l(t, n_l))}{\partial n_l} = -s_l^2(t, n_l)g_l(t, x_l(t, n_l)) \tag{9}$$

Eq. (9) is mathematically consistent with the findings of van Wageningen-Kessels et al. (2013) and the simplified results presented by Laval et al. (2016). However, a key difference lies in the function $g_l(t, x_l(t, n_l))$ utilized in each study. Our model explicitly incorporates discretionary lane changes within $g_l(t, x_l(t, n_l))$, whereas van Wageningen-Kessels et al. (2013) primarily focus on merging and diverging flows. Laval et al. (2016) consider merging, diverging, and lane-changing behaviors as source terms, but do not provide specific functional expressions for them. Their study explores the impact of source terms on the representation of traffic flow using variational theory, concluding that analytic solutions within this theory may not be applicable to source terms in Lagrangian coordinates, whether exogenous or endogenous. However, their work does not delve into the specific discretization and implementation of these source terms in Lagrangian coordinates. Our contribution lies in explicitly integrating lane-changing behaviors within a Lagrangian framework and developing a numerical solution for the model.

The term $g_l(t, x_l(t, n_l))$ in Eq. (9) represents the net lane-changing rate of lane l, which is given by:

$$g_l(t, x_l(t, n_l)) = F_l(t, x_l(t, n_l))H_l(x_l(t, n_l)) \tag{10}$$

with $\int_{-\infty}^{\infty} H_l(x_l(t, n_l))\mathrm{d}(x_l(t, n_l)) = 1$. Moreover, $g_l(t, x_l(t, n_l))$ needs to be expressed in terms of t and n_l. To achieve this, we solve the relations $\frac{\partial x_l}{\partial n_l} = -s_l(t, n_l)$ and $\frac{\partial x_l}{\partial t} = v_l(t, n_l)$ to obtain $x_l(t, n_l)$:

$$x_l(t, n_l) = x(t^0, n_l^0) + \int_{t^0}^{t} v_l(t', n_l^0)dt' - \int_{n_l^0}^{n} s_l(t, n')dn' \tag{11}$$

where $x(t^0, n_l^0) + \int_{t^0}^{t} v_l(t', n_l^0)dt'$ represents the position corresponding to the vehicle index n_l^0 at time t, while $\int_{n_l^0}^{n} s_l(t, n')dn'$ represents the distance between n_l^0 and n_l at time t.

2.2. Boundary conditions

The boundary conditions should satisfy the inflow (or outflow) restrictions at the inflow boundary $x_l = 0$ (or the outflow boundary $x_l = L$). Specifically, at the inflow boundary $x_l = 0$ for lane l, we define the demand function $\mathrm{A}_l(0^-, t)$ as the number of vehicles per unit time that intend to enter the road section from the position upstream of it.

$$\mathrm{A}_l(0^-, t) = \begin{cases} q_l(0^-, t), \ \rho_l(0^-, t) \le \rho_{cri,l} \\ Q_l, \ \rho_l(0^-, t) > \rho_{cri,l} \end{cases} \tag{12}$$

where $0^- = lim_{\xi \uparrow 0}\xi$ represents the position immediately upstream of the inflow boundary, and $q_l(0^-, t)$ and $\rho_l(0^-, t)$ represent the flow and density, respectively, at the position $x_l = 0^-$ and time t. $\rho_{cri,l}$ denotes the critical density, while Q_l represents the capacity of lane l.

Similarly, the supply function $\mathrm{B}_l(0^+, t)$ is defined as the number of vehicles per unit time that can still fit at the position downstream of the inflow boundary. The supply function is given by:

$$\mathrm{B}_l(0^+, t) = \begin{cases} Q_l, \ \rho_l(0^+, t) \le \rho_{cri,l} \\ q_l(0^+, t), \ \rho_l(0^+, t) > \rho_{cri,l} \end{cases} \tag{13}$$

where $0^+ = lim_{\xi \downarrow 0}\xi$ denotes the position immediately downstream of the inflow boundary. $q_l(0^+, t)$ and $k_l(0^+, t)$ represent the flow and density, respectively, at the position $x_l = 0^+$ and time t.

Hence, the inflow boundary condition can be described as follows:

$$q_l(0,t) = \min(\mathrm{A}_l(0^-,t), \mathrm{B}_l(0^+,t)) \tag{14}$$

The inflow boundary condition indicates that the inflow is constrained by both the traffic conditions upstream and downstream of the inflow boundary.

Similarly, for the outflow boundary $x_l = L$ in lane l, the outflow boundary condition can be derived as:

$$q_l(L,t) = \min(\mathrm{A}_l(L^-,t), \mathrm{B}_l(L^+,t)) \tag{15}$$

where $L^- = lim_{\xi\uparrow 0}(L-\xi)$ represents the position immediately upstream of the outflow boundary, and $L^+ = lim_{\xi\downarrow 0}(L+\xi)$ represents the position immediately downstream of the outflow boundary.

3. Discretization of the multilane Lagrangian model

To facilitate the computer implementation of the proposed model, it is necessary to discretize the continuous Lagrangian model. This involves dividing vehicles in any lane l into multiple groups of fixed size Δn, where Δn does not have to be an integer. Each vehicle group is assigned a number, i, based on its position in descending order. Additionally, time is divided into intervals of Δt. Each time interval is assigned a number, k. In this section, we will first discuss the discretization of the continuous multilane Lagrangian model. Next, we will present a solution to address the issue of changing vehicle group sizes. Finally, we will discuss the discretization of the boundary conditions.

3.1. Discretization of the Lagrangian conservation law

The continuous multilane Lagrangian model described in Eq. (9) can be discretized as:

$$\frac{s_{i,l}^{k+1} - s_{i,l}^{k}}{\Delta t} + \frac{v_{i,l}^{k} - v_{i-1,l}^{k}}{\Delta n} = -\left(s_{i,l}^{k}\right)^2 g_{i,l}^{k} \tag{16}$$

By rearranging Eq. (16), we obtain the discretized spacing dynamics as:

$$s_{i,l}^{k+1} = s_{i,l}^{k} + \frac{\Delta t}{\Delta n}\left(v_{i-1,l}^{k} - v_{i,l}^{k}\right) - \left(s_{i,l}^{k}\right)^2 g_{i,l}^{k}\Delta t \tag{17}$$

where $s_{i,l}^{k+1} = s_l\left(t_{k+1}, n_{i,l}\right) = s_l\left(t^0 + (k+1)\Delta t, n_l^0 + i\Delta n\right)$ represents the spacing associated with the i^{th} vehicle group in lane l at time step $k+1$, while $v_{i,l}^{k}$ represents the speed associated with i^{th} vehicle group in lane l at time step k. $g_{i,l}^{k}$ is a discrete variable representing the net lane-changing rate of the i^{th} vehicle group in lane l from time step k to $k+1$. According to Eq. (10), $g_{i,l}^{k}$ can be expressed as:

$$g_{i,l}^{k} = F_{i,l}^{k} H_{i,l}^{k} \tag{18}$$

In Eq. (18), the first term on the right-hand side, $F_{i,l}^{k}$, represents the net lane-changing flow of the i^{th} vehicle group in lane l from time step k to $k+1$. This term is given by:

$$F_{i,l}^{k} = \frac{\varphi_{i,l}^{k}}{\Delta t} \tag{19}$$

where $\varphi_{i,l}^{k}$ represents the net lane-changing number of the i^{th} vehicle group in lane l from time step k to $k+1$. The specific details of $\varphi_{i,l}^{k}$ and its relationship to the lane-changing behavior will be discussed in Section 4.

The second term on the right-hand side of Eq. (18) involves $H_{i,l}^{k}$, which represents the distribution of the lane-changing position for the i^{th} vehicle group in lane l from time step k to $k+1$. We assume that the lane-changing position of the i^{th} vehicle group is uniformly distributed between points $x_{i,l}^{k}$ and $x_{i-1,l}^{k}$. Referring to Eq. (3), $H_{i,l}^{k}$ can be expressed as:

$$H_{i,l}^{k} = \frac{1}{x_{i-1,l}^{k} - x_{i,l}^{k}} = \frac{1}{s_{i,l}^{k}\Delta n} \tag{20}$$

where $x_{i,l}^{k}$ and $x_{i-1,l}^{k}$ denote the position of the tail of the vehicle group i and $i-1$ in lane l at time step k. Therefore, the difference $x_{i-1,l}^{k} - x_{i,l}^{k}$ can be written as $x_{i-1,l}^{k} - x_{i,l}^{k} = s_{i,l}^{k}\Delta n$.

By substituting Eqs. (19) and (20) into (17), we obtain

$$s_{i,l}^{k+1} = s_{i,l}^{k} + \frac{\Delta t}{\Delta n}\left(v_{i-1,l}^{k} - v_{i,l}^{k}\right) - \frac{s_{i,l}^{k}}{\Delta n}\varphi_{i,l}^{k} \tag{21}$$

It is important to note that the stability domain of Eq. (21) is defined by the Courant-Friedrichs-Lewy (CFL) condition (Courant et al., 1967), as given by:

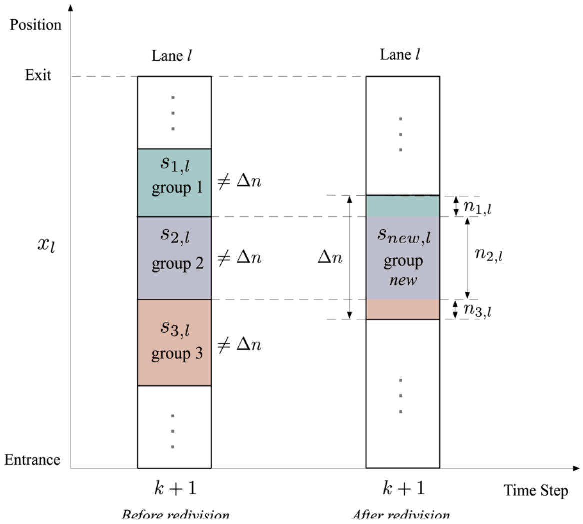


Fig. 1. Illustration of vehicle group re-division.

$$\max\left|\frac{dv_l}{ds_l}\right|\Delta t \leq \Delta n \tag{22}$$

The CFL condition ensures the stability of the model and states that the distance a vehicle group can travel within one time step should not exceed the length of one vehicle group. Therefore, appropriate values for Δt and Δn are chosen to ensure compliance with the CFL condition.

A problem arises that the number of vehicles in the vehicle group i may change due to lane-changing behavior, resulting in a group size that does not necessarily equal the fixed group size of Δn. This issue is related to the third term on the right-hand side of Eq. (21). If $\varphi_{i,l}^k$ is not equal to 0, it indicates that the i^{th} vehicle group experiences a net inflow (when $\varphi_{i,l}^k > 0$) or net outflow (when $\varphi_{i,l}^k < 0$) of vehicles from time step k to $k+1$. To address this problem, we have employed an appropriate method to re-divide the vehicle groups, which will be described in the following Section 3.2.

3.2. Re-division of vehicle groups

The number of vehicles within a vehicle group can change due to lane changes, which may cause Eq. (21) to violate the CFL condition, potentially leading to numerical instabilities. This issue is similar to what van Wageningen-Kessels et al. (2013) identified, where merging and diverging in the main lane could disrupt the CFL condition. There are two potential solutions to address this problem. The first solution involves renumbering the vehicles and forming new groups. The second approach, used by van Wageningen-Kessels et al. (2013), allows inflows at the source location to be added to the main lane flow as a new vehicle group only once the cumulative inflow reaches Δn. This method works well for fixed-location merging but is not suitable for our scenario, where lane changes can occur throughout the entire roadway. Therefore, we opt for the first solution, which involves redistributing the vehicles to form new groups. Fig. 1 illustrates an example of this re-division of vehicle groups in lane l at time step $k+1$.

In the Fig. 1, the newly formed vehicle group is composed of vehicles from the previous vehicle groups 1, 2 and 3. The spacing of vehicle groups 1, 2 and 3 corresponds to $s_{1,l}$, $s_{2,l}$ and $s_{3,l}$, respectively. The number of vehicles in vehicle groups 1, 2 and 3 is not necessarily equal to Δn. The variables $n_{1,l}$, $n_{2,l}$ and $n_{3,l}$ represent the number of vehicles from vehicle groups 1, 2 and 3, respectively, that contribute to the new group. In the example shown in Fig. 1, the number of vehicles in the new group is Δn, and its spacing can be calculated as:

$$s_{new,l} = \frac{s_{1,l}n_{1,l} + s_{2,l}n_{2,l} + s_{3,l}n_{3,l}}{\Delta n} \tag{23}$$

This vehicle group re-division approach can be readily extended to form any new group i. In the following section, we will discuss the discretization of the boundary conditions.

3.3. Discretization of the boundary conditions

At the inflow boundary, located at $x_l = 0$, the flow entering the boundary at any consecutive moment t is determined by the constraint described in Eq. (14). As time is divided into intervals of Δt, the inflow boundary condition is discretized as:

$$q_l^k(0) = \min\left(\mathrm{A}_l^k(0^-), \mathrm{B}_l^k(0^+)\right) \tag{24}$$

The number of vehicles crossing the inflow boundary in a discrete-time interval Δt from time step k to $k+1$ is given by:

$$q_l^k(0)\Delta t = \min\left(\mathrm{A}_l^k(0^-)\Delta t, \mathrm{B}_l^k(0^+)\Delta t\right) \tag{25}$$

where $A_l^k(0^-) = A_l(0^-, t_k) = A_l\left(0^-, t^0 + k\Delta t\right)$ represents the flow being sent upstream of the inflow boundary at time step k, as given by:

$$\mathrm{A}_l^k(0^-) = \begin{cases} q_l^k(0^-), \ \rho_l^k(0^-) \le \rho_{cri,l} \\ Q_l, \ \rho_l^k(0^-) > \rho_{cri,l} \end{cases} \tag{26}$$

where $q_l^k(0^-)$ and $\rho_l^k(0^-)$ represent the flow and density upstream of the inflow boundary, respectively, at time step k. Therefore, in a discrete-time interval Δt, the number of vehicles being sent upstream of the inflow boundary is $A_l^k(0^-)\Delta t$.

The term $B_l^k(0^+) = B_l(0^+, t_k) = B_l\left(0^+, t^0 + k\Delta t\right)$ represents the flow being received downstream of the inflow boundary at time step k, as given by:

$$\mathrm{B}_l^k(0^+) = \begin{cases} Q_l, \ \rho_l^k(0^+) \le \rho_{cri,l} \\ q_l^k(0^+), \ \rho_l^k(0^+) > \rho_{cri,l} \end{cases} \tag{27}$$

where $q_l^k(0^+)$ and $\rho_l^k(0^+)$ represent the flow and density downstream of the inflow boundary, respectively, at time step k. As downstream of the inflow boundary is associated with the last vehicle group, $B_l^k(0^+)$ can be written as:

$$\mathrm{B}_l^k(0^+) = \mathrm{B}_l^k\left(x_l\left(I_l^k\right)\right) = \mathrm{B}_{I_l^k,l}^k = \begin{cases} \dfrac{V_l^*\left(s_{cri,l}\right)}{s_{cri,l}}, \ s_{I_l^k,l}^k \ge s_{cri,l} \\ \dfrac{V_l^*\left(s_{I_l^k,l}^k\right)}{s_{I_l^k,l}^k}, \ s_{I_l^k,l}^k < s_{cri,l} \end{cases} \tag{28}$$

where I_l^k represents the maximum number of vehicle groups in lane l at time step k, and $x_l\left(I_l^k\right)$ represents the position of the last vehicle group in lane l at time step k. $s_{cri,l} = 1/\rho_{cri,l}$ represents the critical spacing. In a discrete-time interval Δt, the number of vehicles being received downstream of the inflow boundary is $B_l^k(0^+)\Delta t$.

Eq. (25) indicates that when the downstream inflow boundary is congested, or the number of vehicles being sent is higher than the number of vehicles being received, the vehicle groups will wait to cross the inflow boundary.

At the outflow boundary, located at $x_l = L$, the outflow boundary constraint described in Eq. (15) applies. The discretized version of Eq. (15) is given by:

$$q_l^k(L) = \min\left(\mathrm{A}_l^k(L^-), \mathrm{B}_l^k(L^+)\right) \tag{29}$$

The number of vehicles crossing the outflow boundary in a discrete-time step Δt is given by:

$$q_l^k(L)\Delta t = \min\left(\mathrm{A}_l^k(L^-)\Delta t, \mathrm{B}_l^k(L^+)\Delta t\right) \tag{30}$$

where $A_l^k(L^-) = A_l(L^-, t_k) = A_l\left(L^-, t^0 + k\Delta t\right)$ represents the flow being sent upstream of the outflow boundary at time step k, as given by:

$$\mathrm{A}_l^k(L^-) = \begin{cases} q_l^k(L^-), \ \rho_l^k(L^-) \le \rho_{cri,l} \\ Q_l, \ \rho_l^k(L^-) > \rho_{cri,l} \end{cases} \tag{31}$$

where $q_l^k(L^-)$ and $\rho_l^k(L^-)$ represent the flow and density upstream of the outflow boundary, respectively, at time step k. Since upstream of the outflow boundary is associated with the first vehicle group, $A_l^k(L^-)$ can be further expressed as:

$$\mathrm{A}_l^k(L^-) = \mathrm{A}_l^k(x_l(1)) = \mathrm{A}_{1,l}^k = \begin{cases} \dfrac{V_l^*\left(s_{1,l}^k\right)}{s_{1,l}^k}, \ s_{1,l}^k \ge s_{cri,l} \\ \dfrac{V_l^*\left(s_{cri,l}\right)}{s_{cri,l}}, \ s_{1,l}^k < s_{cri,l} \end{cases} \tag{32}$$

where $x_l(1)$ is the position of the first vehicle group in lane l at time step k. Therefore, in a discrete-time interval Δt, the number of vehicles being sent upstream of the outflow boundary is $A_l^k(L^-)\Delta t$.

The term $B_l^k(L^+) = B_l(L^+, t_k) = B_l\left(L^+, t^0 + k\Delta t\right)$ determines the flow being received downstream of the outflow boundary at time step k, as given by:

$$B_l^k(L^+) = \begin{cases} Q_l, & \rho_l^k(L^+) \le \rho_{cri,l} \\ q_l^k(L^+), & \rho_l^k(L^+) > \rho_{cri,l} \end{cases} \tag{33}$$

where $q_l^k(L^+)$ and $\rho_l^k(L^+)$ represent the flow and density downstream of the outflow boundary, respectively, at time step k. In a discrete-time interval Δt, the number of vehicles being received downstream of the outflow boundary is $B_l^k(L^+)\Delta t$.

Thus, Eq. (30) indicates that when the number of vehicles being sent $A_l^k(L^-)\Delta t$ is lower than the number of vehicles being received $B_l^k(L^+,t)\Delta t$, all vehicle groups will be allowed to exit the link. Conversely, if not all vehicle groups are permitted to exit, congestion will propagate upstream.

So far, we have completed the discretization of the multilane Lagrangian model and the boundary conditions. However, in Eq. (21), the net lane-changing number of vehicles $\varphi_l^i(t)$ needs to be further estimated in order to derive the spacing dynamics. In the next section, we will introduce our proposed model for estimating the net lane-changing number.

4. Net lane-changing number estimation model

4.1. Assumptions

In order to estimate the net lane-changing number $\varphi_{i,l}^k$ in Eq. (21), we make the following assumptions:

(1) The motivation for drivers to change lanes is based on their desire to improve their driving experience, specifically by increasing their driving speed.
(2) Vehicles are not allowed to cross more than one lane within a single time step.
(3) The traffic is homogeneous, consisting of a single type of vehicle (e.g., cars).

4.2. Lane choice probability

Based on assumption 1, the lane choice probability is estimated by considering the following three aspects:

(1) The lane choice probability is proportional to the positive speed difference between the target lane and the current lane, which is consistent with Laval and Daganzo (2006). This implies that drivers are more likely to switch to a lane with a higher speed than their current one.
(2) The lane choice probability is closely related to the speed of the target lane and the current lane. For example, in congested traffic, a driver might switch from a lane moving at 30 km/h to an adjacent lane with a speed of 40 km/h. In uncongested traffic, they might switch from 60 km/h to 70 km/h. Although the speed improvement in is 10 km/h in both scenarios, the relative benefit is greater in congested conditions. Consequently, drivers are more likely to change lanes in congested traffic compared to uncongested traffic.
(3) The lane choice probability is also influenced by the level of congestion in the target lane. A congested lane is less attractive, thus reducing the probability of drivers choosing it.

Based on the considerations mentioned above, we propose a lane choice probability function as given by:

$$p_{i,l\to j,l'}^k = \frac{\max\left(0, v_{j,l'}^k - v_{i,l}^k\right)}{v_{j,l'}^k + v_{i,l}^k} \frac{1}{\tau\left(s_{j,l'}^k\right)} \tag{34}$$

where the term $p_{i,l\to j,l'}^k$ represents the proportion of drivers who desire to change lanes from vehicle group i in lane l to vehicle group j in lane l' per unit time. Consequently, during the time interval Δt, the proportion of drivers wishing to change lanes from vehicle group i in lane l to vehicle group j in lane l' is $p_{i,l\to j,l'}^k\Delta t$. The term $\max\left(0, v_{j,l'}^k - v_{i,l}^k\right)$ represents the positive speed difference between the target vehicle group j in lane l' and the current vehicle group i in lane l at time step k, which corresponds to aspect (1). The term $v_{j,l'}^k + v_{i,l}^k$ represents the sum of the speed of the target vehicle group j and the current vehicle group i at time step k. The fraction $\max\left(0, v_{j,l'}^k - v_{i,l}^k\right)/\left(v_{j,l'}^k + v_{i,l}^k\right)$ indicates that drivers are more inclined to change lanes to achieve higher speed in congested traffic compared to uncongested traffic, corresponding to aspect (2).

The term $\tau\left(s_{j,l'}^k\right)$ extends previous work by Laval and Daganzo (2006), which initially defined the time required for a driver to decide and execute a lane change when the origin lane is stopped, and the target lane is freely flowing. In our study, we expand its definition to a state-dependent variable that represents the time required for a driver to decide and execute a lane change across different traffic conditions in the target lane l'. This variable primarily correlates with the traffic density in the target lane (Toledo and Zohar, 2007; Wang et al., 2014). In denser traffic conditions, drivers take longer to find a gap and complete a lane change, requiring more time (Toledo and Zohar, 2007; Yang et al., 2019), corresponding to aspect (3). Consequently, $\tau\left(s_{j,l'}^k\right)$ decreases as the average

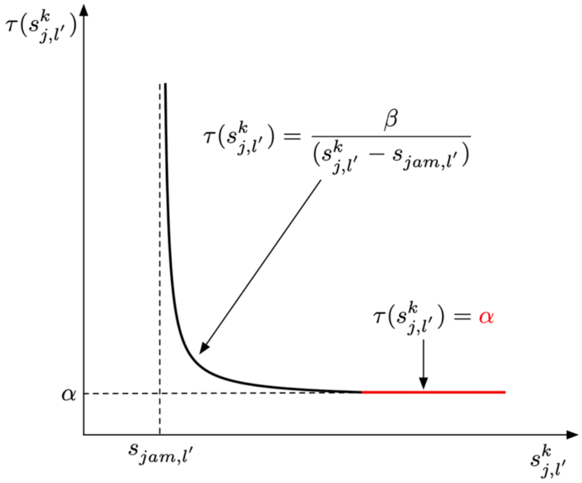


Fig. 2. State-dependent variable $\tau\left(s^k_{j,l'}\right)$ and its functional representation.

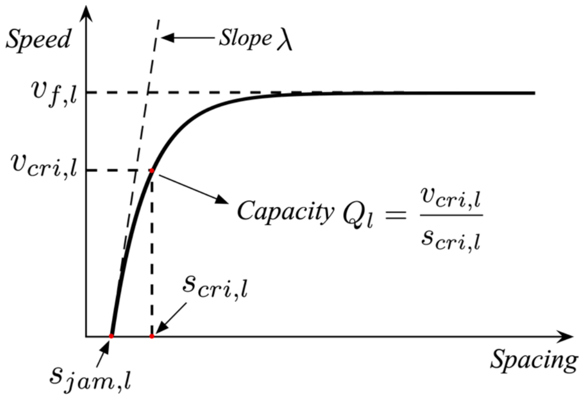


Fig. 3. The fundamental diagram of Newell's speed-spacing relation.

spacing in the target lane increases. However, a minimum threshold exists to prevent $\tau\left(s^k_{j,l'}\right)$ from decreasing indefinitely as spacing increases (Li et al., 2021; Wang et al., 2014). This threshold is essential because a minimum time is required for drivers to decide and complete a lane change, such as when the current lane is congested and the target lane is flowing freely, as discussed in Laval and Daganzo (2006). Any function meeting these criteria can theoretically describe $\tau\left(s^k_{j,l'}\right)$. For simplicity, we propose the following definition:

$$\tau\left(s^k_{j,l'}\right) = \max\left(\alpha, \frac{\beta}{\left(s^k_{j,l'} - s_{jam,l'}\right)}\right) \quad \alpha, \beta > 0 \tag{35}$$

where α denotes the minimum time required for a driver to decide and execute a lane change (in units of time), and β is an adjustment coefficient that regulates the sensitivity of $\tau\left(s^k_{j,l'}\right)$ to changes in spacing (in units of time–distance/vehicle). A higher β amplifies response of the $\tau\left(s^k_{j,l'}\right)$ to spacing variations, increasing its sensitivity, while a lower β reduces this sensitivity. Fig. 2 illustrates the functional representation of $\tau\left(s^k_{j,l'}\right)$.

The calibration of parameters α and β typically involves several key steps: dividing the dataset into training and testing sets, establishing a calibration objective, choosing an optimization method (such as genetic algorithms or quasi-Newton methods), iteratively adjusting the parameters on the training set, and validating the outcomes with the testing set. To evaluate the portability of parameters α and β, cross-validation is employed. This process generally involves dividing the dataset into K groups, with each group further split into a training set and a test set according to a predefined ratio. This ratio can be determined either empirically or based on the specific characteristics of the group. During each round of cross-validation, parameters are calibrated using one training set and their performance is evaluated across all K testing sets. This procedure is repeated for K rounds until all training sets have been used. To assess the portability of parameters across different testing sets, performance metrics such as Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) for spacing or the number of vehicles are recorded.

Additionally, the speed v_l is calculated using the speed-spacing relation $V_l^*(s_l)$ which can be described by Newell's speed-spacing relation (Newell, 1961), as shown in Eq. (36).

$$v_l = V_l^*(s_l) = v_{f,l}\left[1 - e^{\frac{-\lambda_l}{v_{f,l}}(s_l - s_{jam,l})}\right] \tag{36}$$

where $v_{f,l} > 0$ represents the free flow speed (in km/h) in lane l, $s_{jam,l}$ is the minimum safety distance (in km) in lane l, and λ_l is the slope (in veh/h) at $V_l^*(s_{jam,l}) = 0$. Fig. 3 provides an illustration of the speed-spacing relation.

In Fig. 3, $s_{cri,l}$ and $v_{cri,l}$ represent the critical spacing and critical speed of lane l, respectively. When the spacing between vehicles (or vehicle groups) reaches the critical spacing, the road volume reaches its maximum (capacity).

4.3. Net lane-changing number

4.3.1. Desired number of lane changes

Eq. (21) ensures that vehicles are conserved with lane changes. To estimate the number of lane changes $\varphi_{i,l}^k$, we introduce the lane-changing demand function $D_{i,l}^k$. This function is adapted from the traditional Eulerian sending function, as outlined by Daganzo (1994) and Munoz et al. (2003). It represents the number of vehicles per unit time from vehicle group i in lane l that intend to make forward moves, as given by:

$$D_{i,l}^k = \begin{cases} Q_l, s_{i,l}^k \le s_{cri,l} \\ \dfrac{V_l^*\left(s_{i,l}^k\right)}{s_{i,l}^k}, s_{i,l}^k > s_{cri,l} \end{cases} \tag{37}$$

Similarly, the lane-changing supply function $R_{i,l}^k$ is adapted from the traditional Eulerian receiving function, as discussed by Daganzo (1994) and Munoz et al. (2003). This function refers to the available space for vehicle group i in lane l. It is defined as the number of vehicles per unit time that can enter vehicle group i in lane l as given by:

$$R_{i,l}^k = \begin{cases} \dfrac{V_l^*\left(s_{i,l}^k\right)}{s_{i,l}^k}, s_{i,l}^k \le s_{cri,l} \\ Q_l, s_{i,l}^k > s_{cri,l} \end{cases} \tag{38}$$

where Q_l represents the capacity of lane l, $s_{i,l}^k$ represents the spacing of vehicle group i in lane l at time step k.

By utilizing the lane-changing probability described in Eq. (34), the number of drivers who desire to change lanes from vehicle group i in lane l to vehicle group j in the adjacent lane l' from time step k to $k+1$ can be calculated as:

$$L_{i,l\to j,l'}^k = D_{i,l}^k p_{i,l\to j,l'}^k \frac{Len_{(i,l),(j,l')}^k}{Len_{i,l}^k} \Delta t^2 \tag{39}$$

where $Len_{i,l}^k$ represents the length of vehicle group i in lane l at time step k, and $Len_{(i,l),(j,l')}^k$ represents the overlapping length between vehicle group i in lane l and vehicle group j in the adjacent lane l' at time step k.

Similarly, the number of drivers who desire to change lanes from vehicle group j in lane l' to vehicle group i in the adjacent lane l from time step k to $k+1$ can be calculated as:

$$L_{j,l'\to i,l}^k = D_{j,l'}^k p_{j,l'\to i,l}^k \frac{Len_{(j,l'),(i,l)}^k}{Len_{j,l'}^k} \Delta t^2 \tag{40}$$

4.3.2. Calculation of lane-changing number

The total demand of vehicles desiring to enter vehicle group i from adjacent lanes from time step k to $k+1$ is given by:

$$E_{i,l}^k = \sum_{l' \in \Omega} \sum_{j \in \Phi_{l'}} L_{j,l'\to i,l}^k \tag{41}$$

where Ω represents the set of adjacent lanes to which a vehicle from vehicle group i in lane l can change from time step k to $k+1$, and $\Phi_{l'}$ denotes the set of adjacent vehicle groups in lane l' that a vehicle from vehicle group i in lane l can join from time step k to $k+1$.

The total demand $E_{i,l}^k$ refers to the number of vehicles that vehicle group i in lane l can supply over the time interval Δt. The

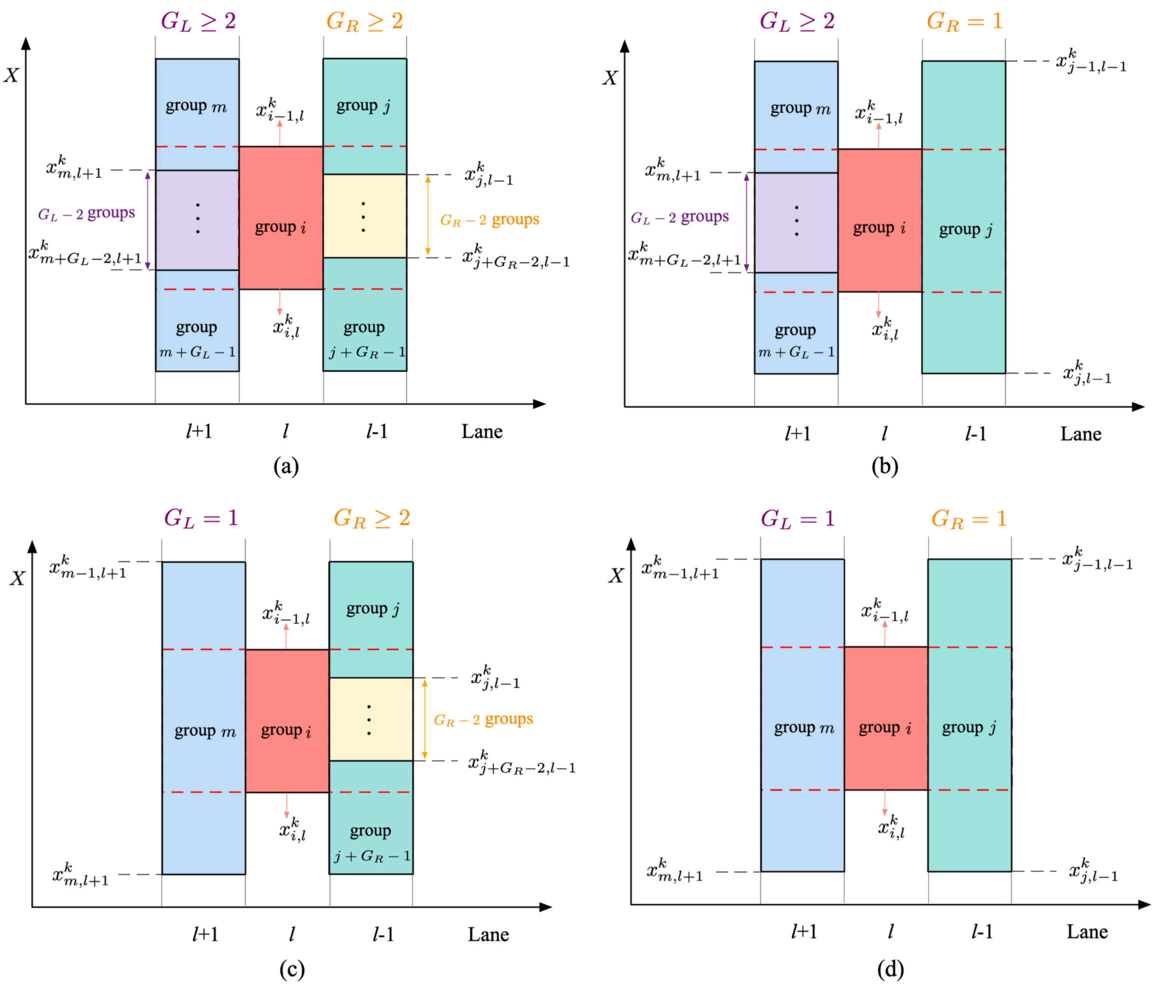


Fig. 4. Relative position of vehicle groups in adjacent lanes (The length between the two red dotted lines in each adjacent lane represents the portion of vehicles performing lane changes between vehicle group i in lane l and vehicle groups in adjacent lanes $l-1$ and $l+1$).

available capacity $R^k_{i,l}\Delta t$ represents the maximum number of vehicles that can be received by vehicle group i in lane l during the same time interval. To compute the number of vehicles entering $\left(M^k_{i,l}\right)$ and leaving $\left(N^k_{i,l}\right)$ vehicle group i in lane l, we determined the proportion of demand that can be accommodated. This approach is adapted from the incremental transfer principle in Eulerian coordinates, as described by Daganzo et al. (1997). The equation is provided as follows:

$$\gamma^k_{i,l} = \min\left(1, \frac{R^k_{i,l}\Delta t}{E^k_{i,l}}\right) \tag{42}$$

Eq. (42) indicates that if the total demand $E^k_{i,l}$ is less than the available capacity $R^k_{i,l}\Delta t$, all demands can be met. Otherwise, the available capacity $R^k_{i,l}\Delta t$ is distributed among the different originating vehicle groups according to their demands.

Therefore, the total number of vehicles entering vehicle group i in lane l from time step k to $k+1$ is given by:

$$M^k_{i,l} = \gamma^k_{i,l}\sum_{l'\in\Omega}\sum_{j\in\Phi_{l'}} L^k_{j,l'\to i,l} \tag{43}$$

The total number of vehicles leaving from vehicle group i in lane l to the adjacent lanes from time step k to $k+1$ is given by:

$$N^k_{i,l} = \sum_{l'\in\Omega}\sum_{j\in\Phi_{l'}}\left(\gamma^k_{j,l'} L^k_{i,l\to j,l'}\right) \tag{44}$$

The specific forms of $M^k_{i,l}$ and $N^k_{i,l}$ depend on the relative positions between vehicle group i and its adjacent vehicle groups in the adjacent lanes. For example, consider vehicle group i in the current lane l, with adjacent lanes $l-1$ (to the right) and $l+1$ (to the left). The values of $M^k_{i,l}$ and $N^k_{i,l}$ can be specified for four different cases, depending on the number of adjacent vehicle groups in lanes $l-1$ and

$l+1$. These cases are illustrated in Fig. 4.

Fig. 4(a) illustrates the case where multiple vehicle groups exist in both adjacent lanes. The horizontal axis represents the lane number, and the vertical axis represents the longitudinal position along the lane. Different vehicle groups are distinguished by color. In this case, vehicle group i is adjacent to vehicle groups j to $j+G_R-1$ in the right lane and groups m to $m+G_L-1$ in the left lane, where G_R and G_L represent the number of adjacent vehicle groups in the right and left lanes, respectively. This implies vehicle group i can perform lane changes with these groups during the time step from k to $k+1$. Similarly, Fig. 4(b) shows the case with a single vehicle group in the adjacent right lane and multiple vehicle groups in the adjacent left lane. Fig. 4(c) depicts the case with multiple vehicle groups in the adjacent right lane and a single vehicle group in the adjacent left lane, and Fig. 4(d) shows the situation with a single vehicle group in both the left and right adjacent lanes. Despite the differences in these scenarios, all four cases can be represented by a unified expression for calculating $M_{i,l}^k$ and $N_{i,l}^k$.

The number of vehicles desiring to change lanes from vehicle group h in lane $l-1$ to vehicle group i in lane l from time step k to $k+1$ can be calculated as:

$$L_{h,l-1\to i,l}^k=\begin{cases} D_{h,l-1}^k p_{h,l-1\to i,l}^k \dfrac{x_{i-1,l}^k-x_{i,l}^k}{s_{h,l-1}^k\Delta n}\Delta t^2 & , \quad h=j \text{ and } G_R=1 \\ D_{h,l-1}^k\cdot p_{h,l-1\to i,l}^k\cdot \dfrac{x_{i-1,l}^k-x_{h,l-1}^k}{s_{h,l-1}^k\Delta n}\Delta t^2 & , \quad h=j \text{ and } G_R\ge 2 \\ D_{h,l-1}^k\cdot p_{h,l-1\to i,l}^k\Delta t^2 & , \quad j<h<j+G_R-1 \text{ and } G_R\ge 2 \\ D_{h,l-1}^k\cdot p_{h,l-1\to i,l}^k\cdot \dfrac{x_{h-1,l-1}^k-x_{i,l}^k}{s_{h,l-1}^k\Delta n}\Delta t^2 & , \quad h=j+G_R-1 \text{ and } G_R\ge 2 \end{cases} \tag{45}$$

Similarly, the number of vehicles wishing to change lanes from vehicle group h' in lane $l+1$ to vehicle group i in lane l from time step k to $k+1$ can be calculated as:

$$L_{h',l+1\to i,l}^k=\begin{cases} D_{h',l+1}^k p_{h',l+1\to i,l}^k \dfrac{x_{i-1,l}^k-x_{i,l}^k}{s_{h',l+1}^k\Delta n}\Delta t^2 & , \quad h'=m \text{ and } G_L=1 \\ D_{h',l+1}^k\cdot p_{h',l+1\to i,l}^k\cdot \dfrac{x_{i-1,l}^k-x_{h',l+1}^k}{s_{h',l+1}^k\Delta n}\Delta t^2 & , \quad h'=m \text{ and } G_L\ge 2 \\ D_{h',l+1}^k\cdot p_{h',l+1\to i,l}^k\Delta t^2 & , \quad m<h'<m+G_L-1 \text{ and } G_L\ge 2 \\ D_{h',l+1}^k\cdot p_{h',l+1\to i,l}^k\cdot \dfrac{x_{h'-1,l+1}^k-x_{i,l}^k}{s_{h',l+1}^k\Delta n}\Delta t^2 & , \quad h'=m+G_L-1 \text{ and } G_L\ge 2 \end{cases} \tag{46}$$

where $G_R=1$ and $G_R\ge 2$ correspond to the cases with a single vehicle group and multiple vehicle groups in the adjacent right lane, respectively. Similarly, $G_L=1$ and $G_L\ge 2$ represent the cases of a single vehicle group and multiple vehicle groups in the adjacent left lane, respectively. $D_{h,l-1}^k$ and $D_{h',l+1}^k$ represent the demand of vehicle groups h and h' in lanes $l-1$ and $l+1$, respectively, as given by Eq. (37). The terms $p_{h,l-1\to i,l}^k$ and $p_{h',l+1\to i,l}^k$ represent the proportion of drivers who desire to change lanes from vehicle groups h and h' in lanes $l-1$ and $l+1$ to vehicle group i in lane l per unit time, as given by Eq. (34). Additionally, $x_{i,l}^k$ represents the position of the tail of vehicle group i in lane l at time step k, which can be obtained from the discretized form of Eq. (11), as given by:

$$\begin{cases} x_{i,l}^k=x_{i,l}^{k-1}+v_{i,l}^{k-1}\Delta t, & i=1 \\ x_{i,l}^k=x_{i-1,l}^k-s_{i,l}^k\Delta n, & i=2,3,4,\cdots,I_l^k \end{cases} \tag{47}$$

$x_{h,l+1}^k$ and $x_{h',l+1}^k$ can be obtained through similar methods.

Due to capacity restrictions, the number of vehicles entering vehicle group i in lane l from time step k to $k+1$ is limited and can be calculated as:

$$M_{i,l}^k=\gamma_{i,l}^k\left(\sum_{h=j}^{j+G_R-1}L_{h,l-1\to i,l}^k+\sum_{h'=m}^{m+G_L-1}L_{h',l+1\to i,l}^k\right) \tag{48}$$

where $\gamma_{i,l}^k$ represents the proportion of demand able to advance to vehicle group i in lane l from time step k to $k+1$, as given by Eq. (42).

Next, the number of vehicles leaving from vehicle group i in lane l to the adjacent lanes $l-1$ and $l+1$ is determined. The number of vehicles desiring to leave from vehicle group i in lane l to vehicle group h in the adjacent lane $l-1$ is given by:

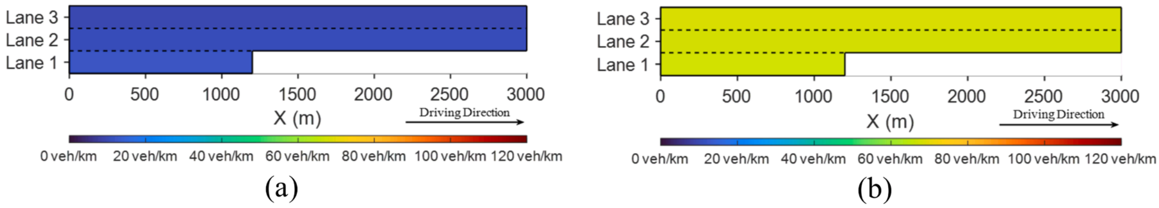


Fig. 5. Test scenario settings with initial snapshots: (a) low density; (2) high density.

$$L^k_{i,l\to h,l-1}=\begin{cases} D^k_{i,l}p^k_{i,l\to h,l-1}\Delta t^2 & , \quad h=j \text{ and } G_R=1\\ D^k_{i,l}\cdot p^k_{i,l\to h,l-1}\cdot\dfrac{x^k_{i-1,l}-x^k_{h,l-1}}{s^k_{i,l}\Delta n}\Delta t^2 & , \quad h=j \text{ and } G_R\geq 2\\ D^k_{i,l}\cdot p^k_{i,l\to h,l-1}\cdot\dfrac{s^k_{h,l-1}}{s^k_{i,l}}\Delta t^2 & , \quad j<h<j+G_R-1 \text{ and } G_R\geq 2\\ D^k_{i,l}\cdot p^k_{i,l\to h,l-1}\cdot\dfrac{x^k_{h-1,l-1}-x^k_{i,l}}{s^k_{i,l}\Delta n}\Delta t^2 & , \quad h=j+G_R-1 \text{ and } G_R\geq 2 \end{cases}\tag{49}$$

Similarly, the number of vehicles desiring to leave from vehicle group i in lane l to vehicle group h' in the adjacent lane $l+1$ is calculated as:

$$L^k_{i,l\to h',l+1}=\begin{cases} D^k_{i,l}p^k_{i,l\to h',l+1}\Delta t^2 & , \quad h'=m \text{ and } G_L=1\\ D^k_{i,l}\cdot p^k_{i,l\to h',l+1}\cdot\dfrac{x^k_{i-1,l}-x^k_{h',l+1}}{s^k_{i,l}\Delta n}\Delta t^2 & , \quad h'=m \text{ and } G_L\geq 2\\ D^k_{i,l}\cdot p^k_{i,l\to h',l+1}\cdot\dfrac{s^k_{h',l+1}}{s^k_{i,l}}\Delta t^2 & , \quad m<h'<m+G_L-1 \text{ and } G_L\geq 2\\ D^k_{i,l}\cdot p^k_{i,l\to h',l+1}\cdot\dfrac{x^k_{h'-1,l+1}-x^k_{i,l}}{s^k_{i,l}\Delta n}\Delta t^2 & , \quad h'=m+G_L-1 \text{ and } G_L\geq 2 \end{cases}\tag{50}$$

The total number of vehicles that successfully make lane changes from vehicle group i in lane l to lanes $l-1$ and $l+1$, considering capacity restrictions, is calculated as:

$$N^k_{i,l}=\sum_{h=j}^{j+G_R-1}\gamma^k_{h,l-1}L^k_{i,l\to h,l-1}+\sum_{h'=m}^{m+G_L-1}\gamma^k_{h',l+1}L^k_{i,l\to h',l+1}\tag{51}$$

where $\gamma^k_{h,l-1}$ and $\gamma^k_{h',l+1}$ represent the proportions of demand that can advance to lanes $l-1$ and $l+1$ from time step k to $k+1$, respectively, given by Eq. (42).

Eqs. (48) and (51) provide the specific forms of $M^k_{i,l}$ and $N^k_{i,l}$ for vehicle group i in lane l with adjacent lanes $l-1$ and $l+1$. This example can be easily extended to scenarios with only one adjacent lane or more than two adjacent lanes.

4.3.3. Calculation of net lane-changing number

The previously discussed Section 4.3.2 provided the calculation for the total number of vehicles entering (or leaving) vehicle group i, denoted as $M^k_{i,l}$ (or $N^k_{i,l}$). Based on this, the net lane-changing number for the vehicle group i in lane l from time step k to $k+1$, represented as $\varphi^k_{i,l}$, can be expressed as:

$$\varphi^k_{i,l}=M^k_{i,l}-N^k_{i,l}\tag{52}$$

By substituting Eq. (52) into Eq. (21), the discretized spacing dynamics can be rewritten as:

$$s^{k+1}_{i,l}=s^k_{i,l}+\frac{\Delta t}{\Delta n}\left(v^k_{i-1,l}-v^k_{i,l}\right)-\frac{s^k_{i,l}}{\Delta n}\left(M^k_{i,l}-N^k_{i,l}\right)\tag{53}$$

Eq. (53) allows for the dynamic iteration of the spacing of each vehicle group in each lane over time.

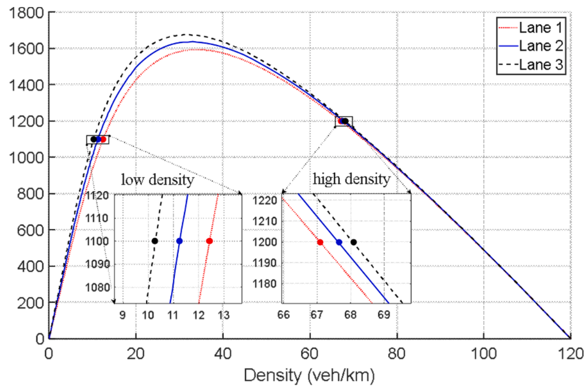


Fig. 6. Fundamental diagram of three lanes.

Table 1
Parameters and numerical settings.

Parameters	Value
Free flow speed of the outer lane (lane 1), v_{f1}	100 km/h
Free flow speed of the middle lane (lane 2), v_{f2}	110 km/h
Free flow speed of the inner lane (lane 3), v_{f3}	120 km/h
Minimum safety distance of lane l, $s_{jam,l}(l = 1, 2, 3)$	$8.33 \times 10^{-3} km$
Critical spacing of the outer lane, $s_{cri,1}$	0.0294 km
Critical spacing of the middle lane, $s_{cri,2}$	0.0306 km
Critical spacing of the inner lane, $s_{cri,3}$	0.0316 km
Slope at $V_l^*(s_{jam,l}) = 0$, $\lambda_l(l = 1, 2, 3)$	3000 veh/h
Time steps	3000
Time step size Δt	0.2 s
Vehicle group size Δn	1

5. Numerical example

5.1. Scenario settings and parameters

In this section, we present a synthetic example utilizing a $L = 3km$ three-lanes highway section with a lane-drop occurring at $X = 1.2km$, as depicted in Fig. 5. Each lane follows the fundamental diagram defined by Eq. (36). We examine two cases: one with low-density traffic and another with high-density traffic, corresponding to Fig. 5(a) and Fig. 5(b) respectively, as shown in Fig. 6.

Case 1: For the low-density condition, we set the initial densities for the outer lane (lane 1), middle lane (lane 2) and inner lane (lane 3) to 12.4 veh/km, 11.2 veh/km and 10.2 veh/km, respectively. These density values correspond to a traffic volume of 1100veh/h on the left side of the fundamental diagram presented in Fig. 6. Regarding the inflow boundary described by Eq. (25), the inflow demand remains constant, given by $A_l^k(0^-) = 1000veh/h$, where $l = 1,2,3$. On the other hand, the outflow boundary, depicted in Eq. (30), maintains a constant outflow supply with a capacity of $B_l^k(3^+) = Q_l$, where $l = 2, 3$. In lane 1 downstream of the lane-drop location ($X = 1.2km$), the outflow supply is given by $B_l^k(1.2^+) = 0$, where $l = 1$. The value of $B_l^k(0^+)$, where $l = 1,2,3$, depends on the spacing of the last vehicle group in each lane. $A_l^k(1.2^-)$, where $l = 1$ and $A_l^k(3^-)$, where $l = 2, 3$ depend on the spacing of the first vehicle group in each lane.

Case 2: For the high-density condition, we set the initial densities for the outer lane (lane 1), middle lane (lane 2) and inner lane (lane 3) to 67.1 veh/km, 67.6 veh/km and 68.1 veh/km, respectively. These density values correspond to a traffic volume of 1200 veh/h on the right side of the fundamental diagram shown in Fig. 6. Regarding the inflow boundary, we maintain a constant inflow demand, denoted by $A_l^k(0^-) = 1200veh/h$, where $l = 1,2,3$. Similarly, the outflow boundary maintains a constant outflow supply with a capacity of $B_l^k(3^+) = Q_l$, where $l = 2,3$. In lane 1 downstream of the lane-drop location ($X = 1.2km$), the outflow supply is given by $B_l^k(1.2^+) = 0$, where $l = 1$. The value of $B_l^k(0^+)$, where $l = 1, 2, 3$, depends on the spacing of the last vehicle group in each lane. Additionally, $A_l^k(1.2^-)$, where $l = 1$ and $A_l^k(3^-)$, where $l = 2, 3$ depend on the spacing of the first vehicle group in each lane.

The parameters α and β of $\tau\left(s_{j,l'}^k\right)$ in Eq. (35) represent the minimum time required for a driver to decide and execute a lane change, and an adjustment coefficient that regulates the sensitivity of $\tau\left(s_{j,l'}^k\right)$ to changes in spacing, respectively. These parameters can be calibrated using field data, such as vehicle trajectory data. In this numerical example, we set α to be 2 s and β for 0.2 $km\cdots/veh$ Table 1 provides the parameters and numerical settings used in this study.

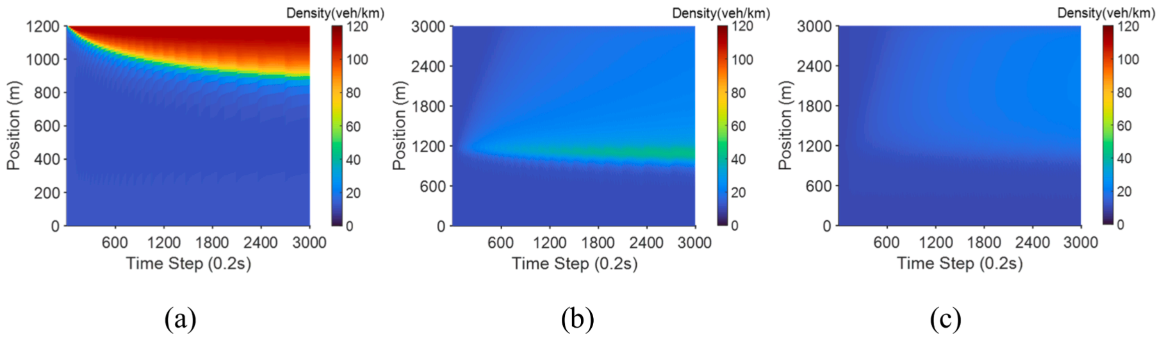


Fig. 7. Density dynamics of different lanes for case 1: (a) Lane1; (b) Lane 2; (c) Lane 3.

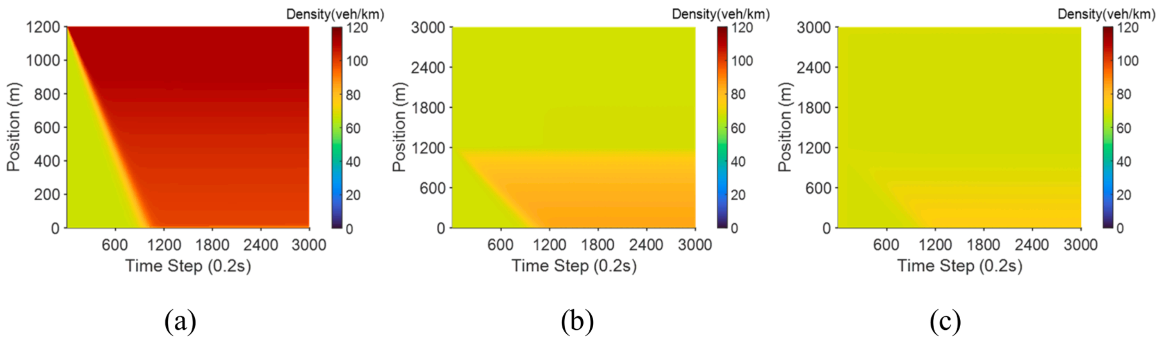


Fig. 8. Density dynamics of different lanes for case 2: (a) Lane1; (b) Lane 2; (c) Lane 3.

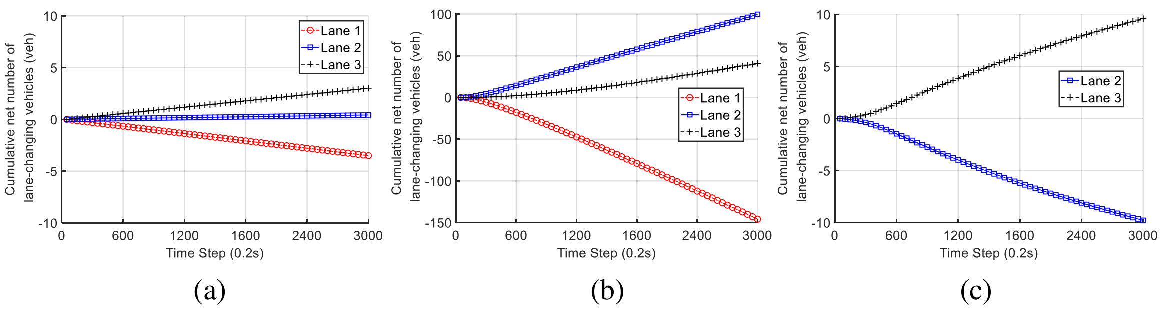


Fig. 9. Cumulative net number of the lane-changing vehicles over time on three highway sections for case 1: (a) upstream of the bottleneck (0.2–0.6 km); (b) near the bottleneck (0.8–1.2 km); (c) downstream of the bottleneck (1.4–1.8 m).

5.2. Discussion of simulation results

5.2.1. Impact of lane-changing behavior on traffic flow dynamics

In this experiment, vehicles in lane 1 are required to change lanes at the bottleneck of the lane-drop, potentially leading to traffic congestion in lane 1 and its propagation upstream. Additionally, lane-changing maneuvers can disturb the traffic flow in lanes 2 and 3. The density dynamics of each lane in the low-density condition (case 1) are depicted in Fig. 7. As shown in Fig. 7(a), the proposed Lagrangian model accurately predicts traffic congestion at the bottleneck around 1200 m in lane 1, with the congestion shockwave gradually propagating upstream. Furthermore, the proposed model captures the traffic perturbation caused by lane-changing behavior from lane 1 to lane 2, as evident in Fig. 7(b), and this perturbation propagates downstream. However, the perturbation in lane 3 is minimal, as depicted in Fig. 7(c). Fig. 8 presents the density dynamics of each lane in the high-density condition (case 2). In this case, the proposed model successfully predicts traffic congestion at the bottleneck of lane 1, with a distinct congestion shockwave boundary (less diffusion compared to existing Eulerian models) as shown in Fig. 8(a). This observation aligns with previous studies (Van Wageningen-Kessels et al., 2010; van Wageningen-Kessels et al., 2009), which suggest that the Lagrangian formulation reduces numerical diffusion. The model also predicts traffic congestion in lane 2 resulting from lane-changing behavior from lane 1 to lane 2 (see Fig. 8). However, the perturbation in lane 3 is minimal, indicating that only a few vehicles opt to switch from lane 2 to lane 3, as

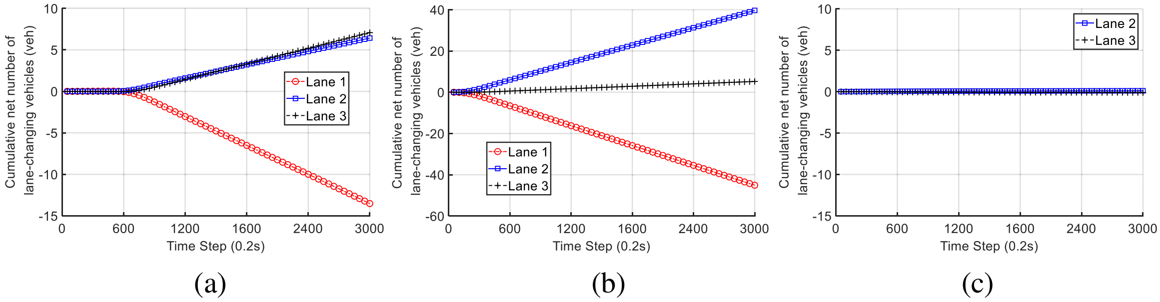


Fig. 10. Cumulative net number of the lane-changing vehicles over time on three highway sections for case 2: (a) upstream of the bottleneck (0.2–0.6 km); (b) near the bottleneck (0.8–1.2 km); (c) downstream of the bottleneck (1.4–1.8 m).

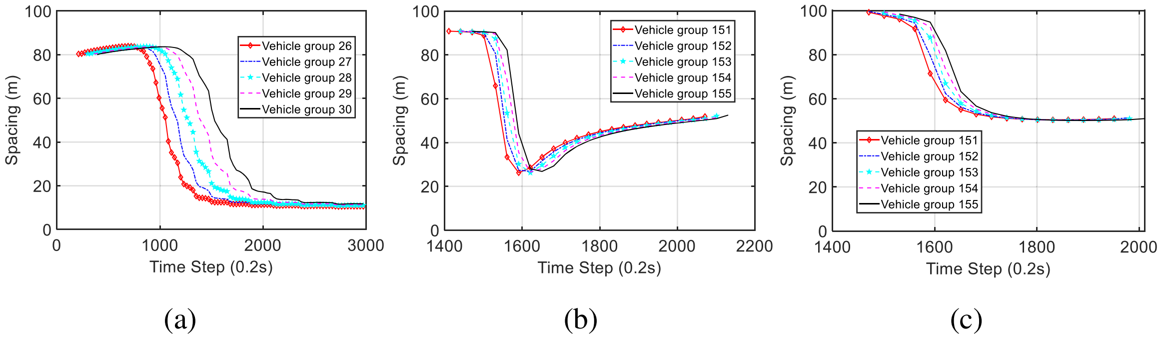


Fig. 11. Spacing dynamics of the 5 vehicle groups for case 1: (a) Lane 1; (b) Lane 2; (c) Lane 3.

depicted in Fig. 8.

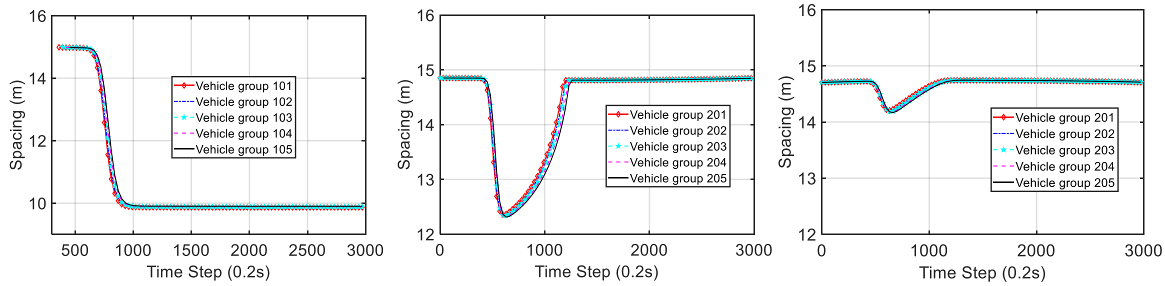
5.2.2. *Lane-changing characteristics*

To investigate the lane-changing characteristics on different sections of the highway, we examine three sections: upstream (0.2 - 0.6 km), near the bottleneck (0.8 - 1.2 km) and downstream (1.4 - 1.8 km). We define the cumulative net number of lane-changing vehicles, $\eta^{k}_{hs,l}$, in lane l at time step k for highway section hs ($hs\epsilon\{upstream, bottleneck,\ downstream\}$), as given by:

$$\eta^{k}_{hs,l} = \sum_{k'=0}^{k} \left(M^{k'}_{hs,l} - N^{k'}_{hs,l} \right) \tag{54}$$

where $M^{k'}_{hs,l}$ and $N^{k'}_{hs,l}$ represent the number of vehicles entering and leaving lane l of section hs at time step k', respectively.

We plot the cumulative net number of the lane-changing vehicles, where the ascending curve above zero represents the cumulative net number of vehicles entering the current lane, while the descending curve represents the cumulative net number of vehicles leaving the current lane. Fig. 9 shows the cumulative net number of the lane-changing vehicles for each lane in the low-density condition (case 1). As observed in Fig. 9(a), vehicles in the far upstream section tend to stay in the current lane and make very few lane changes since no congestion propagates to this section, which is also evident in Fig. 7(a). However, the vehicles in lane 1 at the bottleneck section are required to make lane changes to the adjacent lane (e.g., lane 2) due to the lane-drop. Consequently, the cumulative net number of the lane-changing vehicles for lane 1 shows a steep descending curve, while an ascending curve can be observed for lane 2. The impact of the lane-drop on the traffic dynamics of lane 3 is less significant compared to lane 2, as shown in Fig. 9(b). At the downstream section of the bottleneck, vehicles tend to change lanes from lane 2 to lane 3, as depicted in Fig. 9(c). This behavior is attributed to the higher density in lane 2 resulting from vehicles changing lanes from lane 1 to lane 2 at the bottleneck, which prompts vehicles to change from lane 2 to lane 3 in pursuit of higher speeds. Fig. 10 presents the cumulative curve of the net number of lane-changing vehicles in the high-density condition (case 2). At the upstream of the bottleneck (see Fig. 10(a)), vehicles initially continue driving in their current lanes. However, as congestion from the downstream bottleneck propagates upstream, vehicles in lane 1 change lanes to lanes 2 or 3, resulting in a descending curve for lane 1 and ascending curves for lanes 2 and 3. At the bottleneck section, as shown in Fig. 10(b), the influence of the lane-drop on traffic dynamics of lane 3 is minimal, with only a few lane-changing vehicles. Meanwhile, at the downstream section of the bottleneck, the traffic reaches an equilibrium state, where the cumulative net number of lane-changing vehicles is zero, indicating that vehicles in both lanes tend to continue driving in their current lanes, as observed in Fig. 10(c).



(a) (b) (c)

Fig. 12. Spacing dynamics of the 5 vehicle groups for case 2: (a) Lane 1; (b) Lane 2; (c) Lane 3.

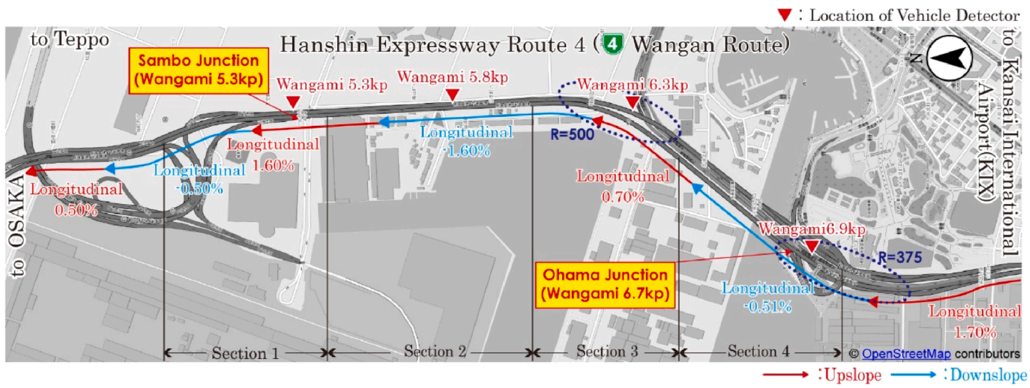


Fig. 13. Study area.

5.2.3. Spacing dynamics

Fig. 11 depicts the spacing dynamics of five vehicle groups in lane 1 (Fig. 11(a)), lane 2 (Fig. 11(b)), and lane 3 (Fig. 11(c)) in the low-density condition (case 1). In Fig. 11(a), the vehicle groups in lane 1 initially maintain a stationary state with nearly constant spacing until the congestion shockwave propagates upstream from the bottleneck. As a result, the spacings decrease to low values corresponding to the congested condition. Fig. 11(b) shows that initially, the vehicle groups in lane 2 drive with a constant spacing. However, when these vehicle groups encounter the perturbation caused by lane-changing maneuvers of vehicles moving from lane 1 to lane 2, the spacings in lane 2 substantially decrease. After passing through the bottleneck, the spacings gradually increase to a constant value that is lower than the initial spacing. This phenomenon can be attributed to the higher traffic density downstream of the bottleneck (with 2 lanes) compared to upstream (with 3 lanes). Fig. 11(c) illustrates the spacing dynamics of the vehicle groups in lane 3. Here, the spacings decrease due to vehicles changing from lane 2 to lane 3 at the bottleneck section, resulting in an increased density in lane 3. Subsequently, the vehicle groups maintain nearly constant spacing downstream of the bottleneck section.

Fig. 12 presents the spacing dynamics of five vehicle groups for each lane in the high-density condition (case 2). In Fig. 12(a), a significant decrease in the spacings of vehicle groups in lane 1 is observed at the bottleneck section. Fig. 12(b) shows that the vehicle groups in lane 2 initially maintain a stationary state with constant spacings, which sharply decrease at the bottleneck section due to vehicles changing from lane 1 to lane 2, leading to increased density in lane 2. Downstream of the bottleneck, the spacings of all vehicle groups increase to a constant value, indicating a steady state. Similar results can be observed in Fig. 12(c), where the spacings of vehicle groups decrease due to vehicles changing from lane 2 to lane 3 upstream of the bottleneck.

6. Experiment with ZTD data

In this section, we validate the proposed model using Zen Traffic Data (ZTD) (Japan Hanshin Expressway Company, 2019), which provides detailed vehicle trajectory data from the Hanshin Expressway Company. Firstly, we estimate the parameters of the fundamental diagram (FD) using the vehicle trajectory data. Then, we divide the vehicle trajectory data into a training data set and a testing data set. The training dataset is used to calibrate the parameters of the state-dependent variable $\tau\left(s_{j,l}^{k}\right)$ under different traffic conditions, while the testing dataset is used to verify the proposed model. To measure the quantitative performance of the proposed model, we calculate the Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) for the lane volume and vehicle spacing. Finally, we compare the cumulative net number of lane-changing vehicles derived from our model with the field data from ZTD.

Table 2
Estimated parameters of FD.

	$v_{f,l}(km/h)$	$\lambda_l(veh/h)$	$s_{jam,l}(km)$	$s_{cri,l}(km)$	$v_{cri,l}(km/h)$
Lane 1	66.3	3384.8	8.0×10^{-3}	0.023	36.1
Lane 2	85.6	4139.4	7.9×10^{-3}	0.024	45.6

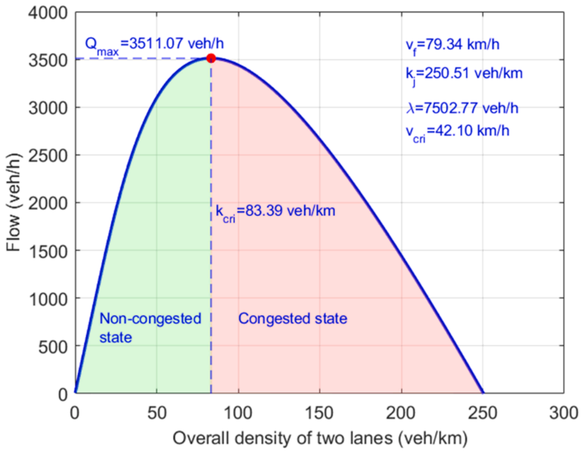


Fig. 14. Newell's fundamental diagram of two lanes.

Table 3
The period spanned by each subset.

Datasets	Non-congestion		Congestion	
	Training set	Testing set	Training set	Testing set
L002_F001	0 - 1200 (s)	1200 - 1500 (s)	2100 - 3000 (s)	3000 - 3300 (s)
L002_F002	0 - 1200 (s)	1200 - 1500 (s)	2400 - 3300 (s)	3300 - 3600 (s)
L002_F003	0 - 600 (s)	600 - 900 (s)	1200 - 2100 (s)	2100 - 2400 (s)
L002_F004	1500 - 2400 (s)	2400 - 2700 (s)	300 - 1200 (s)	1200 - 1500 (s)
L002_F005	0 - 1200 (s)	1200 - 1500 (s)	2100 - 3000 (s)	3000 - 3300 (s)

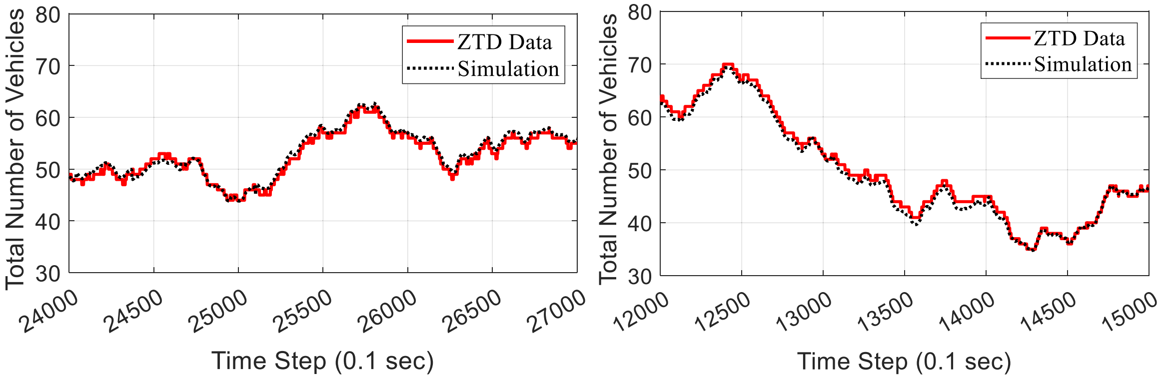
Table 4
Parameters of lane choice probability.

Datasets	Non-congestion		Congestion	
	$\alpha(s)$	$\beta(km \cdot s/veh)$	$\alpha(s)$	$\beta(km \cdot s/veh)$
L002_F001	1.3	0.03	2.1	0.07
L002_F002	1.5	0.02	1.9	0.06
L002_F003	1.4	0.01	1.8	0.05
L002_F004	1.2	0.01	1.7	0.04
L002_F005	1.5	0.01	1.8	0.06

6.1. Parameters calibration

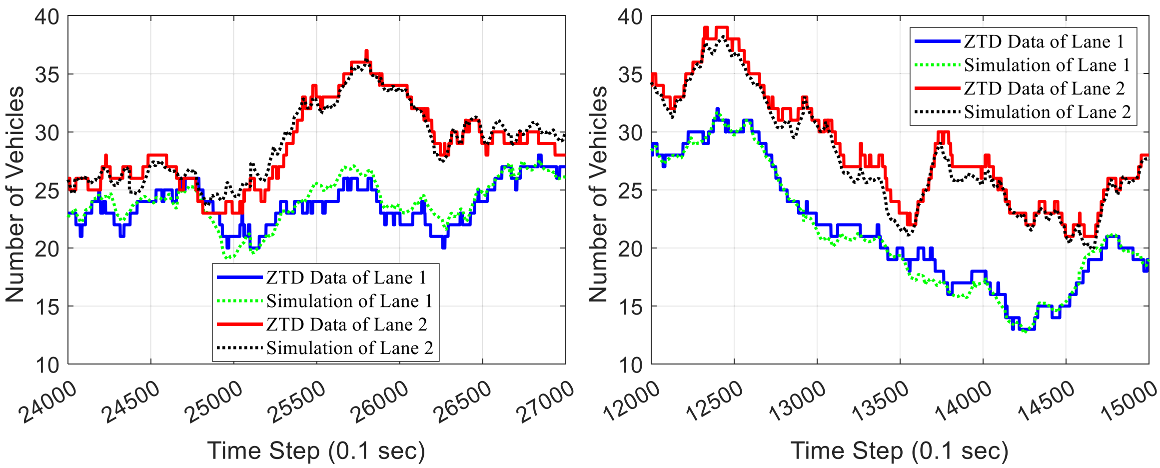
We utilize the vehicle trajectory data from the ZTD dataset, collected on the Osaka-bound direction of Wangan Route at the Ohama-Sambo section, ranging from 6.4 *kp* (distance from starting point of the expressway route, Unit: kilometers) to 5.6 *kp* in Japan (as shown in Fig. 13). The data has a high temporal resolution of 0.1 s. The selected section excludes the on-ramp and off-ramp since the proposed model does not consider merges and diverges. The ZTD dataset consists of five one-hour vehicle trajectory datasets, namely L002_F001, L002_F002, L002_F002, L002_F003, L002_F004, and L002_F005. Each dataset provides detailed vehicle trajectory data for approximately an hour (from 7:00 AM to 8:00 AM) over a 1.7-km distance. It should be noted that the proportion of large vehicles in each dataset is extremely low (*<0.58 %*), thereby satisfying assumption 4 from Section 4.1. The lanes are numbered from left to right, with lane 1 being the leftmost and lane 2 being the rightmost.

Two types of parameters are calibrated based on the vehicle trajectory data. The first type includes the parameters of the fundamental diagram (FD), namely the free flow speed $v_{f,l}$, the minimum safety distance $s_{jam,l}$, and the slope at $V_l^*(s_{jam,l}) = 0$, λ_l. We assume



(a) (b)

Fig. 15. Comparison between observations from L002_F004 dataset and the predicted total number of vehicles in: (a) the non-congested condition; (b) the congested condition.



(a) (b)

Fig. 16. Comparison between observations from the L002_F004 dataset and model predicted number of vehicles for each lane: (a) in the non-congested condition; (b) in the congested condition.

the FD (speed-spacing) relation follows Newell's car following model, as described by Eq. (36).

However, the parameter values of the FD cannot be directly calibrated using vehicle trajectory data. To address this issue, we convert vehicle trajectory data into aggregated data that can be directly used by the FD model, such as the space mean speed and average spacing. We employ a generalized definition of traffic states in a time-space region *Reg*, proposed by Edie (1963), given by:

$$v(\mathrm{Reg}) = \frac{d(\mathrm{Reg})}{t(\mathrm{Reg})} \tag{55}$$

$$s(\mathrm{Reg}) = \frac{|\mathrm{Reg}|}{t(\mathrm{Reg})} \tag{56}$$

where $v(Reg)$ represents the space mean speed of all vehicles in region *Reg* (km/h), $d(Reg)$ is the total distance traveled by all vehicles in region *Reg* (km), $t(Reg)$ is the total time spent by all vehicles in region *Reg* (h), $s(Reg)$ is the average spacing of all vehicles in region *Reg* (km), and $|Reg|$ is the area of region *Reg* ($km \cdot h$).

Before using the vehicle trajectory dataset, we first identified and flagged any anomalous trajectories. These anomalies were categorized into two types: acceleration anomalies, where the acceleration values were outside the range of [-6 m/s^2, 5 m/s^2] as recommended by Montanino and Punzo (2013), and fragmented data segments lasting $<$5 s. Trajectories flagged with these anomalies were excluded from the analysis to ensure data quality.

For lane l, we divide the space-time area ($0.8\ km \times\ 5\ h$), comprising five one-hour vehicle trajectory datasets, into several

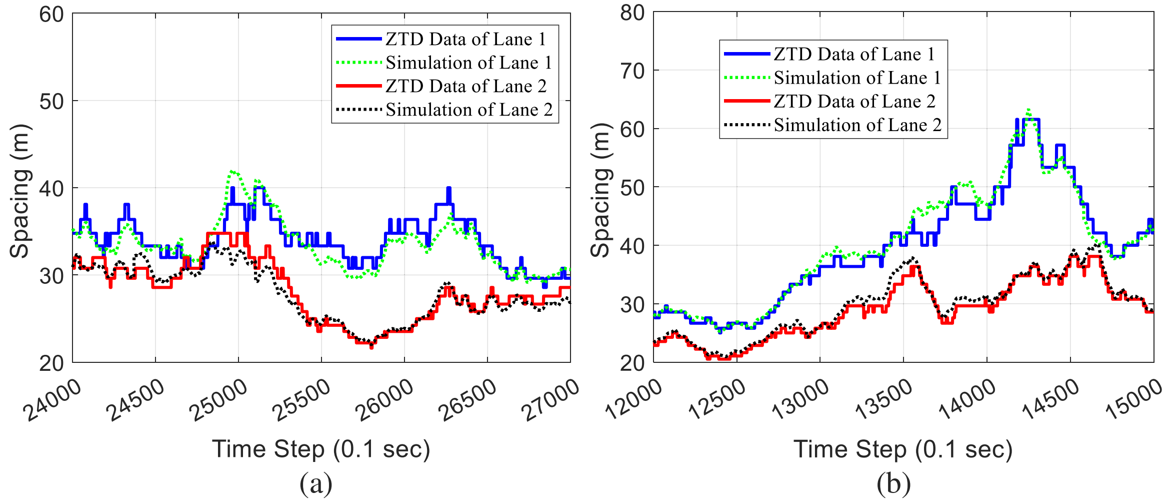


Fig. 17. Comparison between observations from the L002_F004 dataset and model predicted spacings for each lane: (a) in the non-congested condition; (b) in the congested condition.

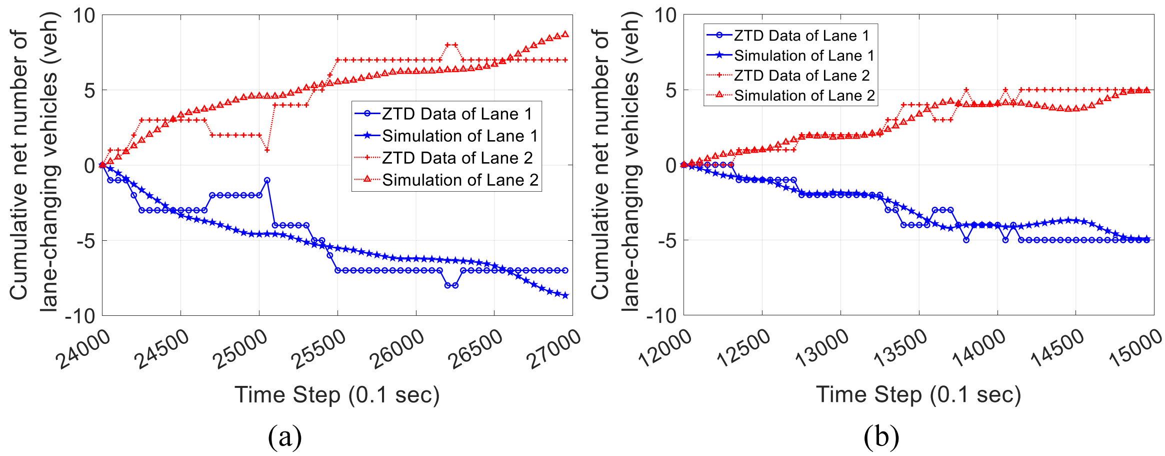


Fig. 18. Cumulative net number of lane-changing vehicles over time: (a) non-congested condition of L002_F004 dataset; (b) congested condition of L002_F004 dataset.

Table 5
Performance metrics for the estimated number of vehicles from the L002_F001 dataset across 200-meter intervals.

	Non-congestion				congestion			
	Lane1		Lane2		Lane1		Lane2	
	RMSE (veh)	MAPE (%)	RMSE (veh)	MAPE (%)	RMSE (veh)	MAPE (%)	RMSE (veh)	MAPE (%)
0–200m	0.56	8.09	0.76	7.73	0.54	6.10	0.70	7.04
200–400m	0.58	8.10	0.54	6.25	0.60	6.93	0.74	7.25
400–600m	0.52	7.95	0.53	6.19	0.55	5.64	0.41	5.45
600–800m	0.36	5.66	0.40	4.82	0.52	5.83	0.51	6.87

subregions with a distance interval of $dx = 200m$ and a time interval of $dt = 5s$. A subregion is excluded from the analysis and deemed unreliable if >10 % of its vehicle trajectories are anomalous. The remaining valid subregions are then counted, resulting in a total number denoted as Z_l. For each subregion $z \in Z_l$ in lane l, we can calculate the space mean speed v_{zl} and spacing s_{zl} using Eqs. (55) and (56), respectively. We denote the unknown parameter vector in the speed-spacing relation $V_l^*(s_l)$ of lane l as $\overrightarrow{\Psi} = [v_{f,l}, \lambda_l, s_{jam,l}]$. The estimated parameter vector $\overrightarrow{\Psi^*}$ is obtained by minimizing the objective function:

Table 6
Performance metrics for the estimated spacings from the L002_F001 dataset across 200-meter intervals.

	Non-congestion				congestion			
	Lane1		Lane2		Lane1		Lane2	
	RMSE (m)	MAPE (%)	RMSE (m)	MAPE (%)	RMSE (m)	MAPE (%)	RMSE (m)	MAPE (%)
0–200m	3.90	7.81	3.84	6.64	2.03	6.77	2.40	7.16
200–400m	4.71	8.38	3.56	6.57	3.22	7.75	3.04	6.36
400–600m	4.51	7.94	2.73	6.51	2.85	6.74	2.64	5.47
600–800m	3.12	5.64	1.93	4.91	2.78	6.16	4.03	7.62

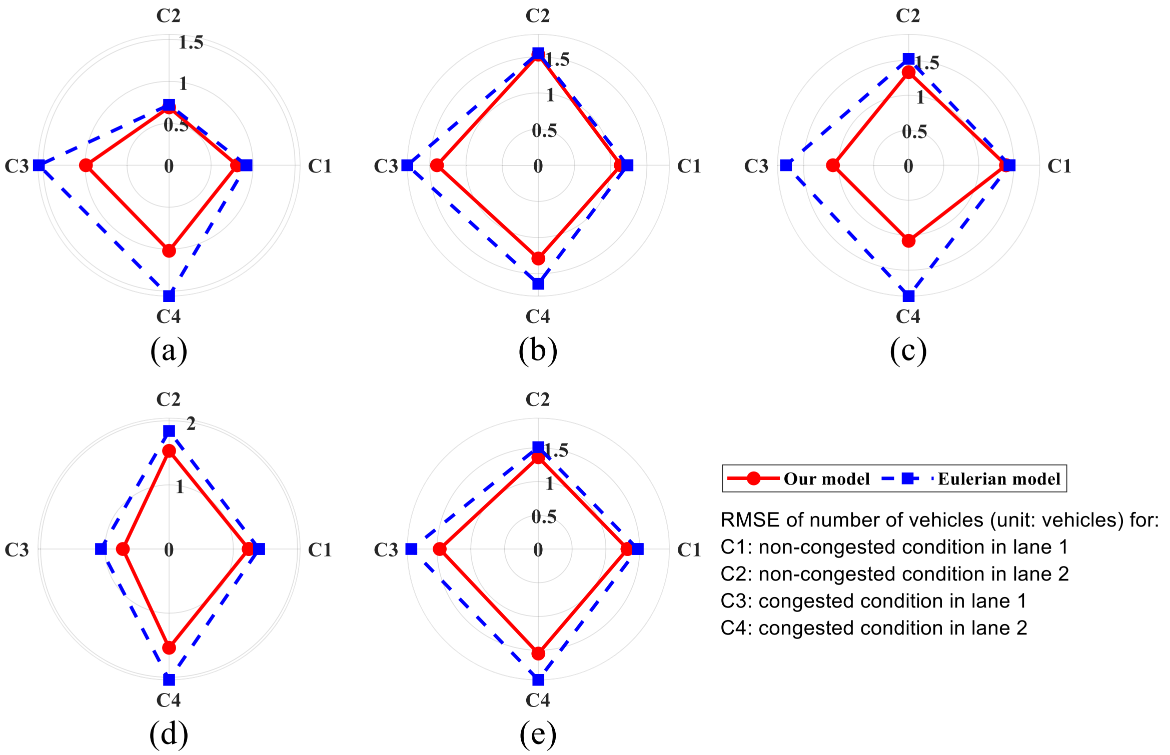


Fig. 19. Comparison of RMSE for the number of vehicles across different lanes and traffic conditions in the following datasets: (a) L002_F001 dataset; (b) L002_F002 dataset; (c) L002_F003 dataset; (d) L002_F004 dataset; (e) L002_F005 dataset.

$$\overrightarrow{\Psi^*} = \mathrm{argmin}_{\overrightarrow{\Psi}}\left[\sum_{z=1}^{Z_l}\left\{v_{zl} - V_l^*(s_{zl}|\overrightarrow{\Psi})\right\}^2\right] \tag{57}$$

To solve Eq. (57), we apply the quasi-Newton method with the initial value set to $\overrightarrow{\Psi^0} = [100, 3000, 5 \ \times 10^{-3}]$ for lane $l = 1{,}2$. The estimated results are shown in Table 2.

The second type of parameters consists of the parameters of the state-dependent variable $\tau\left(s_{j,l'}^k\right)$, denoted as $\overrightarrow{\Phi} = [\alpha, \beta]$. We denote the estimated number of vehicles in lane l as $Num_{est,l}^k(\overrightarrow{\Phi})(l = 1, 2)$ and the observed number of vehicles in lane l as $Num_{obs,l}^k(l = 1, 2)$ in lane l at time step k.The parameters $\overrightarrow{\Phi^*}$ are estimated using the following objective function:

$$\overrightarrow{\Phi^*} = \mathrm{argmin}_{\overrightarrow{\Phi}}\left[\sum_{l=1}^{2}\sum_{k_1}^{k_2}\left\{Num_{obs,l}^k - Num_{est,l}^k(\overrightarrow{\Phi})\right\}^2\right] \tag{58}$$

where $\overrightarrow{\Phi} = [\alpha, \beta]$ represents the unknown parameter vector. The time step k corresponds to time $t = t_0 + k\Delta t$. In this case, t_0 and Δt take the values of $0s$ and $0.1s$, respectively. k_1 and k_2 represent the time steps at the beginning and end of the given dataset, respectively. We employ the quasi-Newton method to solve Eq. (58), with the initial value set to $\overrightarrow{\Phi^0} = [2, 0.2]$.

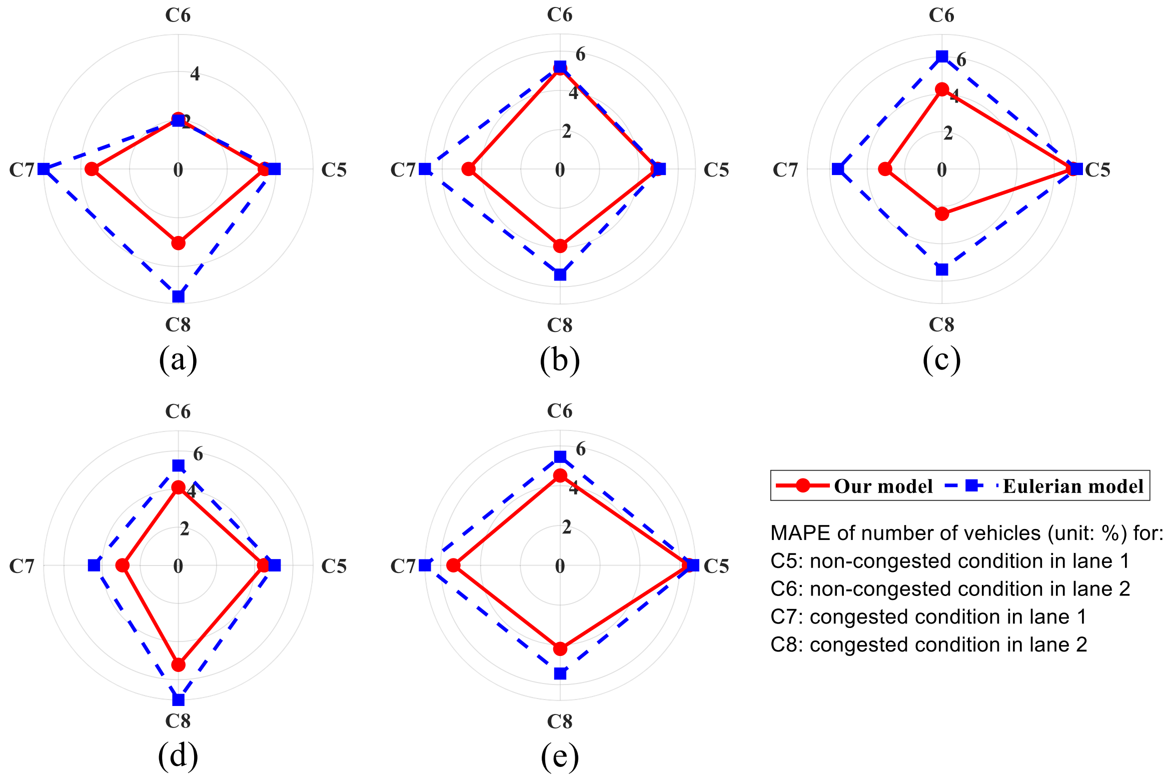


Fig. 20. Comparison of MAPE for the number of vehicles across different lanes and traffic conditions in the following datasets: (a) L002_F001 dataset; (b) L002_F002 dataset; (c) L002_F003 dataset; (d) L002_F004 dataset; (e) L002_F005 dataset.

The parameters of the state-dependent variable $\tau\left(s_{j,l'}^{k}\right)$ are calibrated considering two different traffic states: congestion and non-congestion. A congested state is defined by a density greater than $83.39veh/km$, while densities below this threshold are classified as non-congested. This classification is based on the criteria established in Fig. 14, where Newell's FD (Newell, 1961) is calibrated using five one-hour vehicle trajectory datasets. Unlike the single-lane calibration results shown in Table 2, Fig. 14 integrates data from two lanes, treating them as a single unit or 'pipe'. This combined analysis facilitates the segmentation of the dataset into distinct traffic states. Each one-hour dataset is further divided into 5-minute intervals. Intervals with an average density above $83.39veh\ /km$ are classified as congested, while others are classified as non-congested. Consecutive intervals of the same traffic state are merged into unified entries, resulting in the final datasets for congested and non-congested states. Each traffic state dataset is then divided into a training set for model calibration and a testing set for validation, yielding four sub-datasets: training and testing sets for both the non-congested and congested conditions. The specific time periods for each subset are provided in Table 3. Certain time intervals are excluded from the analysis, such as the 1500 - 1800 s interval in L002_F001, due to the presence of abnormal vehicle trajectory data. Specifically, intervals are deemed unreliable and exclude if the proportion of abnormal trajectories exceeds 10 % of the total vehicle count within that interval. For example, the aforementioned interval contains 17.3 % (13 trajectories) of abnormal vehicle trajectories. This exclusion criterion is consistently applied throughout the dataset to maintain the quality and reliability of the data used in our study.

Table 4 presents the results of the estimated parameter vector $\overrightarrow{\Phi^*}$ for the five datasets under both non-congested and congested conditions. As shown in Table 4, the α and β values are larger under congestion compared to non-congestion, indicating that drivers require more time to make and execute lane changes in congested traffic. Fig. 15 provides a comparison between the total number of vehicles observed in the L002_F004 dataset and the predictions made by the model for both traffic conditions. The model's predictions closely match the field observations.

6.2. Experimental results

6.2.1. Performance evaluation of the proposed model

To assess the model's performance, we evaluated its ability to estimate the number of vehicles and spacing using the L002_F004 dataset as an example. Fig. 16 shows a comparison between the observed and estimated vehicle counts over time for each lane under different traffic conditions. The results indicate a high level of consistency. Fig. 17 displays the model's performance in estimating spacings, demonstrating accurate estimations.

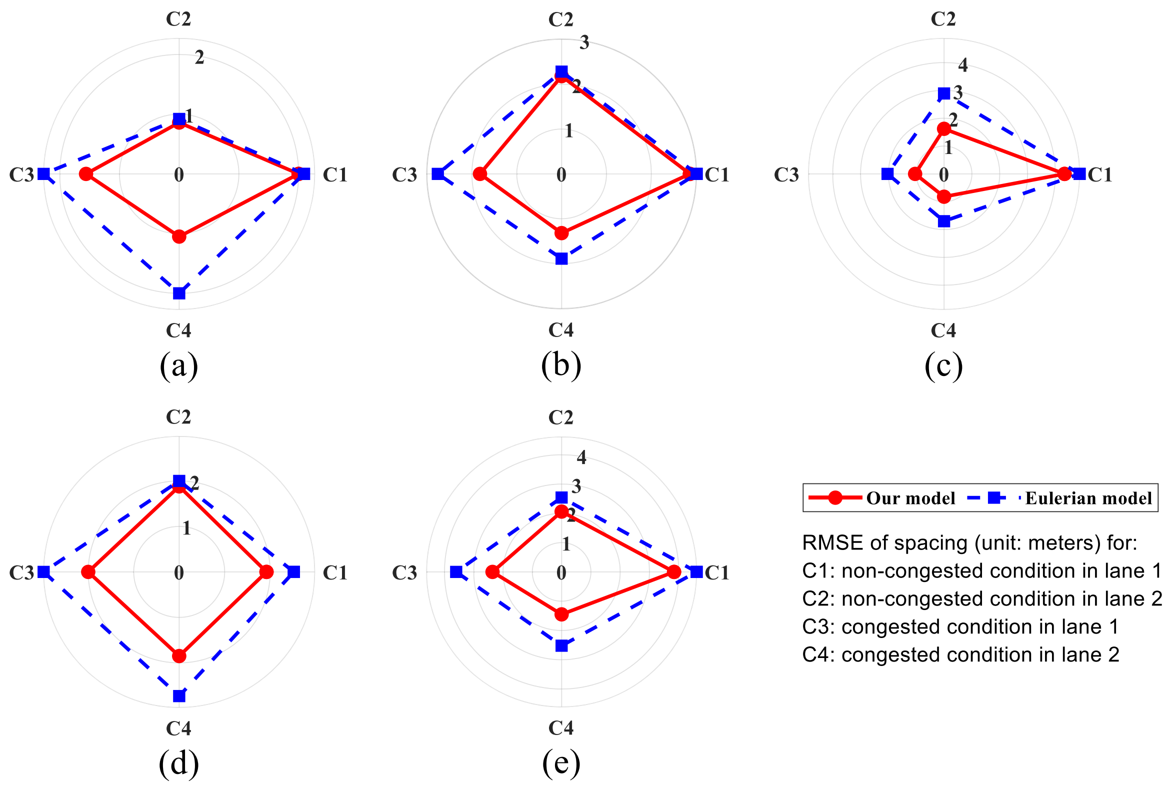


Fig. 21. Comparison of RMSE for spacing across different lanes and traffic conditions in the following datasets: (a) L002_F001 dataset; (b) L002_F002 dataset; (c) L002_F003 dataset; (d) L002_F004 dataset; (e) L002_F005 dataset.

Fig. 18 illustrates the cumulative net number of lane-changing vehicles for each lane over time. The rising line above zero indicates the cumulative net number of vehicles entering lane 2, while the descending line below zero indicates the cumulative net number of vehicles leaving lane 1. The blue circle and pentagram lines represent the cumulative net number of lane-changing vehicles in lane 1, as calculated from the ZTD datasets and the proposed model, respectively. Similarly, the red plus sign and triangle dotted lines represent the cumulative net number of lane-changing vehicles in lane 2, derived from the ZTD datasets and the proposed model, respectively.

The close alignment between the estimated cumulative number of lane-changing vehicles and the observed data from the ZTD dataset, as illustrated in Fig. 18, confirms the model's accuracy. A noticeable trend is that more vehicles tend to change lanes from lane 1 to lane 2 under both traffic conditions. The variation in the number of lane-changing vehicles is more significant under non-congested conditions, as shown in Fig. 18(a), compared to congested conditions, as seen in Fig. 18(b). This suggests that lane changes are more frequent in non-congested traffic conditions.

Figs. 16–18 illustrate the overall performance of the model across the entire 800-meter study area. For a more granular analysis, the data has been further disaggregated into 200-meter intervals. To quantitatively assess the model's performance, we use the Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) as performance metrics, defined as follows:

$$RMSE = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\left(observation^k - estimation^k\right)^2} \tag{59}$$

$$MAPE = \frac{100}{K}\sum_{k=1}^{K}\left|\frac{observation^k - estimation^k}{observation^k}\right| \tag{60}$$

where $observation^k$ represents the observed physical quantity at time step k, such as the number of vehicles or spacing within a specific interval (e.g., 0–200 m) on lane l at time step k, while $estimation^k$ denotes the corresponding estimated value at time step k.

Table 5 presents the estimation performance of the model for the number of vehicles in each lane, using RMSE and MAPE as evaluation metrics. The model demonstrates robust performance across the 200-meter intervals, with RMSE values ranging from 0.36 to 0.76 vehicles and MAPE values ranging from 4.82 % and 8.10 % under non-congested conditions. Under congested conditions, the RMSE values range from 0.41 to 0.74 vehicles and MAPE values range from 5.45 % to 7.25 %. Table 6 provides a summary of the RMSE and MAPE for spacing, indicating that RMSE values remain below 4.71 m and MAPE values are below 8.38 %. These results demonstrate that the proposed model offers good estimation accuracy for both vehicle counts and spacing across the 200-meter

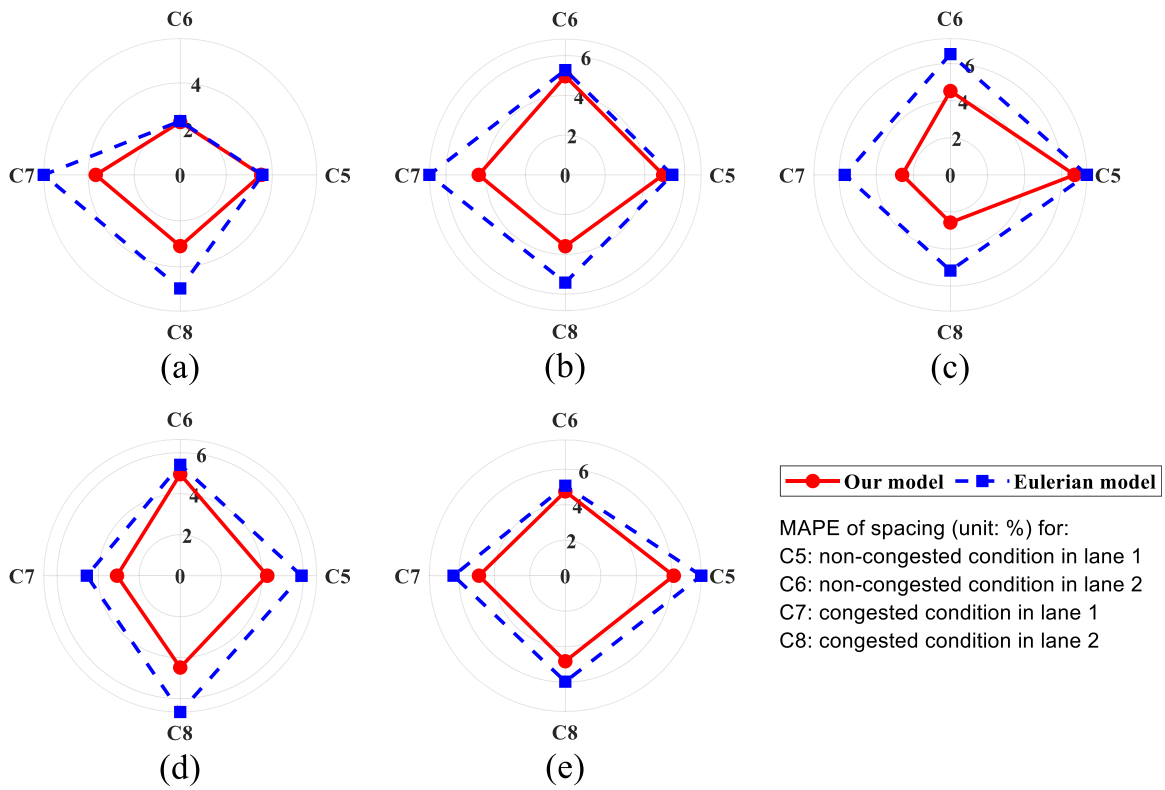


Fig. 22. Comparison of MAPE for spacing across different lanes and traffic conditions in the following datasets: (a) L002_F001 dataset; (b) L002_F002 dataset; (c) L002_F003 dataset; (d) L002_F004 dataset; (e) L002_F005 dataset.

intervals.

6.2.2. *Comparison with simulations of the Eulerian multilane model*

The Eulerian multilane LWR model, described in Eq. (1), builds on the traditional LWR model by incorporating multiple lanes and accounts for interactions between lane-specific traffic flows. In this study, we applied the Eulerian multilane model developed by Laval and Daganzo (2006), following their numerical methods. For a comprehensive description of these methods, please refer to Laval and Daganzo (2006). We evaluated the model's effectiveness in estimating traffic flow states and compared its performance with the model proposed in our research.

We calibrated the Eulerian model using the ZTD dataset as detailed in Section 6.1. To assess the performance of the calibrated Eulerian model, we conducted simulations across various testing datasets described in Section 6.1. The RMSE and MAPE for vehicle count and spacing were chosen as the metrics for comparison.

Figs. 19 and 20 present a quantitative comparison between our proposed model and the Eulerian model across different lanes and traffic conditions. Fig. 19 shows the RMSE for the number of vehicles, while Fig. 20 displays the MAPE for the number of vehicles. The results demonstrate that the Eulerian model generally has higher RMSE and MAPE values than our Lagrangian model, particularly in congested conditions, as shown in categories C3 and C4 in Fig. 19 and categories C7 and C8 in Fig. 20. Similar trends are observed in Figs. 21 and 22, which display the RMSE and MAPE for spacing, respectively. These findings suggest that the Lagrangian model performs better in predicting both the number of vehicles and spacing, especially in congested scenarios. This improved performance could be attributed to the Lagrangian model's capability to track individual vehicle groups, resulting in a more precise representation of traffic dynamics. This ability is especially beneficial in the complex conditions of congested traffic compared to non-congested traffic.

7. Discussion and conclusions

In this paper, we propose a Lagrangian formulation of a multilane traffic flow model, which differs from the traditional Eulerian coordinate-based models. We derive the conservation law for multilane traffic in Lagrangian coordinates and discretize the model to capture spacing dynamics. By discretizing the proposed model, we vividly illustrate how changes in the net lane-changing number (or lane-changing rate) in Lagrangian coordinates influence vehicle spacing. Additionally, a lane choice probability model is proposed to estimate the lane-changing number by considering the relative positions between vehicle groups in adjacent lanes. This estimation allows for the update of spacing in each lane at every time step. We also discuss the boundary conditions of both the continuous and discrete Lagrangian multilane models.

To validate the effectiveness of our model, we conducted a numerical simulation of a three-lane highway section with a lane-drop scenario. The results demonstrate that the proposed model accurately simulates the generation and propagation of congestion at the bottleneck, as well as the traffic perturbation caused by the lane-changing behaviors. Additionally, the proposed model effectively captures the lane-changing characteristics and spacing dynamics of lane-changing vehicle groups over space and time. To further validate the proposed model, we compared its outputs with data from an 800-meter, two-lane freeway section in Japan, known as the ZTD dataset. The comparison reveals that the proposed multilane Lagrangian model accurately reflects traffic dynamics and provides precise estimates of traffic volume and vehicle spacing. Furthermore, the analysis of the cumulative net number of lane-changing vehicles demonstrates the model's capability in capturing lane-changing dynamics. The model not only performs well across the entire 800-meter study area but also shows strong results at finer spatial resolutions, such as 200-meter intervals. Moreover, when compared to the traditional Eulerian multilane model, our Lagrangian model shows superior performance in predicting both the number of vehicles and spacing, especially under congested conditions. This improved performance is likely due to the Lagrangian model's capability to track individual vehicle groups, resulting in a more precise representation of traffic dynamics.

Our proposed model holds significant potential for implementation in highway management and control, particularly with the increasing availability of vehicle trajectory data. Integration with data assimilation techniques enables online lane-based traffic estimation and dynamic lane-based control. Moreover, our proposed model has good scalability from the traditional human-driven traffic to the mixed traffic flow environment, where human-driven vehicles and autonomous vehicles coexist. These are promising avenues for our future research.

CRediT authorship contribution statement

Liang Lu: Conceptualization, Data curation, Writing – original draft, Software, Formal analysis, Methodology. **Fangfang Zheng:** Conceptualization, Funding acquisition, Methodology, Supervision, Writing – original draft, Writing – review & editing. **Xiaobo Liu:** Funding acquisition, Validation, Writing – review & editing. **Henry X. Liu:** Validation, Writing – review & editing. **Yufei Yuan:** Validation, Writing – review & editing.

Declaration of competing interest

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

Acknowledgment

This work was supported by the National Natural Science Foundation of China under grants 52072315 and Science and Technology Program of Sichuan Province under grants 2025YFHZ0193.

References

Chanut, S., Buisson, C., 2003. Macroscopic model and its numerical solution for two-flow mixed traffic with different speeds and lengths. Transp. Res. Rec. 1852, 209–219. https://doi.org/10.3141/1852-26.

Coifman, B., 2003. Estimating density and lane inflow on a freeway segment. Transp. Res. A: Policy Pract. 37, 689–701. https://doi.org/10.1016/S0965-8564(03)00025-9.

Coifman, B., 2002. Estimating travel times and vehicle trajectories on freeways using dual loop detectors. Transp. Res. A: Policy Pract. 36, 351–364. https://doi.org/10.1016/S0965-8564(01)00007-6.

Courant, R., Friedrichs, K., Lewy, H., 1967. On the partial difference equations of mathematical physics. IBM J. Res. Dev. 11, 215–234. https://doi.org/10.1147/rd.112.0215.

Daganzo, C.F., 2002a. A behavioral theory of multi-lane traffic flow. I: Long homog. free. sect. 28.

Daganzo, C.F., 2002b. A behavioral theory of multi-lane traffic flow. II: Merges Congest. 11.

Daganzo, C.F., 1995. The cell transmission model, part II: network traffic. Transp. Res. B: Methodol. 29, 79–93.

Daganzo, C.F., 1994. The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory. Transp. Res. B: Methodol. 28, 269–287. https://doi.org/10.1016/0191-2615(94)90002-7.

Daganzo, C.F., Lin, W.-H., Del Castillo, J.M., 1997. A simple physical principle for the simulation of freeways with special lanes and priority vehicles. Transp. Res. B: Methodol. 31, 103–125. https://doi.org/10.1016/S0191-2615(96)00032-X.

Delle Monache, M.L., Liard, T., Piccoli, B., Stern, R., Work, D., 2019. Traffic reconstruction using autonomous vehicles. SIAM. J. Appl. Math. 79, 1748–1767. https://doi.org/10.1137/18M1217000.

Duret, A., Yuan, Y., 2017. Traffic state estimation based on eulerian and lagrangian observations in a mesoscopic modeling framework. Transp. Res. B: Methodol. 101, 51–71. https://doi.org/10.1016/j.trb.2017.02.008.

Edie, L.C., 1963. Discussion of traffic stream measurements and definitions. Port of New York Authority New York.

Han, Y., Wang, M., He, Z., Li, Z., Wang, H., Liu, P., 2021. A linear lagrangian model predictive controller of macro- and micro- variable speed limits to eliminate freeway jam waves. Transp. Res. C: Emerg. Technol. 128, 103121. https://doi.org/10.1016/j.trc.2021.103121.

Herrera, J.C., Work, D.B., Herring, R., Ban, X.J., Jacobson, Q., Bayen, A.M., 2010. Evaluation of traffic data obtained via GPS-enabled mobile phones: the Mobile Century field experiment. Transp. Res. C: Emerg. Technol. 18, 568–583. https://doi.org/10.1016/j.trc.2009.10.006.

Japan Hanshin Expressway Company, 2019. Zen traffic data (ZTD) vehicle trajectories and supporting data. from. https://zen-traffic-data.net/ (Accessed 6 September 2022).

Laval, J.A., Costeseque, G., Chilukuri, B., 2016. The impact of source terms in the variational representation of traffic flow. Transp. Res. B: Methodol. 94, 204–216. https://doi.org/10.1016/j.trb.2016.09.011.

Laval, J.A., Daganzo, C.F., 2006. Lane-changing in traffic streams. Transp. Res. B: Methodol. 40, 251–264. https://doi.org/10.1016/j.trb.2005.04.003.

Laval, J.A., Leclercq, L., 2013. The Hamilton–Jacobi partial differential equation and the three representations of traffic flow. Transp. Res. B: Methodol. 52, 17–30. https://doi.org/10.1016/j.trb.2013.02.008.

Leclercq, L., Laval, J., Chevallier, E., 2007. The Lagrangian coordinates and what it means for first order traffic flow models 18.

Levin, M.W., Boyles, S.D., 2016. A cell transmission model for dynamic lane reversal with autonomous vehicles. Transp. Res. C: Emerg. Technol. 68, 126–143. https://doi.org/10.1016/j.trc.2016.03.007.

Li, Y., Li, L., Ni, D., Zhang, Y., 2021. Comprehensive survival analysis of lane-changing duration. Measurement 182, 109707. https://doi.org/10.1016/j.measurement.2021.109707.

Lighthill, M.J., Whitham, G.B., 1955a. On kinematic waves I. Flood movement in long rivers. Proc. R. Soc. Lond. A 229, 281–316. https://doi.org/10.1098/rspa.1955.0088.

Lighthill, M.J., Whitham, G.B., 1955b. On kinematic waves II. A theory of traffic flow on long crowded roads. Proc. R. Soc. Lond. A 229, 317–345. https://doi.org/10.1098/rspa.1955.0089.

Ma, W., Qian, S., 2019. High-resolution traffic sensing with autonomous vehicles.

Mehran, B., Kuwahara, M., Naznin, F., 2012. Implementing kinematic wave theory to reconstruct vehicle trajectories from fixed and probe sensor data. Transp. Res. C: Emerg. Technol. 20, 144–163. https://doi.org/10.1016/j.trc.2011.05.006.

Molnár, T.G., Upadhyay, D., Hopka, M., Van Nieuwstadt, M., Orosz, G., 2021. Delayed lagrangian continuum models for on-board traffic prediction. Transp. Res. C: Emerg. Technol. 123, 102991. https://doi.org/10.1016/j.trc.2021.102991.

Montanino, M., Punzo, V., 2013. Making NGSIM data usable for studies on traffic flow theory: multistep method for vehicle trajectory reconstruction. Transp. Res. Rec. 2390, 99–111. https://doi.org/10.3141/2390-11.

Munoz, L., Sun, X., Horowitz, R., Alvarez, L., 2003. Traffic density estimation with the cell transmission model. In: Proceedings of the 2003 American Control Conference, 2003. Presented at the 2003 American Control Conference, IEEE, Denver, CO, USA, pp. 3750–3755. https://doi.org/10.1109/ACC.2003.1240418.

Newell, G.F., 1993. A simplified theory of kinematic waves in highway traffic, part I: general theory. Transp. Res. B: Methodol. 27, 281–287.

Newell, G.F., 1961. Nonlinear effects in the dynamics of car following. Oper. Res. 9, 209–229. https://doi.org/10.1287/opre.9.2.209.

Orosz, G., Ge, J.I., He, C.R., Avedisov, S.S., Qin, W.B., Zhang, L., 2017. Seeing beyond the line of site – Controlling connected automated vehicles. Mech. Eng. 139, S8–S12. https://doi.org/10.1115/1.2017-Dec-8.

Richards, P.I., 1956. Shock waves on the highway. Oper. Res. 4, 42–51. https://doi.org/10.1287/opre.4.1.42.

Toledo, T., Zohar, D., 2007. Modeling duration of lane changes. Transp. Res. Rec. 1999, 71–78. https://doi.org/10.3141/1999-08.

van Wageningen-Kessels, F., Hoogendoorn, S., van Lint, H., Vuik, K., 2009. Multiple user classes in the kinematic wave model in lagrangian coordinates. Present. Traff. Granul. Flow 7.

Van Wageningen-Kessels, F., Van Lint, H., Hoogendoorn, S.P., Vuik, K., 2010. Lagrangian formulation of multiclass kinematic wave model. Transp. Res. Rec. 2188, 29–36. https://doi.org/10.3141/2188-04.

van Wageningen-Kessels, F., Yuan, Y., Hoogendoorn, S.P., van Lint, H., Vuik, K., 2013. Discontinuities in the Lagrangian formulation of the kinematic wave model. Transp. Res. C: Emerg. Technol. 34, 148–161. https://doi.org/10.1016/j.trc.2011.08.004.

Wang, Q., Li, Z., Li, L., 2014. Investigation of discretionary lane-change characteristics using next-generation simulation data sets. J. Intell. Transp. Syst. 18, 246–253. https://doi.org/10.1080/15472450.2013.810994.

Wong, G.C.K., Wong, S.C., 2002. A multi-class traffic flow model – an extension of LWR model with heterogeneous drivers. Transp. Res. A: Policy Pract. 36, 827–841. https://doi.org/10.1016/S0965-8564(01)00042-8.

Work, D.B., Blandin, S., Tossavainen, O.P., Piccoli, B., Bayen, A.M., 2010. A traffic model for velocity data assimilation. Appl. Math. Res. eXpress abq002. https://doi.org/10.1093/amrx/abq002.

Work, D.B., Tossavainen, O.-P., Blandin, S., Bayen, A.M., Iwuchukwu, T., Tracton, K., 2008. An ensemble Kalman filtering approach to highway traffic estimation using GPS enabled mobile devices. In: 2008 47th IEEE Conference on Decision and Control. Presented at the 2008 47th IEEE Conference on Decision and Control, IEEE, Cancun, Mexico, pp. 5062–5068. https://doi.org/10.1109/CDC.2008.4739016.

Yang, M., Wang, X., Quddus, M., 2019. Examining lane change gap acceptance, duration and impact using naturalistic driving data. Transp. Res. C: Emerg. Technol. 104, 317–331. https://doi.org/10.1016/j.trc.2019.05.024.

Yuan, Y., van Lint, J.W.C., Wilson, R.E., van Wageningen-Kessels, F., Hoogendoorn, S.P., 2012. Real-time lagrangian traffic State estimator for freeways. IEEE Trans. Intell. Transport. Syst. 13, 59–70. https://doi.org/10.1109/TITS.2011.2178837.

Zheng, F., Jabari, S.E., Liu, H.X., Lin, D., 2018. Traffic state estimation using stochastic lagrangian dynamics. Transp. Res. B: Methodol. 115, 143–165. https://doi.org/10.1016/j.trb.2018.07.004.