https://doi.org/10.1016/j.trc.2016.05.002

Modelling the impacts of mandatory and discretionary lane-changing maneuvers

T. L. Pan[a], William H.K. Lam[a], A. Sumalee[a], R. X. Zhong[b,c,*]

[a]*Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, China.*
[b]*School of Engineering, Sun Yat-Sen University, Guangzhou, China.*
[c]*Guangdong Provincial Key Laboratory of Intelligent Transportation Systems, Guangzhou, China.*

Abstract

In this paper, a novel mesoscopic multilane model is proposed to enable simultaneous simulation of mandatory and discretionary lane-changing behaviors to realistically capture multilane traffic dynamics. The model considers lane specific fundamental diagrams to simulate dynamic heterogeneous lane flow distributions on expressways. Moreover, different priority levels are identified according to different lane-changing motivations and the corresponding levels of urgency. Then, an algorithm is proposed to estimate the dynamic mandatory and discretionary lane-changing demands. Finally, the lane flow propagation is defined by the reaction law of the demand-supply functions, which can be regarded as an extension of the Incremental-Transfer and/or Priority Incremental-Transfer principles. The proposed mesoscopic multilane cell transmission model is calibrated and validated on a complex weaving section of the State Route 241 freeway in Orange County, California, showing both the positive and negative impact of lane changing maneuvers, e.g., balancing effect and capacity drop, respectively. Moreover, the empirical study verifies that the model requires no additional data other than the cell transmission model does. Thus, the proposed model can be deployed as a simple simulation tool for accessing dynamic mesoscopic multilane traffic state from data available to most management centers, and also the potential application in predicting the impact of traffic incident or lane control strategy.

Keywords: Mesoscopic multilane traffic model, minimum gap acceptance criterion, lane specific fundamental diagram, mandatory lane changing demand estimation.

1. Introduction

The kinematic wave (LWR-KW) model (Lighthill and Whitham, 1955; Richards, 1956) and its discretized version the cell transmission model (CTM) (Daganzo, 1994; Szeto, 2008; Sumalee et al., 2011), which adopt the triangular fundamental diagram (Newell, 1993), is recognized as the simplest means to explain the evolution of traffic dynamics and features. However, these macroscopic models simply assume that traffic flow is uniformly distributed over lanes by integrating traffic streams travelling on different lanes into a single flow stream with uniform lateral distribution (Munjal and Pipes, 1971). This uniform assumption may not be appropriate in the sense that heterogeneous traffic flow distribution, such as lane specific flow, density, speed and vehicle type, can be easily observed on multilane expressways (Carter et al., 1999; Cassidy and Rudjanakanoknad, 2005; Gunay, 2007; Duret et al., 2012). Therefore, single lane

*Corresponding author. Tel:+86 3933-2772-503; Fax: +86 3933-2775
Email addresses: glorious9009@gmail.com (T. L. Pan), cehklam@polyu.edu.hk (William H.K. Lam), ceasumal@polyu.edu.hk (A. Sumalee), cezhong@gmail.com (R. X. Zhong)

traffic models may fail to capture more complex traffic features on multiple lane roadways, such as heterogeneous traffic flow distribution, capacity drop, moving bottlenecks, and so on. Moreover, recent research has revealed that vehicle lane-changing (LC) has significant impacts on traffic safety since accidents tend to happen in lane-changing areas such as weaving sections and interchanges (Golob et al., 2004; Cassidy and Rudjanakanoknad, 2005; Lee and Cassidy, 2009; Srivastava and Geroliminis, 2013).

In order to quantify complex features of multilane traffic, understanding lane-changing is essential (Mauch and Cassidy, 2002; Ahn and Cassidy, 2007). According to different decision making processes and their impacts on surrounding traffic, lane-changing maneuvers are usually classified as mandatory lane-changing (MLC) and discretionary lane-changing (DLC) (Yang and Koutsopoulos, 1996). A MLC is executed when a driver must change lane to follow a certain path to his/her destination. A DLC occurs when a driver seeks for better driving conditions to gain a speed (or travel time) advantage.

Modeling the lane-changing maneuvers is a key issue in the development of microscopic traffic simulation tools (Pipes, 1967; Toledo et al., 2005; Kesting et al., 2007; Sun and Elefteriadou, 2010). Under microscopic simulation framework, lane-changing is usually modeled in a sequence of three steps: 1) lane-changing necessity checking, 2) target lane choice, and 3) gap acceptance decision. The two most popular microscopic lane-changing algorithms are rule-based models and discrete choice-based (DCB) models (Toledo et al., 2005; Ben-Akiva et al., 2006; Choudhury et al., 2007; Kesting et al., 2007). Rule-based algorithms model lane-changings from the perspective of heterogenous drivers with different gap-acceptance conditions[1] and different lane-changing behavior for various situations. The DCB algorithms simulate driver behavior using Logit or Probit models. These microscopic lane-changing algorithms can describe lane-changing behaviors in detail, however, they usually contain a large number of parameters and cannot provide intuitive descriptions of system-level effects of lane-changing traffic. In addition, calibration and validation require high resolution vehicle trajectory data. Kesting et al. (2007) argued that the empirical investigations of lane-changing behavior for these algorithms are even more difficult than that for car-following behavior.

On the other hand, many studies have been carried out to understand various characteristics of lane-changing traffic at the macroscopic level by extending the celebrated kinematic wave (KW) theory. Laval and Leclercq (2010) proposed a mechanism to describe how stop-and-go oscillations may result from the reactions of heterogeneous drivers to minor speed variations on a hill. Ahn and Cassidy (2007) and Wang and Coifman (2008) demonstrated that lane-changing maneuvers contribute to traffic oscillations. In Laval and Daganzo (2006), a hybrid model of lane-changing traffic theory, also termed multilane hybrid (MH) theory, was proposed. The MH theory models each lane-changing maneuver as a moving bottleneck with bounded acceleration. It was shown that the bounded acceleration of a lane-changing into a faster-moving lane can explain traffic instabilities at lane-drop and moving bottlenecks, and also at merge bottlenecks (Laval et al., 2007). Cassidy and Rudjanakanoknad (2005) found that lane-changing maneuvers can cause capacity drop at a merge bottleneck. Coifman et al. (2006) found that lane-changing maneuvers cause additional delay in queues. Patire and Cassidy (2011) investigated the lane-changing behavior for an uphill expressway. Carey et al. (2015) extended the CTM to consider lane-changing maneuvers for dynamic traffic assignment purpose. A lane change model was proposed from a combination of the route, speed and keep-right incentives by influencing car-following behavior for relaxation and synchronization (Schakel et al., 2012). The model fits lane flow distributions well under free flow condition while the fit under congestion condition is unclear.

Lane-changing maneuvers, on the other hand, has positive effects on the performance of transportation

[1] Gap acceptance models are used to model the execution of lane changings. Before making a lane-changing, the driver compare the available gaps to her smallest acceptable gap (also known as critical gap). A lane-changing is executed if the adjacent gaps for subject vehicles are greater than the critical gaps, i.e. the gaps are acceptable.

systems. It is deemed that, e.g. Shvestsov and Helbing (1999); Laval and Daganzo (2006); Jin (2010a) and Patire and Cassidy (2011), the heterogeneity among different lanes is usually a reason motivating lane-changing. While lane-changing maneuvers, on the other hand, might have balancing effect, e.g. lane-changing could smooth out differences between lanes under certain situations. This balancing effect could be beneficial to the whole traffic system in achieving higher efficiency. For example, Cheu et al. (2009) found that lane-changing could reduce the overall system queuing delay by simulation. As deemed by Jin (2010a), that all these studies, however, do not provide a simple approach for analyzing the impacts of lane-changing maneuvers and the corresponding traffic dynamics at the aggregate level. Based on the observation that, when changing its lane, a vehicle affects traffic on both its current and target lanes, Jin (2010a) also proposed an extended kinematic wave model to capture such lateral interactions by introducing a new lane-changing intensity variable, i.e. lane-changing traffic causes effective additional density which is determined by drivers' lane-changing choices and characteristics in a road section during a time interval. With a modified fundamental diagram, it was claimed that the impacts of lane-changing traffic on overall traffic flow can be captured by the extended kinematic wave model. Jin (2010b) used a set of vehicle trajectory data collected by the NGSIM project to calibrate the lane-changing intensity. It was found that the lane-changing intensity is an exponential function of the traffic density. However, Jin's approach is mainly developed for freeway segments without on-/off- ramps and enforces several restrictive assumptions (Zheng, 2014). Tang et al. (2009) proposed a macroscopic model of lane-changing and showed its consistence with car-following behavior on a two-lane highway. The effect of lane-changing on the stable region and the propagation speeds of the first-order and second-order waves was investigated by using linear stability theory. Both the models proposed in Jin (2010a) and Tang et al. (2009) did not represent lane-specific behavior. Sheu and Ritchie (2001) and Sheu (2004) defined lane-changing fraction from the original lane to adjacent lanes, and return-lane-changing fraction from adjacent lanes to the original lane in the downstream using the corresponding upstream and downstream traffic counts detected directly from point detectors in state space formulation, wherein Kalman filtering approach was adopted to estimate all traffic variables including traffic volumes and lane-changing fractions.

Remarkably, existing approaches for simulating vehicle lane-changing maneuvers (both microscopic and macroscopic) typically differentiate DLC and MLC whilst developing different models for these two LC behaviors separately. For example, some models concentrate on DLC typically lack a MLC component (Laval and Daganzo, 2006) and vice versa (Hou et al., 2015). Hence, it is not easy to relate these models to each other or furthermore to give a comprehensive model for simulating LC maneuvers at one scale. As reviewed by Zheng (2014), there is a clear need to develop a comprehensive model that captures the impact of MLC and DLC maneuvers on surrounding traffic while maintaining the balance between maximizing the model's predictive and explanatory power and minimizing the model's complexity.

As previously discussed, vehicle lane-changing maneuvers often result in instability in traffic flow. Recent research has found that even more serious impact on surrounding traffic can be caused by the lane-changing maneuvers of heavy vehicles although they only account for a minority of traffic stream. Heavy vehicles produce a disproportionate effect for passenger vehicle drivers especially under heavy traffic conditions. Despite the increasing number of heavy vehicles on freeways and its pronounced effect on traffic flow, previous studies have focused on the lane-changing behavior of passenger car drivers wherein the differences between heavy vehicles and passenger cars is primarily accounted for through differences in vehicle length and acceleration/deceleration capabilities (Gipps, 1986; Ahmed, 1999; Hidas, 2005). From an empirical aspect, Knoop et al. (2012) investigated the relationship between the number of lane changes as function of the operational characteristics of the origin and target lane. The empirical findings by Duret et al. (2012) highlighted that the distribution of flow over different lanes (or lane flow distribution (LFD))

should be included for a multi-lane traffic model to take into account the lateral feature of traffic flow due to the heterogeneity occurs in diver behaviors, vehicle types and the corresponding vehicle-class specific control, e.g. driving ban for trucks (DBT). Keyvan-Ekbatani et al. (inpress) found distinct strategies regarding speed and lane choice. It was found that the desired speed changes with the lane and with traffic conditions while the relation between speed and lanes is currently not in most models. In light of these and parallel to conventional macroscopic traffic flow models which adopt an uniform fundamental diagram to represent the longitudinal dynamics of the traffic flow, the speed-density relationship should be in a lane specific manner when developing multi-lane traffic model.

This paper develops a comprehensive mesoscopic multilane traffic model that simulates DLC and MLC maneuvers simultaneously. The rest of this paper is organized as follows.

1. A model for propagating mesoscopic multilane dynamics is developed in Section 2, which enables:
   - Determining the minimum gap acceptance criteria for different lane-changing intentions;
   - Evaluating whether or not the traffic flows sent from different directions can be accepted by a downstream section via an extended Incremental-Transfer (IT) and/or Priority IT (PIT) principle that defines demand-supply reaction laws;
   - Propagating traffic states on multi-lanes in both temporal and spatial dimensions.
2. A dynamical lane-changing demand estimation algorithm is developed in Section 3, including:
   - An estimation of the longitudinal distribution of MLC demand and DLC demand;
   - A process to dynamically refine the MLC and DLC based on the execution of lane changing flows.
3. An empirical study on a complex weaving section of SR-241 freeway in Orange County, California is conducted in Section 4 by using the traffic data provided by the Caltrans Performance Measurement System (PeMS)[2]. Conclusions are drawn and future works are discussed in Section 5.

2. Development of the mesoscopic multilane cell transmission model

Modelling lane-changing is a key procedure when developing multilane traffic model since heterogeneous traffic flow on multilane freeways is closely related to lane-changing maneuvers. On the other hand, both lane-changing decision making progress and lane-changing impact are continuously affected by the heterogeneous traffic flow on multilane freeway, as well as exogenous elements such as freeway geometry and traffic control schemes. This section presents a models for simulating mesoscopic traffic state in temporal, longitudinal and lateral dimensions. The model consists of a traffic flow component governing the lane specific traffic flow propagation and a lane-changing acceptance mechanism based on gap assessment rules. The traffic flow component extends the sending and receiving functions of the CTM based on lane specific fundamental diagrams and lane changing demand. In addition, levels of urgency determining the lane-changing priority is proposed.

[2] The Freeway Performance Measurement System (PeMS: http://pems.dot.ca.gov/) is conducted by the University of California, at Berkeley, with the cooperation of the California Department of Transportation (Caltrans), California Partners for Advanced Transit and Highways, and Berkeley Transportation Systems (Chen, 2003).

2.1. Lane-specific traffic flow characteristics

Parallel to the conventional macroscopic traffic flow models which adopt an uniform **macroscopic fundamental diagram** (MFD) to propagate the longitudinal dynamics of traffic flow, a properly defined lane specific fundamental diagram would be a better choice for simulation of multilane traffic. In this paper, the triangular lane specific flow-density equation is adopted:

$$q_m(t) = \begin{cases} v_{f,m} \cdot \rho_m(t), & \text{if } \rho_m(t) \leq \rho_{c,m} \\ w_{c,m} \cdot (\rho_{J,m} - \rho_m(t)), & \text{if } \rho_m(t) > \rho_{c,m} \end{cases}, \tag{1}$$

in conjunction with a flow-speed-density relationship:

$$v_m(t) = q_m(t)/\rho_m(t) \tag{2}$$

where $\rho_m(t)$ $(PCE/mile/lane)$, $v_m(t)$ $(mile/hour)$ and $q_m(t)$ $(PCE/hour/lane)$ denote the traffic density, speed and flow of lane m at time t, respectively. While $v_{f,m}$ $(mile/hour)$, $w_{c,m}$ $(mile/hour)$, $\rho_{c,m}$ $(PCE/mile)$ and $\rho_{J,m}$ (PCE/mile) denote free flow speed, wave-back speed of congestion, critical density and jam density of the fundamental diagram of lane m, respectively. The FD is built to map the average flow (or speed) to the total effective density, which is a weighted summation over all vehicle type-specific densities with respect to the passenger car equivalent (PCE) values.

To elaborate the necessity of lane specific FDs, historical traffic data is extracted from ten detector stations embedded in a 2.1-mile section on the State Route 241 in Organge County, California from the PeMS database. The data collected by six detector stations which are beyond 1 mile away from the off-ramp intersection is used to calibrate the FDs of non-weaving section, while the data by the detectors within 1 mile from the intersection is adopted to calibrate the FDs of the weaving section. Details of the test site will be outlined in the empirical study. Based on the lane specific vehicle count and occupancy observed on March 20, 2014 and June 2, 2015, we calibrate a set of FDs by the least squares in line with the literature.

It is clear that the FDs are very different from one lane to another in terms of free-flow speed, capacity etc on non-weaving and weaving section respectively. This is because of the heterogeneity occurs in diver behaviors, vehicle types and the corresponding vehicle-class specific control, e.g. driving ban for trucks[3]. This heterogeneity indicates that a properly defined lane specific FD is essential to consider the effect of multi-class traffic and vehicle-class specific control, whilst the lane specific FD would be an effective calibration and validation tool for a multi-lane traffic model. Since vehicle classes and the corresponding vehicle-class specific control on different lanes significantly affect the FD, it is not surprising that the lane specific FD will be also beneficial to capture the impact of lane based traffic control strategies on traffic patterns (Duret et al., 2012). Due to the interruptions of MLC maneuvers, capacity reduction could obviously be observed from the FDs calibrated on weaving sections (Figure 1b) compared with the FDs of the corresponding lanes on non-weaving section (Figure 1a) (e.g. Cassidy and Bertini (1999)). For example, the capacity of central lane reduces more than 300 PCE/hour in Figure 1b, wherein the FDs are zoomed to highlight the lane heterogeneity and capacity reduction.

2.2. Determination of minimum gap acceptance criterion

The execution of each individual lane-changing maneuver is a trade-off by taking into account necessity, desirability and possibility of all potential lane-changing maneuvers (Zheng et al., 2013; Gipps, 1986), see

[3] For example, General Laws of Massachusetts-Chapter89-Section 4C: On any highway with more than one passing lane in the same direction, heavy commercial vehicles, except buses, shall be restricted in ordinary operation to the right-hand travel lane, and in overtaking and passing shall be restricted to the next adjacent passing or travel lane, and shall not use any other lanes except in an emergency.

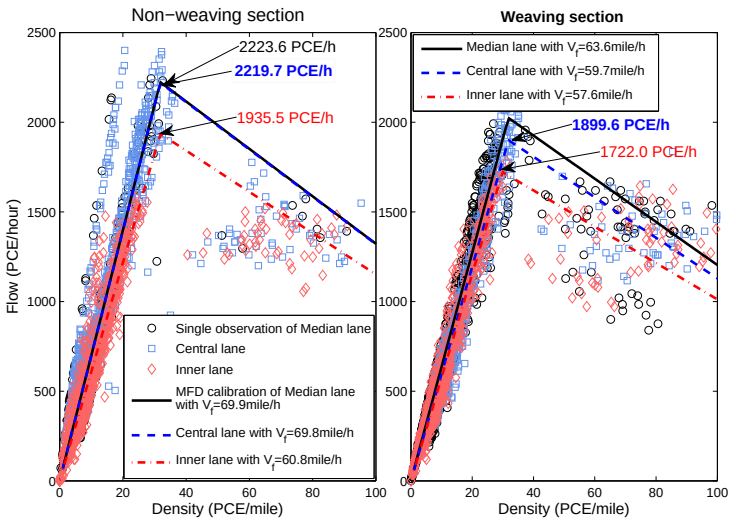

Figure 1: Lane specific FDs.

e.g. Figure 2. Necessity reflects the level of lane-changing urgency, which is related to the distance between the subject vehicle and its target turning point. Desirability is induced by the speed advantage possessed by adjacent vehicles compared with the current speed of subject vehicle. Finally, both the necessity of MLC and desirability of DLC confront the same issue, i.e. whether the prospective space gaps provided by target lane are long enough to ensure a safe lane-changing maneuver.

In Yang and Koutsopoulos (1996), a minimum gap acceptance criterion is first introduced as a benchmark to assess whether the adjacent space gaps on target lane are large enough to accommodate a subject vehicle that intends to change lane. The benchmark involves both the lead gap and lag gap as shown in Figure 2. The lead gap denotes the space between the rear of leading vehicle and the front of subject vehicle while the lag gap denotes the space between the front of following vehicle and the rear of subject vehicle. In the literature, the minimum lead (or lag) gap acceptance criterion is determined by several factors, such as the lane-changing type, the travelling speeds of the lane-changing vehicles, the level of lane-changing urgency, the number of lanes to be crossed, the length and mechanical characteristics of subject vehicles, drivers' attitudes such as patience or aggression, *etc.* (Gipps, 1986; Ahmed, 1999; Hidas, 2005). In this paper, in line with the mesoscopic nature of the proposed traffic flow model and the within cell homogenous assumption of the CTM, the lead gap, lag gap and subject vehicle length are aggregated as a minimum acceptance criterion for lane-changing acceptance assessment, as depicted in Figure 2.

The subject vehicle travelling on current lane m makes its final lane-changing decision towards the target lane β by comparing the minimum gap acceptance criterion $\tilde{g}_m^\beta(t)$ (*feet*) and the actual space gap $\bar{G}_\beta(t)$ (*feet*) provided by the target lane. For the subject vehicle with DLC intention, the minimum gap acceptance criterion $\tilde{g}_m^\beta(t)$ is determined by the minimal safe gap and the speed difference $\Delta v_m^\beta(t)$ (*mile/hour*) between the current lane speed $v_m(t)$ and the target lane speed $v_\beta(t)$[4].

Apart from gap acceptance criterion and lane speed difference, the subject vehicle with MLC intention is also affected by the remaining distance. The remaining distance $x(t)$ (mile) refers to the distance from

[4]In practice, the speeds of leading vehicle and following vehicle are different, and such difference might be either slight or significant, depending on the space gap length between them and the traffic condition. This simplification is due to the within cell homogenous assumption of the CTM.

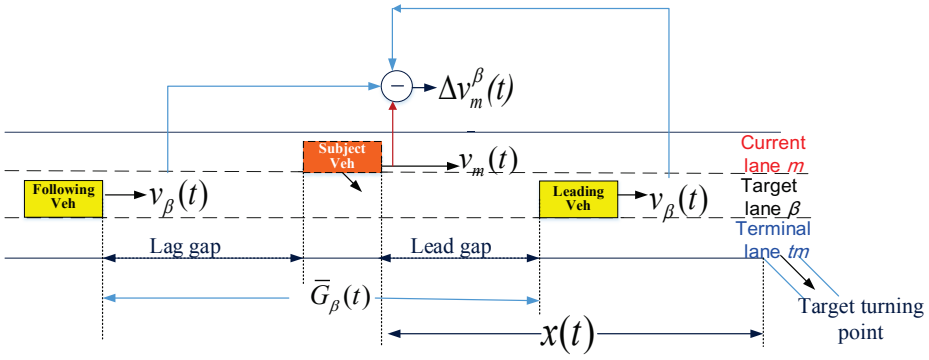

Figure 2: Lead gap, lag gap and adjacent space gap.

the current position of the subject vehicle to its target off-ramp intersection or downstream traffic incident spot. This distance is directly related to the driver's assessment of level of urgency associated with her/his MLC intention (Gipps, 1986). Based on Yang and Koutsopoulos (1996), the level of urgency is considered to follow three sequential stages as the vehicle approaches its target turning point. These three stages are separated by two critical positions x_r and x_c. Take a vehicle intending to execute a MLC as example, the target turning point is considered to be **remote** as long as the remaining distance $x(t) > x_r$, and **close** if $x(t) < x_c$, where x_r and x_c are distances that define the range within which the minimum gap acceptance criterion $\tilde{g}_m^\beta(t)$ linearly varies from the upper bound to the lower one based on current speed difference $\Delta v_m^\beta(t)$:

$$\tilde{g}_m^\beta(t)=\begin{cases} c_l \cdot \lfloor (v_m(t)-v_\beta(t)) \rfloor + c_f \cdot \lfloor (v_\beta(t)-v_m(t)) \rfloor + \bar{g}_{\min}, & \text{if } x(t) > x_r \\ \{c_l \cdot \lfloor (v_m(t)-v_\beta(t)) \rfloor + c_f \cdot \lfloor (v_\beta(t)-v_m(t)) \rfloor\} \cdot \frac{x(t)-x_c}{x_r-x_c} + \bar{g}_{\min}, & \text{if } x_c \le x(t) \le x_r \\ \bar{g}_{\min}, & \text{if } x(t) < x_c \end{cases} \quad (3)$$

where the operational symbol $\lfloor z \rfloor$ is defined as below:

$$\lfloor z \rfloor = \begin{cases} z, & \text{if } z > 0 \\ 0, & \text{if } z \le 0 \end{cases}.$$

The constant $\bar{g}_{\min}$ (feet) is the minimal safe gap to be provided by the target lane for the subject vehicle. c_l and c_f (feethour/mile) are constant presenting the relationship between speed difference with extra leading gap and extra lag gap, when the subject vehicle is still not too aggressive to execute a lane change.

As shown in Equation (3), in the remote stage, unless the speed difference between related lanes under consideration is negligible, i.e. $\Delta v_m^\beta(t)=v_\beta(t)-v_m(t) \approx 0$, the driver usually prefers a relatively larger gap at the commencement of lane-changing maneuver. Such a large gap is due to a risk-adverse attitude under non-urgent situation. In the second stage, the minimum gap acceptance criterion decreases linearly with respect to the reduction of remaining distance from x_r to x_c. Finally, after passing the critical distance point x_c, i.e. the vehicle gets so close to the target turning point that the MLC has to be executed, the minimum gap acceptance criterion achieves the lowest value which can just assure the subject vehicle can be safely accommodated by its target lane.

The decrease in the minimum gap acceptance criterion in the last two stages is based on the driver's expectation that the following vehicle in target lane will passively (or voluntarily) slow down to increase the gap before or after the subject vehicle (with MLC intention) enters (Hidas, 2005; Yang and Koutsopoulos,

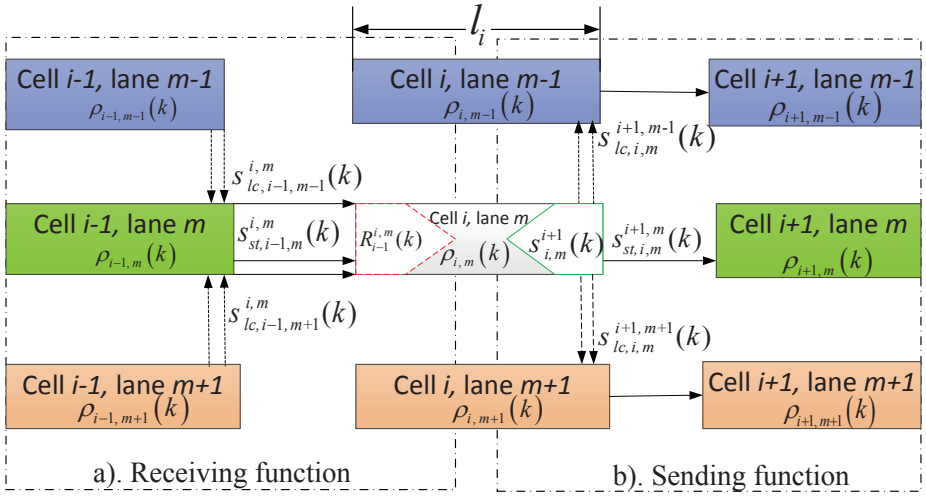

Figure 3: Merging and diverging induced by lane-changing maneuvers.

1996). However, for vehicles with DLC intention, the minimum gap acceptance criterion does not decrease as MLC does, because DLC is not related with level of urgency. In conclusion, a trade-off between minimum gap acceptance criterion and adjacent gap length is the prerequisite in determining the acceptance/rejection result for each lane-changing intention, no matter how the minimum gap acceptance criterion varies with the level of lane-changing urgency or speed difference. On average it would take 2-4 seconds (which will be adopted in the empirical study) for a driver to execute a lane change when the origin lane is stopped and the target lane is freely flowing, as indicated by empirical studies in Laval and Daganzo (2006); Laval and Leclercq (2008); Jin (2010b).

2.3. *Sending function and lane-changing fraction*

The continuum kinematic wave theory (KW) was first applied to model the multilane dynamics by Munjal and Pipes (1971), and was extended to a discrete version known as hybrid CTM by Laval and Daganzo (2006). These KW multilane models are capable of modelling the macroscopic dynamics with balancing the flow intended to merge and the available space provided by the target cell using incremental-transfer (IT) theory. However, there are some unmodeled elements would potentially hinder these models representing real traffic:

1. The heterogeneity in the lane specific fundamental diagrams is not considered.
2. The IT theory does not consider the minimum gap acceptance criterion.
3. Unacceptable small gaps scattered among vehicles might be aggregated and considered as an acceptable space by using conventional macroscopic multilane traffic flow models based on IT principle.
4. No account is taken to model different levels of priority associated with lane-changing intentions and straightforward flows when all of these streams are expecting to merge into the same lane within a prescribed short section.

The foregoing two sections intend to address the first three issues while this and the forthcoming sections tries to handle the last one.

To begin with, consider a multilane freeway section shown in Figure 3 which is divided into several cell packages, such as i-1, i and i+1, along the longitudinal dimension. Each cell package includes multiple cells associated with lanes ranging from the shoulder to median lane, such as m-1, m and m+1. Based on

the lane specific fundamental diagrams, the sending function $s_{i,\,m}^{i+1}(k)$ (*PCE/hour*) intending to leave cell i on lane m, which will be denoted as *cell* ($i,\ m$) hereafter, during $[kT_s, (k+1)T_s)$ can be calculated as:

$$s_{i,\,m}^{i+1}(k) = \begin{cases} v_{f,i,m} \cdot \rho_{i,m}(k), & \text{if } \rho_{i,m}(k) < \rho_{c,i,m} \\ Q_{i,m}, & \text{if } \rho_{i,m}(k) \geq \rho_{c,i,m} \end{cases} \tag{4}$$

where $\rho_{i,m}(k)$ (*PCE/mile*) denotes the dynamic traffic density of *cell* ($i,\ m$) estimated on time step k, and T_s denotes the duration of a simulation time step. $s_{i,m}^{i+1}(k)$ would move towards all possible directions associated with *cell package* (i+1), i.e., *cell* (i+1, m-1), *cell* (i+1, m), and *cell* (i+1, m+1). This algorithm of multilane sending function extends the sending function algorithm developed for single lane section in Daganzo (1994). In this paper, the subscripts of variables denote the lane change type, source cell and lane, while the superscripts denote the corresponding target cell and lane, respectively. According to the types of lane-changing intentions and the target cells, the total sending function $s_{i,m}^{i+1}(k)$ can be evaluated as:

$$s_{i,m}^{i+1}(k) = \sum_{lc=1,2;\,\beta=m-1,m+1} \sum s_{lc,\,i,\,m}^{i+1,\,\beta}(k) + s_{st,\,i,\,m}^{i+1,\,m}(k), \tag{5}$$

where lc=1 and 2 denote MLC and DLC, respectively, whilst $\beta = m \pm 1$ refer to the two adjacent lanes as illustrated in Figure 3b. $s_{lc,\,i,\,m}^{i+1,\,\beta}(k)$ denotes the lane changing demand be of type lc that intends to leave *cell* ($i,\ m$) and be sent towards *cell* (i+1, β) during time interval $[kT_s, (k+1)T_s)$, while $s_{st,i,m}^{i+1,m}(k)$ denotes the flow that intends to leave *cell* ($i,\ m$) to *cell* (i+1, m) . However, these sending functions might not be totally or even partially received considering the gap and the available space provided by target cells. The corresponding flow receiving algorithm will be devised in the following section.

2.4. Allocation of spaces for merging flows and dynamics propagation

In view of the disadvantages of the IT principle as previously discussed, an acceptance/rejection mechanism of lane-changing demand is introduced in this sub-section. Considering *cell* ($i,\ m$) as the target cell as illustrated in Figure 3a, the average space gap (*feet*) between successive vehicles in *cell* (i,m) at time step k is estimated based on the *within cell homogenous assumption* of the CTM:

$$\bar{G}_{i,m}(k) = \frac{5280 \cdot l_i - \rho_{i,m}(k) \cdot l_i \cdot L_p}{\rho_{i,m}(k) \cdot l_i}, \tag{6}$$

where L_p (*feet*) represents the vehicle length of Passenger Car Equivalent (*PCE*), and l_i (*mile*) denotes the longitudinal cell length of *cell package* (i). This equation quantifies the average gap length between successive vehicles traveling on *cell* (i, m) by excluding the space occupied by vehicles in current cell (1 *mile*=5280 *feet*). Based on this gap and the gap acceptance rule, the sending function $s_{lc,i-1,\alpha}^{i,m}(k)$ might be received totally or partially by *cell* ($i,\ m$) if $\bar{G}_{i,m}(k) \geq \tilde{g}_{lc,i-1,\alpha}^{i,m}(k)$, where $\tilde{g}_{lc,i-1,\alpha}^{i,\,m}(k)$ is the minimum gap acceptance criterion required by this stream of sending function $s_{lc,i-1,\alpha}^{i,m}(k)$, and $\tilde{g}_{lc,i-1,\alpha}^{i,m}(k)$ can be estimated by Equation (3). Otherwise, the sending flow will keep the current lane. $\alpha = m \pm 1$ refers to the two adjacent lanes from which the subject vehicles intend to travel towards *cell* ($i,\ m$) as illustrated in Figure 3a. Since a lane changing vehicle generally requires a larger space than a straightforward vehicle which is not intended to change lane, to quantify the space (in terms of *PCE*) required by each stream of the lane changing demand, the minimum gap acceptance criterion $\tilde{g}_{lc,i-1,\alpha}^{i,m}(k)$is further normalized to be the relative minimum gap acceptance criterion factor using Equation (7) as follows:

$$\tilde{\phi}_{lc,\,i-1,\,\alpha}^{i,\,m}(k) = \frac{\tilde{g}_{lc,\,i-1,\,\alpha}^{i,\,m}(k)}{L_p} \tag{7}$$

As it can be inferred from Equation (3), $\tilde{\phi}^{i,\,m}_{lc,\,i-1,\,\alpha}(k) \geq \frac{\bar{g}_{\min}}{L_p} > \frac{L_p}{L_p} = 1$. In the conventional CTM, the sending functions might not be totally received by the downstream section due to the limited available space (Daganzo, 1994; Muñoz et al., 2003). Extending this to the multilane model, whether the sending functions to be sent from different directions, such as *cell* (*i*-1, *m*+1), *cell* (*i*-1, *m*), and *cell* (*i*-1, *m*-1), can be received by the target *cell* (*i, m*) also depends on its available space defined by the receiving function, on the condition that the minimum gap acceptance criterion is fulfilled. Considering the lane-specific fundamental diagram depicted in Figure 1, the receiving function of the target *cell* (*i, m*), $R^{i,\,m}_{i-1}(k)$ (*PCE/hour*) during time interval $[kT_s, (k+1)T_s)$ is given as follow:

$$R^{i,\,m}_{i-1}(k) = \begin{cases} Q_{i,m}, & \text{if } \rho_{i,m}(k) < \rho_{c,i,m} \\ w_{c,i,m} \cdot \left(\rho_{J,i,m} - \rho_{i,m}(k)\right), & \text{if } \rho_{i,m}(k) \geq \rho_{c,i,m} \end{cases} \tag{8}$$

The receiving function $R^{i,\,m}_{i-1}(k)$ quantify the available space provided by *cell* (*i, m*), because this cell is proposed to allocate both the straightforward flow $s^{i,m}_{st,i-1,m}(k)$ and the lane-changing flows $s^{i,m}_{lc,i-1,\alpha}(k)$ (with α=$m \pm 1$, and lc=1, 2) to be sent from upstream cells on adjacent lanes. The flow propagation rule can be defined by extending the IT principle considering the gap acceptance criterion. To be specific:

$$q^{i,\,m}_{lc,\,i-1,\,\alpha}(k) = \begin{cases} s^{i,\,m}_{lc,\,i-1,\,\alpha}(k), & \text{if } U^{i,\,m}_{i-1}(k) \leq R^{i,\,m}_{i-1}(k) \;\&\; \tilde{g}^{i,\,m}_{lc,\,i-1,\,\alpha}(k) \leq \bar{G}_{i,\,m}(k) \\ \frac{s^{i,\,m}_{lc,\,i-1,\,\alpha}(k)}{U^{i,\,m}_{i-1}(k)} R^{i,\,m}_{i-1}(k), & \text{if } U^{i,\,m}_{i-1}(k) > R^{i,\,m}_{i-1}(k) \;\&\; \tilde{g}^{i,\,m}_{lc,\,i-1,\,\alpha}(k) \leq \bar{G}_{i,\,m}(k) \\ 0, & \text{if } \tilde{g}^{i,\,m}_{lc,\,i-1,\,\alpha}(k) > \bar{G}_{\,i,\,m}(k) \end{cases} \tag{9}$$

$$q^{i,\,m}_{st,\,i-1,\,m}(k) = \begin{cases} s^{i,\,m}_{st,\,i-1,\,m}(k), & \text{if } U^{i,\,m}_{i-1}(k) \leq R^{i,\,m}_{i-1}(k) \\ \frac{s^{i,\,m}_{st,\,i-1,\,m}(k)}{U^{i,\,m}_{i-1}(k)} R^{i,\,m}_{i-1}(k), & \text{if } U^{i,\,m}_{i-1}(k) > R^{i,\,m}_{i-1}(k) \end{cases}$$

where $q^{i,m}_{lc,i-1,\alpha}(k)$ and $q^{i,m}_{st,i-1,m}(k)$ denote the flows actually received by *cell* (*i,m*) from the lane changing demands and straightforward demand, respectively. Note that in Equation (9), the sending flows need to compete for the downstream supply of *cell* (*i, m*) when the available space of the target cell is insufficient to accommodate all the qualified the sending functions.

Regarding the minimum gap acceptance criterion required by different movements as the "priority levels" in the conventional IT principle, the competitive strength of $s^{i,m}_{lc,i-1,\alpha}(k)$ for the supply of the target *cell* (*i, m*) can be evaluated by the first term of Equation (9), wherein $U^{i,m}_{i-1}(k)$ is thus defined:

$$U^{i,m}_{i-1}(k) = s^{i,m}_{st,i-1,m}(k) + \sum_{lc=1,2} \sum_{\alpha=m\pm1} s^{i,m}_{lc,i-1,\alpha}(k) \cdot \tilde{\phi}^{i,m}_{lc,i-1,\alpha}(k), \forall s^{i,m}_{lc,i-1,\alpha}(k) \text{ with } \tilde{g}^{i,m}_{lc,i-1,\alpha}(k) \leq \bar{G}_{i,m}(k) \tag{10}$$

Note that the straightforward flow is not scaled by any factor, this is because a lane changing vehicle generally requires a space larger than the straightforward one. Finally, the density $\rho_{i,m}(k+1)$ (*PCE/mile*) of *cell* (*i, m*) is evaluated according to the flow conservation equation of the hybrid CTM as follow:

$$\rho_{i,m}(k+1) = \rho_{i,m}(k) + \frac{T_s}{l_i}\left(q^{i,m}_{st,i-1,m}(k) + \sum_{lc=1,2} \sum_{\alpha=m\pm1} q^{i,m}_{lc,i-1,\alpha}(k) - \left(q^{i+1,m}_{st,i,m}(k) + \sum_{lc=1,2} \sum_{\beta=m\pm1} q^{i,\,\beta}_{lc,i,m}(k)\right)\right), \tag{11}$$

and the evaluation is operated over all cell package on all lanes on the test section.

To sum up, a flow chart of the proposed dynamic mesoscopic multilane CTM is presented in Figure 4. The proposed model enables the traffic state propagation in temporal (the iteration of k), longitude (the iteration of i), and lateral (the iteration of m) dimensions. Firstly, the sending function of a cell is

determined according to the lane-specific fundamental diagrams. For lane-changing flows, the minimum gap acceptance criterion associated with levels of lane-changing urgency are also defined. Next, considering a specific lane changing intention, a gap acceptance/rejection mechanism is developed for assessing the relationship between the minimum gap acceptance criterion and the average space gap provided by the target cell to determine whether this lane changing intention could be executed, as depicted in the purple box in Figure 4. To quantify the space required by a lane changing vehicle in terms of PCE, the relative minimum gap acceptance criterion factor is introduced. In line with the conventional IT principle, this factor can be regarded as "adverse-priority" allocated to the related lane changing demand when competing for the downstream supply. This is consistent with a general traffic rule that lane changing vehicle should not harm the benefit of the following vehicle on the target lane. Finally, a flow conservation equation is used to calculate traffic density.

3. The lane-changing demand estimation algorithm

Note from the model development in Section 2 that the proposed model presumes each lane-changing demand, say $s^{i,m}_{lc,i-1,\alpha}(k)$ or $s^{i,\beta}_{lc,i-1,m}(k)$ as mentioned in Section 2, is given regardless of its moving direction and lane changing category. This is possible if the trajectory of each individual vehicle can be measured. However, measuring the trajectory of each individual vehicle is barely accessible for large scale applications in reality. On the other hand, even the benchmark vehicle trajectory database provided by the NGSIM project is subject to kinds of errors as previously mentioned. Acknowledging the lane changing demand cannot be directly measured in general, it is necessary to establish an algorithm to simultaneously estimate the time-dependent MLC and DLC lane changing demands based on the aggregated dynamic traffic state and historical data. Furthermore, dynamic lane changing demand estimation is also important for the adjustment of lane changing flows considering the space gap acceptance/rejection mechanism.

3.1. Estimation of the initial value of the total MLC demand

The MLC demand depends on the topology of the weaving section as shown in Figure 5. Bottlenecks due to lane drop is a typical scenario investigated in the literature as shown in Figure 5a. Under this scenario, the traffic flow travelling on lane m must cross over lane β before arriving at the bottleneck. In this section, lane tm denotes the terminal lane of a specific MLC maneuvers, while lane β denotes the lane between lane m and lane tm, and adjacent to lane m. The total MLC demand $\tilde{S}^{tm}_{1,m}(t)$ to be sent from lane m to lane tm, and also $\tilde{S}^{tm}_{1,\beta}(t)$ to be sent from lane β to lane tm, can be directly deduced from the dynamic traffic flow on lane m and lane β.

For the off-ramp intersection scenario shown in Figure 5b, the initial value of total MLC demand associated with related lanes can be obtained by solving a system of linear equations given by (12) based on the flow conservation. The solution of the variables $\tilde{S}^{tm\,(0)}_{1,m}(t)$ and $\tilde{S}^{tm\,(0)}_{1,\beta}(t)$ can be identified as the initial values of total MLC demand sent by lane m and lane β, while $\tilde{S}^{f\,(0)}_{tm}(t)$ denotes the estimated sending flow towards the off-ramp. Let iteration number r=0, the lane specific flow conservation is presented as below:

$$\begin{aligned}
\tilde{S}^{tm\,(r)}_{1,m}(t) + e_m(t + t^0_{x_n}) &= u_m(t) \\
\tilde{S}^{tm\,(r)}_{1,\beta}(t) + \tilde{S}^{tm\,(r)}_{1,m}(t) + e_\beta(t + t^0_{x_n}) &= u_\beta(t) + \tilde{S}^{tm\,(r)}_{1,m}(t) \\
\tilde{S}^{f\,(r)}_{tm}(t) + e_{tm}(t + t^0_{x_n}) &= u_\beta(t) + \tilde{S}^{tm\,(r)}_{1,m}(t) + \tilde{S}^{tm\,(r)}_{1,\beta}(t)
\end{aligned} \tag{12}$$

Note that the values of $u_i(t)$ (with i=m, β, or tm), $e_j(t + t^0_{x_n})$ (with j=m, β, or tm) can be regarded as the measured traffic flow rate if there happen to have detectors or the simulated flows by other simulation

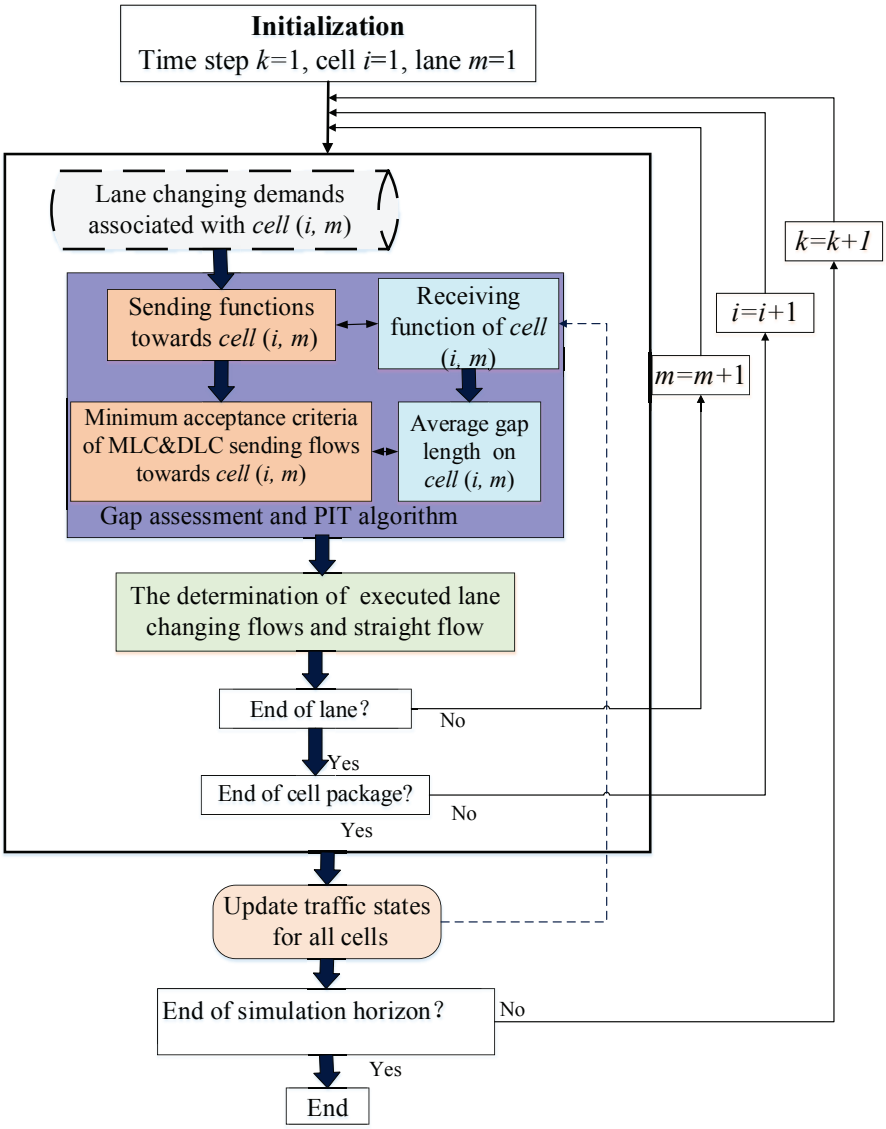

Figure 4: Flow chart of dynamic multilane cell transmission model.

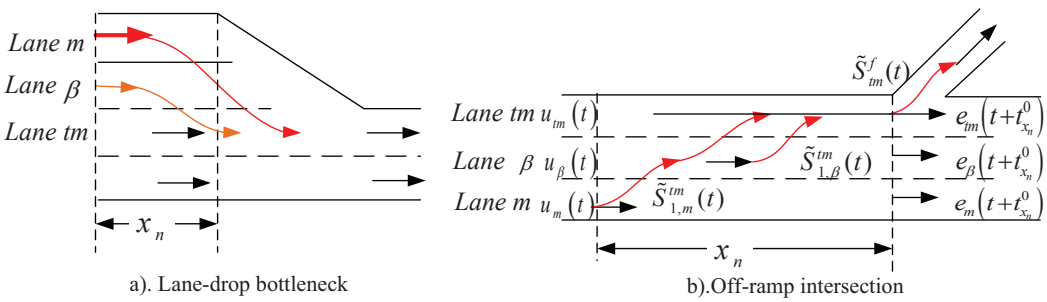

Figure 5: The total MLC demand under different scenarios.

blocks at these boundaries, as it will be shown in the empirical study. x_n is the farthest remaining distance away from the target turning point (which is regarded as the reference point whose remaining distance is 0) where the MLC has just been activated, and $t^0_{x_n}$ is the average time spent by the vehicles to travel from x_n to 0 (location of target turning point) via all the three lanes. Similar to the consistent check of the boundary conditions in the CTM (Muñoz et al., 2003), the values of the total MLC demand still need to be dynamically adjusted by considering the DLC maneuvers to adapt to the current traffic condition. Meanwhile we need to allocate this total demand along a lane to each cell till it arrives the desired turning point. The forthcoming sections will address these issues.

3.2. Longitudinal distribution of MLC and DLC demands

As depicted in Figure 5b, assume the traffic platoon $u_m(t)$ passes a location denoted x_n on lane m at time t, and a part of $u_m(t)$ say $\tilde{S}^{tm\,(r)}_{1,m}(t)$ would like to switch to its terminal lane tm, which is connected to the desired off-ramp. The cumulative MLC demand from lane m to the first target lane β ($\beta = m \pm 1$) follows an exponential distribution with respect to the remaining distance from x_n to another downstream location x, which is defined by $L^{\beta\,(r)}_{1,m}(t, x)$ (PCE/hour) as follows with $x \in [0,\ x_n]$:

$$
\begin{aligned}
L^{\beta\,(r)}_{1,\,m}(t,\ x) &= \begin{cases} e^{\left(-[x-x_c]^2/\left(\sigma^{\beta}_{m}(t)\right)^2\right)} \cdot \tilde{S}^{tm\,(r)}_{1,\,m}(t), & \text{if } x > x_c \\ \tilde{S}^{tm\,(r)}_{1,\,m}(t), & \text{if } x \le x_c \end{cases} \\
\sigma^{\beta}_{m}(t) &= \alpha_0 + \alpha_1 \cdot N^{tm}_{m} + \alpha_2 \cdot \bar{\rho}_{\beta}(t)
\end{aligned}
\tag{13}
$$

where x_c is the critical remaining distance, at which location all the related drivers have to execute their MLC maneuvers (Yang and Koutsopoulos, 1996), was introduced in Section 2.2. The function $\sigma^{\beta}_{m}(t)$ reflects the level of urgency related to the lane change decision making mechanism of $\tilde{S}^{tm\,(r)}_{1,\,m}(t)$ by switching from the initial lane m *to* lane β. This function is affected by two variables, i.e. N^{tm}_{m} and $\bar{\rho}_{\beta}(t)$, where N^{tm}_{m} denotes the number of lanes to be crossed from current lane m to terminal lane tm on the lateral dimension, $\bar{\rho}_{\beta}(t)$ denotes the average density of lane β along the longitudinal dimension from x_n to the target turning point 0 at time t, and α_0, α_1 and α_2 are the related parameters.

Equation (13) is a simplification of the exponential probabilistic distribution of cumulative MLC demand proposed in Yang and Koutsopoulos (1996). In their original definition, it is assumed that the driver can predict the short-term future traffic condition from her/his current location at the very moment till the time she/he reaches the destination (x=0). However, this is unlikely to be true especially under congested traffic condition. This renders the original model difficult to be calibrated. By Equation (13), drivers only use the current traffic condition spatially ahead that they could perceive to make their lane changing decisions. The effect of the perceived traffic condition decreases exponentially with respect to the distance away from her/his current position. While the calibration method for parameters α_0, α_1 and α_2 is not introduced in Yang and Koutsopoulos (1996), we will introduce a curve fitting based algorithm to calibrate these parameters in the empirical study.

Similarly, the longitudinal cumulative probability distribution of MLC demand from lane β to lane tm, with respect to the remaining distance from x_n to another downstream location x is estimated as:

$$
\begin{aligned}
L^{tm\,(r)}_{1,\,\beta}(t, x) &= \begin{cases} e^{\left(-[x-x_c]^2/\left(\sigma^{tm}_{\beta}(t)\right)^2\right)} \cdot \left(\tilde{S}^{tm\,(r)}_{1,\,\beta}(t) + L^{\beta\,(r)}_{m}(t,\ x)\right), & \text{if } x > x_c \\ \tilde{S}^{tm\,(r)}_{1,\,\beta}(t), & \text{if } x \le x_c \end{cases} \\
\sigma^{tm}_{\beta}(t) &= \alpha_0 + \alpha_1 \cdot N^{tm}_{\beta} + \alpha_2 \cdot \bar{\rho}_{tm}(t)
\end{aligned}
\tag{14}
$$

In theory, the proportion of MLC demand assigned to *cell* $(i,\ m)$ (heading for cell $(i{+}1,\ \beta)$) is evaluated as:

$$s_{1,i,m}^{i+1,\,\beta\,(r)}(k) = L_{1,m}^{\beta\,(r)}(k\cdot T_s,\ x_n-(i+1)\cdot l_i) - L_{1,m}^{\beta\,(r)}(k\cdot T_s,\ x_n - i\cdot l_i)\,. \tag{15}$$

However, lane changing demand might not be received by the target cell according to current traffic condition. Under such circumstance, the remaining part has to postpone the lane changing maneuvers. Thus $s_{1,i,m}^{i+1,\,\beta\,(r)}(k)$ is refined as:

$$s_{1,i,m}^{i+1,\,\beta\,(r)}(k) = L_{1,m}^{\beta\,(r)}(k\cdot T_s,\ x_n-(i+1)\cdot l_i) - I_{1,m}^{\beta\,(r)}(k\cdot T_s,\ x_n - i\cdot l_i) \tag{16}$$

where $I_{1,\,m}^{\beta\,(r)}(k\cdot T_s,\ x_n - i\cdot l_i)$ denotes the cumulative MLC demand which was originally proposed at time k and position x_n but actually be executed at the downstream boundary of cell i, which can be evaluated as:

$$I_{1,m}^{\beta\,(r)}(k\cdot T_s,\ x_n - i\cdot l_i) = \sum_{w=1}^{i-1} q_{1,w,m}^{w+1,\,\beta\,(r)}(k) \tag{17}$$

Apart from the MLC demand which have not been executed, the remaining sending flow of *cell* (i, m), i.e. $s_{i,m}^{i+1\ (r)}(k) - \left(\tilde{S}_{1,\,m}^{tm\,(r)}(k\cdot T_s) - I_{1,m}^{\beta\,(r)}(k\cdot T_s,\ x_n - i\cdot l_i)\right)$, should either go straightly towards *cell* (i+1, m) or seek for a better driving condition. Speed difference between adjacent lanes, and drivers' desire for traveling faster might trigger DLC maneuvers. The DLC demand for the 1^{st} round of estimation can be thus estimated:

$$\tilde{s}_{2,i,m}^{i+1,\,\beta\,(r)}(k) = \frac{\max\left(0,\, v_{i,\,\beta}(k) - v_{i,m}(k)\right)}{v_{f,i,m}\cdot\tau}\cdot\left\{s_{i,m}^{i+1\ (r)}(k) - \left(\tilde{S}_{1,\,m}^{tm\ (r)}(k\cdot T_s) - I_{1,m}^{\beta\ (r)}(k\cdot T_s,\ x_n - i\cdot l_i)\right)\right\} \tag{18}$$

where τ can be interpreted as the average time a driver takes to decide and execute a lane change when the origin lane is stopped and the target lane is freely flowing (Laval and Daganzo, 2006; Laval and Leclercq, 2008). Along the longitudinal dimension, the total DLC flow sent by lane m to lane β from remaining distance x_n to 0 is estimated by:

$$\tilde{S}_{2,\,m}^{\beta\,(r)}(t) = \tilde{S}_{2,\,m}^{\beta\,(r)}(k\cdot T_s) = \sum_{w=1}^{i_n-1} \tilde{s}_{2,w,m}^{w+1,\,\beta\ (r)}(k) \tag{19}$$

where the segment from x_n to 0 is partitioned into i_n cells. This estimation of initial values of total DLC demand enables a dynamical adjustment of the MLC demand, which will be depicted in Section 3.3.

3.3. The dynamical adjustment process of lane changing demand

Considering both MLC and DLC maneuvers, the r^{th} iteration of flow conservation for the off-ramp scenario in Figure 5b, can be rewritten as

$$\begin{aligned}
&\tilde{S}_{1,\,m}^{tm\ (r+1)}(t) + \tilde{S}_{2,\,m}^{\beta\ (r)}(t) + e_m(t+t_{x_n}^0) = u_m(t) + \tilde{S}_{2,\,\beta}^{m\ (r)}(t)\\
&\tilde{S}_{1,\,\beta}^{tm\,(r+1)}(t) + \tilde{S}_{1,\,m}^{tm\,(r+1)}(t) + \tilde{S}_{2,\,\beta}^{m\,(r)}(t) + \tilde{S}_{2,\,\beta}^{tm\,(r)}(t) + e_\beta(t+t_{x_n}^0) = u_\beta(t) + \tilde{S}_{1,\,m}^{tm\,(r+1)}(t) + \tilde{S}_{2,\,m}^{\beta\,(r)}(t) + \tilde{S}_{2,\,tm}^{\beta\,(r)}(t)\\
&\tilde{S}_{tm}^{f\,(r+1)}(t) + \tilde{S}_{2,\,tm}^{\beta\,(r)}(t) + e_{tm}(t+t_{x_n}^0) = u_\beta(t) + \tilde{S}_{1,\,m}^{tm\,(r+1)}(t) + \tilde{S}_{1,\,\beta}^{tm\,(r+1)}(t)
\end{aligned} \tag{20}$$

Following the iteration as depicted by Equations (13)-(20), the estimation of vector $\left[\tilde{S}_{1,\,m}^{tm\,(r)}(t),\ \tilde{S}_{1,\,\beta}^{tm\,(r)}(t),\ \tilde{S}_{tm}^{f\,(r)}(t)\right]^T$ becomes $\left[\tilde{S}_{1,\,m}^{tm\,(r+1)}(t),\ \tilde{S}_{1,\,\beta}^{tm\,(r+1)}(t),\ \tilde{S}_{tm}^{f\,(r+1)}(t)\right]^T$ by the end of round r, where r is the iteration counter. If the

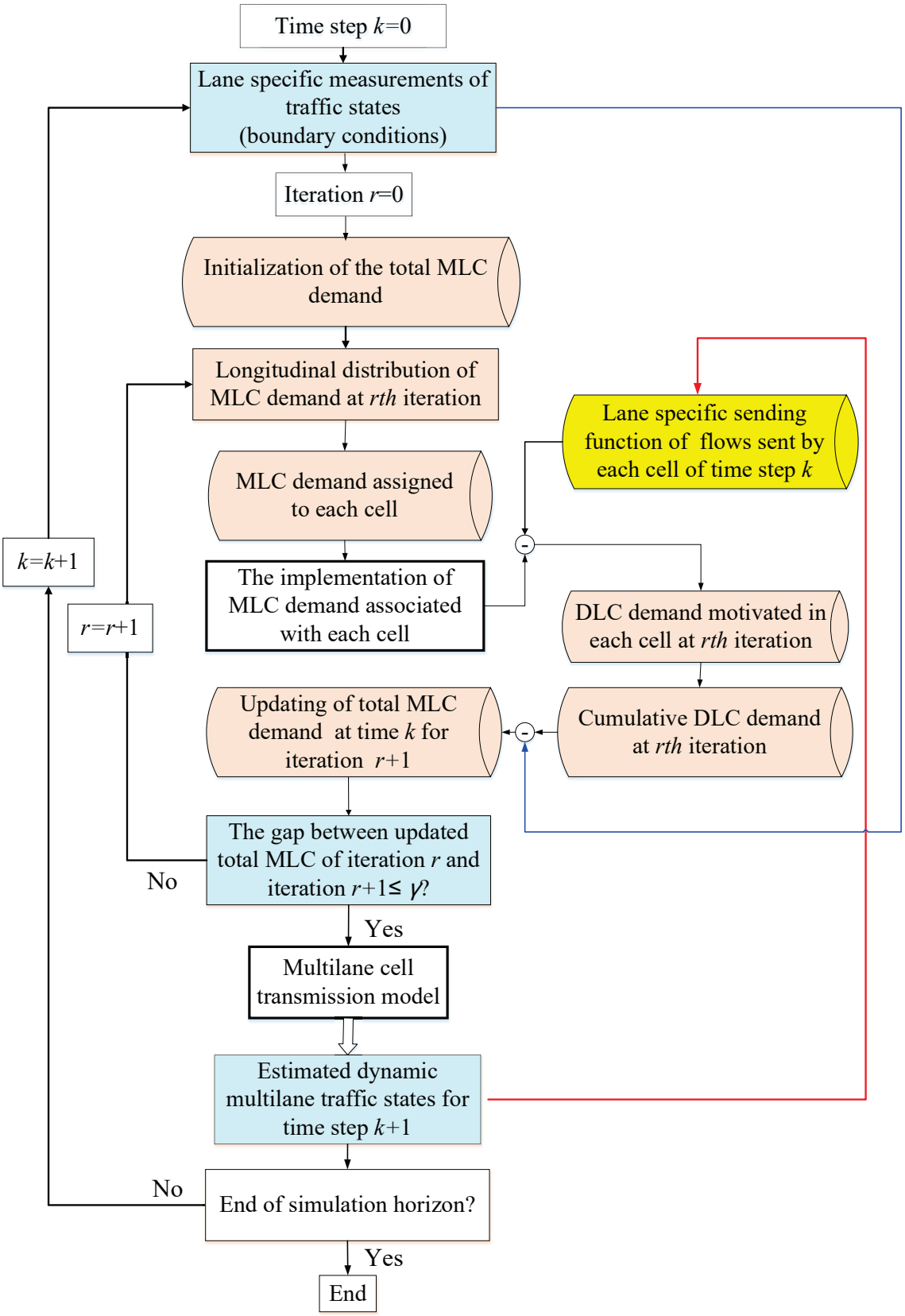

Figure 6: The flow chart for MLC and DLC demand estimation.

gap between these two vectors as mentioned above is less than or equals to the prescribed threshold γ, i.e.

$$\frac{\left\| \left[\tilde{S}^{tm\,(r+1)}_{1,\,m}(t),\ \tilde{S}^{tm\,(r+1)}_{1,\,\beta}(t),\ \tilde{S}^{f\,(r+1)}_{tm}(t) \right]^T - \left[\tilde{S}^{tm\,(r)}_{1,\,m}(t),\ \tilde{S}^{tm\,(r)}_{1,\,\beta}(t),\ \tilde{S}^{f\,(r)}_{tm}(t) \right]^T \right\|}{\left\| \left[\tilde{S}^{tm\,(r)}_{1,\,m}(t),\ \tilde{S}^{tm\,(r)}_{1,\,\beta}(t),\ \tilde{S}^{f\,(r)}_{tm}(t) \right]^T \right\|} \leq \gamma, \tag{21}$$

then the iteration terminates. Otherwise, the algorithm repeats. This dynamical adjustment process of lane changing demand is illustrated in Figure 6.

4. Empirical study

In this section, traffic dynamics of a complex weaving section be of three lanes with significant MLC is investigated. Based on the traffic data provided by lane specific loop detectors, the proposed mesoscopic multilane cell transmission model is calibrated. After that, the proposed model is validated on this section to reveal negative impacts and positive effects of the MLC and DLC maneuvers under various traffic conditions such as congestion onset and dissolve processes. Moreover, this empirical study will also verify that the model requires no additional data other than the CTM. Thus, it can be deployed as a simple simulation tool for accessing traffic state from data available to most management centers.

4.1. Description of the test site

The region of interest for simulation is a section of freeway SR-241 south bound. This section, located in the east of Orange County, California, stretches from 33°47'41.63"N, 117°43'56.92"W (as indicated by **OR** in Figure 7 to 33°47'6.69"N, 117°44'47.81"W (position **EN** in Figure 7. The main road is composed of three lanes at the first 0.9 mile, and reduces to two lanes when SR-241 forks to SR-241 and SR-261 as depicted in Figure 7. After this diverging node, the main section stretches another 0.1 miles until the position EN on SR-241. The first 0.3 miles of SR-261 with two lanes connected to inner lane and central lane of main road is also modeled in this empirical study. An off-ramp to Chapman Avenue is connected to the main road of SR-241 through an intersection that is around 0.2 miles upstream to the diverging node. Overall, a vehicle may exit this weaving section to one of the three directions, i.e. Chapman Avenue towards Santiago Canyon community, State Route-261 towards Irvine city, or continues travelling on State Route-241 towards the southern part of Orange County. The gantry information board, which is located 0.75 mile upstream to the diverging node, reminds the drivers about the remaining distance away from the downstream off-ramp locations and diverging point. Therefore, this 1 mile section between **OR** and **EN** is considered to be the weaving section. For modeling purpose, the main road of the test section is partitioned into 10 cells with equal length of 0.1 miles as demonstrated in Figure 7 to fulfill the Courant–Friedrichs–Lewy (CFL) condition that is a vehicle cannot travel across more than one cell during one simulation time interval (3 seconds in this empirical study) with the maximal free-flow speed (74.3 mile/hour for the median lane). In addition, the off-ramp intersection between the diverging node and detector station 1211563 on SR-261 is partitioned into 3 cells, and so does the off-ramp section on Chapman Avenue by detector station 1211577. To show the capacity difference between the weaving and non-weaving sections, the lane-specific FDs of the 1-mile segment upstream to **OR**, which is densely instrumented with loop detectors, are also calibrated. Based on the lane-based detector stations, the PeMS is capable of updating the aggregated traffic data, including traffic volume, occupancy, speed, and truck proportion every 5 minutes[5].

[5] The raw data of PeMS, such as traffic volume and occupancy, are updated every 30 seconds. However, on the one hand, the raw data does not provide the values of truck proportion; and on the other hand, the data loss of raw data greatly affect the analysis on lane flow distribution.

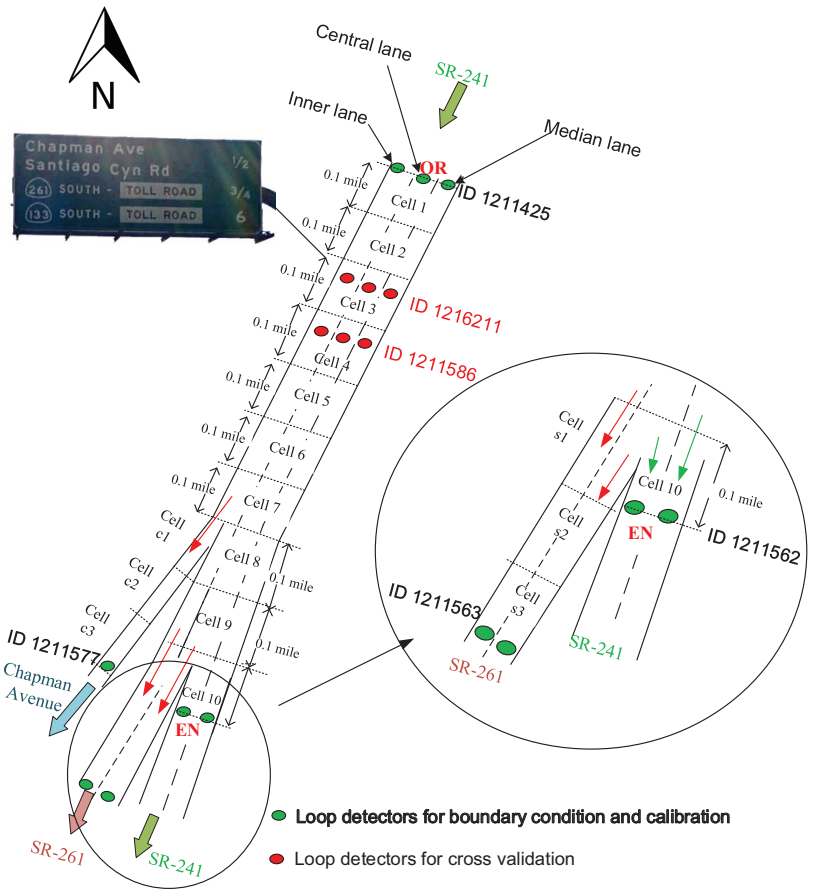

Figure 7: Topology of the test site.

The selected data of the two days covers different traffic conditions, e.g. all lanes were free-flowing on June 2, 2015, while they suffered from serious congestion on the morning of March 20, 2014 due to the congestion spilled back from the SR-261. And in particular, the data recorded on the two days has much less data loss and error due to fewer ill-functioning of detectors compared with most of the other days. This weaving section is chosen since, as it will be shown later, the heavy traffic during the peak hours would trigger an obvious queuing, complex DLC and MLC maneuvers due to queuing and weaving, as well as the subsequent congestion dissolve processes can be observed on March 20, 2014. For example, high MLC demand and DLC demand would induce a drop of discharging rates on the weaving section during the congestion period due to the complex topology and heavy traffic as it will be shown later in Section 4.3.

4.2. Model calibration and estimation of MLC demand

Assume that each pair of lane specific traffic flow and density measured by the same detector satisfies a specific triangular relationship as introduced in Section 2.1. In this empirical study, cell parameters of the main road are calibrated using conventional least squares method in line with the first-order macroscopic traffic flow literature as shown in Table 1. Their graphical counterparts have been demonstrated in Section 2.1. For those cells without detectors, the calibration adopts the data observed by the most adjacent detectors on the main road. Note that some parameters cannot be calibrated due to the limited amount of congestion

Table 1: The calibration of fundamental diagrams.

Detector station ID	Related cells	Inner lane			Central lane			Median lane		
		v_f	Q_m	w_c	v_f	Q_m	w_c	v_f	Q_m	w_c
1211425	cell 1	58.3	1903.1	11.4	65.8	2002.9	11.8	74.3	1857.8	10.6
1216211	cells 2-3	58.5	1804.0	10.7	55.4	1929.7	11.4	66.5	2083.7	11.6
1211586	cells 4-9	59.1	1958.4	11.8	55.4	2204.9	13.8	66.5	1941.4	11.4
1211562	cell 10	—	—	—	45.6	N/A	N/A	59.1	1956.0	11.3
1211563	cells s1-s3	58.2	1476.0	12.8	—	—	—	55.0	1632.0	14.2
1211577	cells c1-c3	77.8	N/A	N/A	84.2	N/A	N/A	79.2	N/A	N/A

data at certain locations. Alternatively, one can apply the automatic calibration algorithm by Zhong et al. (2015) to interpolate the parameters for those with no measurement devices. The FDs of SR-261 and off-ramp sections (after the weaving section) are similarly calibrated.

In order to estimate the initial time dependent longitudinal distribution of MLC demand by Equation (13), the median lane traffic flow data observed during the non-rush hour on June 2, 2015 (excluding March 20, 2014) is firstly applied to calibrate the parameters of this equation, i.e. α_0, α_1 and α_2 and x_c. The main reasons for using this day are: a) The median lane would most possibly send out MLC demand (towards the fork and off-ramp intersection) to central lane without receiving flows with MLC intention in return under light traffic conditions of median lane in conjunction with normal conditions (e.g. incident free conditions) of the central lane; b) According to the lane specific traffic speed measurements by the adjacent detecting stations, the speed difference between the median lane and the central lane during non-rush hour is always less than 5 mile/hour on June 2, 2015, therefore the DLC demand should be negligible based on the assumption that a DLC is mainly motivated by sufficiently large speed advantage of the adjacent lanes. c) The traffic state of the central lane maintains free-flowing for the whole day, whose density is lower than 25 *PCE/mile/lane* during non-rush hour in particular. Therefore the space between the vehicles travelling on central lane would be no less than 200 feet on average during the non-rush hours. While noting that a space of 200 feet (61 meters) is large enough to receive almost all the MLC flows. d) The vehicles traveling on the median lane must move across two lanes if they would like to exit the freeway via Chapman Avenue off-ramp while they only need to cross one lane to the SR 261. This topology feature enables the calibration of the weighting associated with the number of lanes to be crossed by MLC flows, i.e. α_1. To sum up, the gradual reduction of traffic flow on the median lane along the longitudinal dimension is mainly caused by MLC maneuvers towards the fork and off-ramp intersection. In other words, the longitudinal lane flow distribution on the median lane would reflect the cumulative effect of MLC maneuvers for a short time period such as five minutes within which the traffic state would not admit a sharp change.

The calibration procedures for the parameters α_0, α_1 and α_2 can be summarized as follow: The first step is to identify $\sigma(t)$ through a nonlinear curve fitting algorithm based on the longitudinal variation of median lane flows measured at t simultaneously. The curve fitting algorithm tries to identify $\sigma(t)$ in the first equation of (13) such that the distance between the fitted curve (red line) and the measured data (black star) is minimized. Then a linear regression algorithm is applied to identify the parameters α_0, α_1 and α_2 by acknowledging the average dynamic traffic density of central lane $\rho_{central}(t)$ (central lane as target lane) on weaving section and number of lanes N_{ln} to be crossed to reach the terminal lane (central or inner lane).

Figure 8 demonstrates the cumulative impact of MLC maneuvers on the longitudinal lane flow variation of the median lane on SR-241 on June 2, 2015 with respect to the remaining distance to the target turning

Table 2: The parameters in mesoscopic multilane model.

Longitudinal cumulative distribution function of MLC demand (Calibrated based on the PeMS data collected on June 2, 2015)	α_0 α_1 Weighting associated with the number of lanes to be crossed α_2 Weighting associated with density of the target lane x_c Critical remaining distance x_r Remote remaining distance	-55.9 726.9/lane 33.7 mile/PCE 0.05 mile (265 feet) 1.0 mile (5280 feet)
DLC demand determination function (Laval and Daganzo, 2006)	τ the average lane changing reaction time	3s (2-4 s are adopted for the SA)
Minimum gap acceptance criterion function (Hidas, 2005)	c_l Parameter of extra leading gap c_f Parameter of extra lag gap $\bar{g}_{\min}$	1.32 feet hour/mile 1.32 feet hour/mile 37.7 feet

point, e.g. the diverging point to SR 261 or the off-ramp intersection connected to Chapman Avenue. For example, the upmost and the middle figures assume that the MLC is initialized at the first detection station (ID 1211425), while the diverging point before detection station ID 1211562 is regarded as the destination in the morning[6] of June 2, 2015. Comparing the longitudinal lane flow variation of the median lane during 6:40-6:45 and 10:30-10:35, we can observe that drivers with MLC intentions, might decide to change lane earlier during 6:40-6:45 than 10:30-10:35 due to a higher average density on the central lane, which is reflected by a larger σ during 6:40-6:45. To be specific, during 6:40-6:45, 50% MLC demand can be completed with a remaining distance of 0.33 miles to the target turning point while this becomes closer (0.27 miles) to the intended turning point during 10:30-10:35 due to the lower level of urgency (lower traffic density after peak hour). On the other hand, although the average density of central lane observed during 16:15-16:20 is close to that of 10:30-10:35, a farther lateral distance to the target turning point (two lanes away from the off-ramp intersection connected to Chapman Avenue on lateral dimension) might strengthen the drivers' sense of urgency to change lanes. Therefore, 50% MLC demand is completed earlier with a remaining distance up to 0.34 miles to the off-ramp intersection during 16:15-16:20 which is indicated by a larger σ. The average reaction time for a DLC is calibrated by Laval and Daganzo (2006). The minimum safe gap is calibrated by Hidas (2005). Here we adopt their typical values as reported in the literature as indicated in Table 2.

4.3. Simulation results against the measurement

Based on the parameters calibrated in Section 4.2, a model validation is conducted using the inflow profile and boundary conditions observed on the morning of March 20, 2014 as input profiles to simulate the proposed mesoscopic multilane cell transmission model that considers both MLC and DLC maneuvers. It is worthwhile to point out that, in what follows, the measurements in the middle of weaving section, i.e. those given by detector stations **ID 1216211** and ID **1211586** on March 20, 2014, are not regarded as boundary conditions in the simulation of this test day. Instead, we pretend both the detectors were missing, i.e. we do not use their measurements for simulation, but use them for cross-validation.

[6] According to data observation, the mandatory lane changing ratios assigned to Chapman Avenue and SR-261 does not maintain stationary. In the morning, the traffic leaving SR-241 via SR-261 dominates the mandatory lane changing behaviors, while in the afternoon Chapman Avenue becomes busier than SR-261.

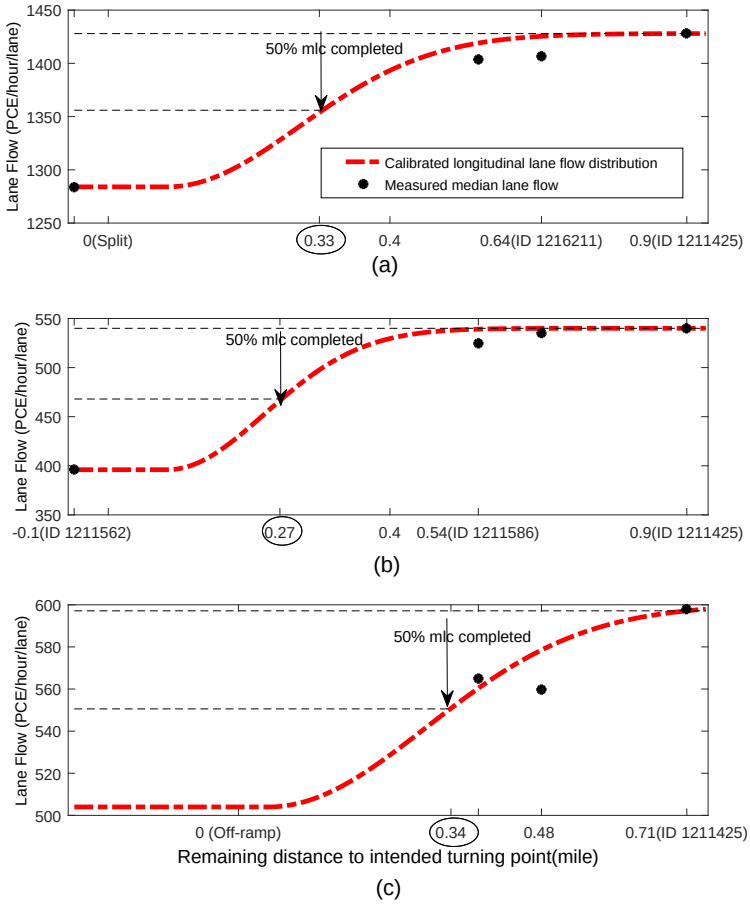

Figure 8: Longitudinal variation of median lane flow due to MLC maneuvers: a) During 6:40-6:45, with $\rho_{central} = 23PCE/mile$, $N_{ln} = 1$ and $\sigma = 1561$; b) During 10:30-10:35, with $\rho_{central} = 11PCE/mile$, $N_{ln} = 1$ and $\sigma = 1149$; c) During 16:15-16:20, with $\rho_{central} = 11PCE/mile$, $N_{ln} = 2$ and $\sigma = 1867$.

Table 3: Cell density estimation against measurement on March 20, 2014.

	Inner lane	Central lane	Median lane	Aggregated	CTM
Simulated traffic density of *cell 3* against the measurement of ID 1216211	16.7%	13.2%	18.3%	**10.7%**	13.4%
Simulated traffic density of *cell 4* against the measurement of ID 1211586	16.0%	16.5%	18.3%	**9.4%**	14.6%

Table 4: The absolute variation percentage of lane flow.

	Inner lane	Central lane	Median lane
The estimated lane flow variation	7.66%	7.60%	2.13%
The measured lane flow variation	6.12%	6.24%	2.55%

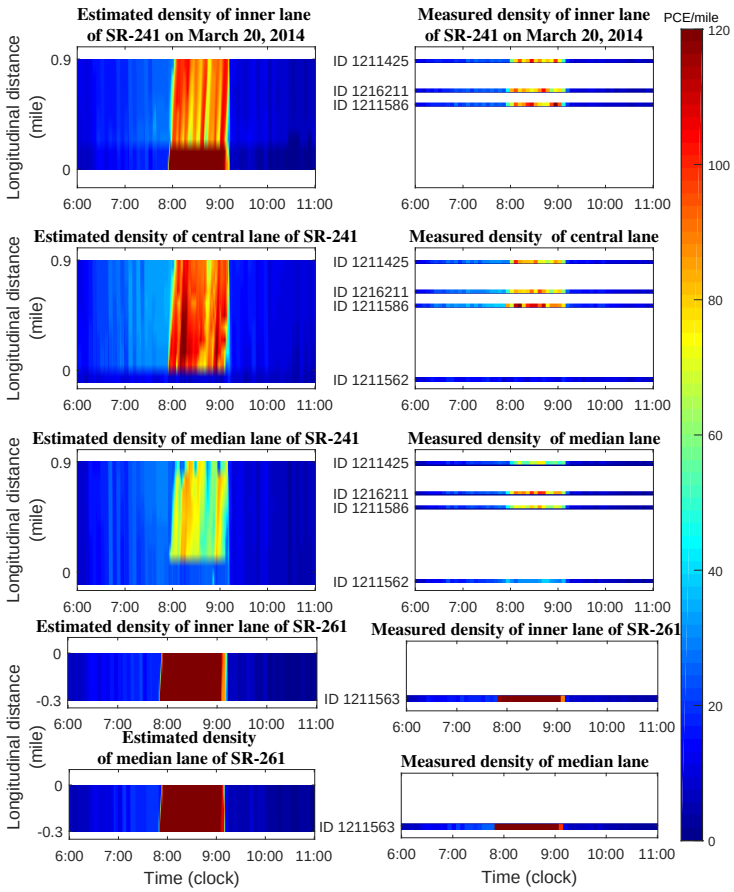

Figure 9: Heat map of estimated multilane density against measurement on March 20, 2014.

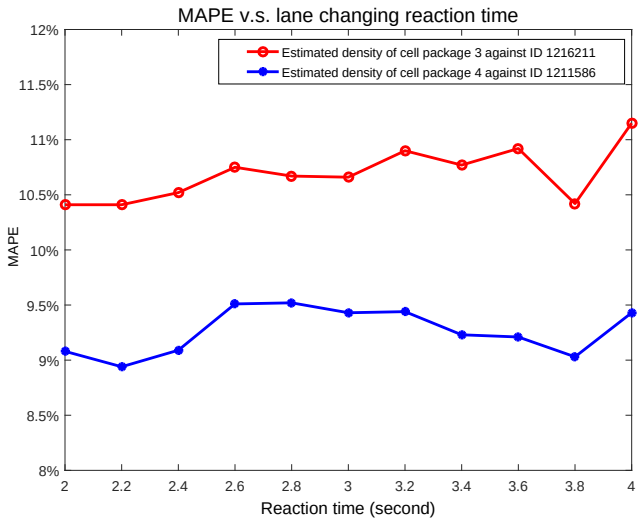

Figure 10: One-at-A-Time sensitivity analysis on the lane changing reaction time.

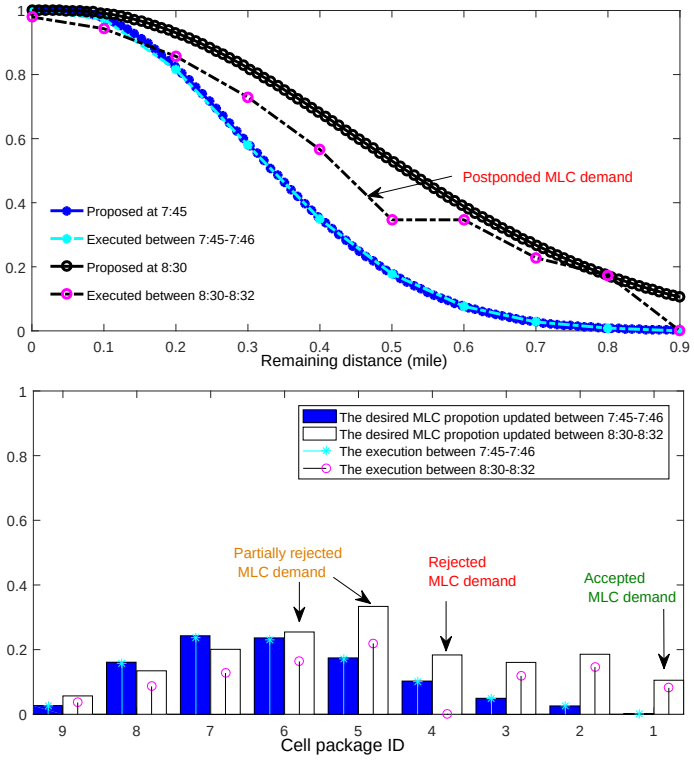

Figure 11: Proposed MLC demand and execution on longitudinal dimension.

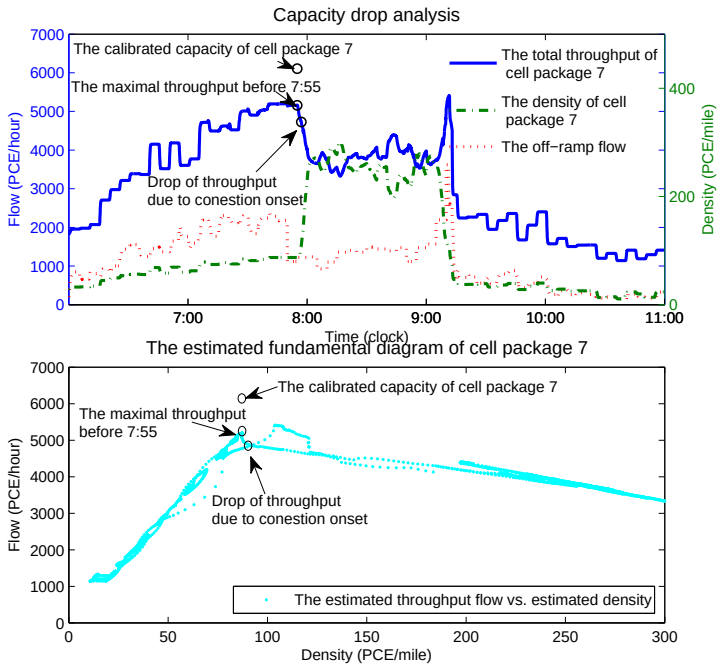

Figure 12: Capacity drop identification at *cell package* 7 within the weaving section.

Figure 9 presents the heat map of the simulated traffic state (left graphs) of the test segment against its observation counterpart (right graphs), wherein the fork of SR-241 and SR-261 is chosen as the reference point (i.e. 0 mile in longitudinal distance). On the whole, the simulated lane-specific traffic density captures the trend of measurements. Generally speaking, each estimated lane based cell density (with a resolution, i.e. the simulation time step of 3 seconds) is close to its measured counterpart (with a resolution of 5 minutes of PeMS data) under both heavy and light traffic conditions from 6:00-11:00 on the test day. In particular, the transient states, i.e. the longitudinal congestion onset, the lateral congestion onset and congestion offset, are also well captured by the proposed model. At the beginning, there was no congestion queue along the whole segment. Under such scenario, the lane-flow distribution was almost stable and evenly distributed until 7:50 as indicated in Figure 9. A congestion was first simultaneously observed around 7:50 on both the two lanes of SR-261, i.e. the downstream section after the weaving area, to be specific the fork of SR-241 and SR-261. Then the weaving section suffered from a severe congestion due to the queue spilled back from SR-261. Unlike the conventional CTM, traffic flow characteristics of different lanes along the main road are with some subtle and important differences as demonstrated by the simulation results presented in the left-hand side of Figure 9. Since SR-261 is directly connected to the main road SR-241 via its inner lane and central lane as shown in Figure 7, the drop of discharging rate due to the congestion on SR-261 directly affected the traffic on these two lanes. In contrast, the traffic on the median lane of SR-241 was not influenced by this congestion during its early stage (7:50-8:00).

The congestion on central lane slows down or partially rejects the MLC flows that intended to merge into the central lane from the median lane, due to the limited available space of the lane and small space gaps between successive vehicles on the congested central lane. Therefore, **the congestion is gradually spread to the median lane**. However, compared with the central lane, the median lane still provides better driving condition. This imbalance between adjacent lanes (median lane and central lane) triggers some vehicles, which were originally traveling on the central lane of SR 241 and did not intend to leave SR-241 within the current weaving section, switched to the median lane at the upstream cells to seek for a better driving condition so as to avoid congestion. Such DLC maneuvers further increased the traffic density on the median lane while slightly alleviated the congestion on the central lane. Therefore, the central lane was less congested than the inner lane, in particular on *cell 8* and *cell 9*. This phenomenon is known as **the balancing effect of DLC**, i.e. lane changes could smooth out the differences between adjacent lanes.

Also since the drivers tend to switch to the median lane long before they could arrive at the congestion site for better driving condition and to avoid congestion as previously discussed, the congestion on the median lane started from upstream cells, see e.g. detector station ID 1216211, rather than the downstream bottleneck location. This congestion did not cause severe impact as shown by the simulation and measurements of the detector stations ID 1211586 and ID 1211425 on the median lane. This result is unlike some existing models which assume drivers do not **perceive the downstream traffic condition** to make their lane changing decisions, i.e. each driver decides to change lane or not based on the perceived traffic condition of her/his current location. Shiomi et al. (2015) realized such an unrealistic assumption on LC maneuver and encouraged future effort to review and to overcome this limitation. The proposed model can be regarded as a possible solution to this stream. Finally the congestion spreads to all lanes. The proposed model can capture all these phenomena.

During the congestion period, the off-ramp connected to Chapman Avenue helps in alleviating the congestion level of cells upstream to cell 7 on inner lane as indicated by the simulation. This is because the vehicles still get a chance to exit the freeway from the off-ramp to Chapman Avenue, while those enter the weaving section of SR-261 do not have such a luxury to escape from the congestion. Note that the section of SR-241 downstream to the diverging point never gets congested (both the central and median lanes). This

is a typical phenomenon that the downstream section of a bottleneck is usually free flowing. After the peak hour period, all the lanes return back to free-flowing condition, and the lane-flow distributions go back to almost stable again.

As mentioned previously the measured data of detector station ID 1216211 and ID 1211586 is used for cross validation, the corresponding lane specific MAPE (Mean Absolute Percentage Error) of the density estimation is reported in Table 3, where the MAPE is evaluated as

$$MAPE_{i,\,m} = \frac{1}{N}\sum_{K=1}^{N}\frac{\left|\bar{\rho}_{i,\,m}(K) - \rho^{M}_{i,\,m}(K)\right|}{\rho^{M}_{i,\,m}(K)} \tag{22}$$

Here $\bar{\rho}_{i,\,m}(K)$ denotes the average density of *cell* $(i,\,m)$ estimated during the time interval $[kT_s, (k+1)T_s)$, in accordance with its measurement counterpart $\rho^{M}_{i,\,m}(K)$ whose resolution is 5 minutes. N denotes the amount of observations during the whole simulation horizon. $\rho^{M}_{i,\,m}(K)$ is provided by the detector station most close to *cell package* i *on* the same lane. The MAPE for lane specific traffic density estimation ranges from around 13.0% to 18.5% which indicate that the proposed model can capture the lane-changing traffic dynamics well. To compare with the CTM without considering lane heterogeneity nor lane changing maneuvers, both the simulated and measured lane specific traffic densities are aggregated, i.e. summing the densities of the three lanes. Defining the aggregated MAPE as the aggregated absolute value of the discrepancy between the simulated data and measured data over the aggregated measured data, it is found that the aggregated MAPE is around **9%-11%**. This result is more accurate than that of traffic density by the **CTM** which is around **13%-15%**. To test the effect of the average reaction time, an One-at-A-Time (OAT) sensitivity analysis is conducted. As shown in Figure 10 that the proposed lane changing model is not sensitive to τ for a small change of τ ranging from 2.0-4.0 sec in the sense that the MAPE of the estimation does not admit an abrupt change.

As it has been discussed in Section 4.1, the longitudinal distribution of MLC demand, which is assumed to be initialized at the beginning of this weaving section (with a remaining distance of 0.9 miles to the diverging point this empirical study), is affected by the perceived level of congestion of the target lane (central lane) as well as the number of lanes to be crossed on the lateral dimension. To demonstrate this phenomenon, the upper graph of Figure 11 depicts the longitudinal cumulative distribution functions of MLC demand initially proposed at different times on the test day, e.g. 7:45 and 8:30, against the executed ones. This graph indicates that drivers intend to execute a MLC maneuver earlier when facing the congested condition (i.e. 8:30 with $\rho_{central} = 76 pcu/mile$) than the free flowing condition (i.e. 7:45 with $\rho_{central} = 32 pcu/mile$). This earlier plan of MLC is proposed to ensure the execution of lane changing maneuvers before arriving the turning point in case that the congested target lane might reject the lane changing demand due to its limited available space as well as the small gaps between vehicles. This phenomenon is demonstrated in the lower graph of Figure 11 that the MLC demand proposed at 8:30 was firstly partially rejected by *cell 2*. Therefore, this rejected MLC demand has to postpone to the downstream cells. However, the updated proportion of MLC demand assigned to *cell 3* and the original assignment were still partially rejected by central lane. Similar observations can be drawn from other cells, especially the completed rejection on cell 4. Finally, cell 9 accepted all the MLC assignment, because people with MLC intension have to squash into the target lane regardless whether there is a minimum acceptable gap or not as discussed in Section 2.2. In conclusion, the MLC demand which was sent from median lane and proposed at 7:45 was well executed during 7:45-7:46 by arriving the diverging node due to the free-flowing traffic condition on central lane. However, due to the saturated traffic condition of the central lane at 8:30, which is nearly 3 times of that at 7:45, the proposed demand was not executed as planned but postponed to downstream cells,

and cost longer time compared with the one proposed at 7:45. Such postponement proves the necessity of proposing an early MLC at upstream during congestion.

To compare the simulated and measured lane changing effects, the cumulative lane flow variation analysis of this weaving section, i.e. from OR to EN, is conducted wherein the Chapman Avenue off-ramp and SR-261 is regarded as the stretch of the inner lane for comparison. The results are depicted in Table 4. It is assumed in this test that the cumulative lane flow difference between the upstream OR and the downstream EN during the simulation horizon is caused by executed MLC and DLC (If there is no lane change, the cumulative upstream flow should equal to the downstream by flow conservation, excluding the flow on the freeway at the initial time and the remaining flow at the end time). As shown in Table 4, the estimated lane flow variations (in terms of percentage) are very close to their measurement counterparts which proves the accuracy of the proposed lane changing model.

A conventional way to observe capacity drops is to compare the flow just upstream of the bottleneck and the flow just downstream of the bottleneck as mentioned by Cassidy and Bertini (1999), e.g. Section 2. Another way is to depict the flow and density relationship at the bottleneck site. In line with Srivastava and Geroliminis (2013), we present a representative throughput time series plot along with the corresponding off-ramp flows and the mainline density at the bottleneck in Figure 12. The total throughput of *cell package* 7 can be seen here to decrease from a maximum of about 5150 PCE/hour between 7:45–7:55 AM, to about 4700 PCE/hour after 7:56 AM, when the density of *cell package* 7 maintains at critical density. The off-ramp flow was increasing until the weaving section entered the congested period, then both the throughput and the off-ramp flow drop. Note that both the ramp flow and the throughput significantly increases before the breakdown. The congestion blocks both the MLC for the off-ramp and the throughput, until both the off-ramp and the throughput admits a sharp increase around 9:15 AM after the congestion was alleviated.

5. Conclusions and future work

Modeling lane-changing maneuvers is essential to capture several important characteristics of multilane traffic flow, e.g. heterogeneous traffic flow distribution, capacity drop, and flow balancing effect. To this end, a novel mesoscopic multilane model was proposed to simulate the the effect of MLC and DLC maneuvers in this paper. The proposed model extends the multilane hybrid (MH) theory by incorporating the lane-based fundamental diagrams to capture the relation between speed and lanes which is believed to be missing in most existing models (Keyvan-Ekbatani et al., inpress). Meanwhile, different priority levels were identified according to the lane-changing motivations and the corresponding levels of urgency. In particular, a recursive lane changing demand estimation algorithm that considers the impact of level of urgency on longitudinal probability distribution function of lane changing maneuvers was devised. Flow propagations of both MLC and DLC maneuvers were calculated by demand-supply reaction laws based on the extended IT and PIT principles.

Using traffic data from the PeMS, the proposed mesoscopic multilane cell transmission model was calibrated and validated on a complex weaving section of the SR241 freeway in Orange County, California. The results indicated that the proposed model can capture the impacts of lane changing maneuvers on the temporal and spatial traffic state, especially the lateral lane flow distribution and the queuing effect on longitudinal dimension in conjunction with congestion spreading to adjacent lanes. The MLC demand estimation algorithm and the lane changing probability distribution function on longitudinal dimension fitted their measurement counterparts in a satisfactory manner. This implies the model can be also used to infer turning ratios (e.g. off-ramp demand) for ramps. Unlike some existing models which assume drivers do not perceive the downstream traffic condition to make their lane changing decisions, which is regarded as unrealistic by Shiomi et al. (2015), the proposed model assumes that the drivers use the traffic condition

spatially ahead that they could perceive to make their lane changing decisions. The effect of the perceived traffic condition decreases exponentially with respect to the distance away from her/his current position. This exponential probabilistic distribution function of the cumulative MLC was calibrated and validated by the empirical results. With the same data source, the proposed model outperforms the CTM in terms of accuracy. Meanwhile some important effects of lane changing such as capacity drop and flow balancing effect of DLC were presented in empirical study which are essentially in agreement with previous findings in the literature. As proven by the empirical study, the proposed models do not require high resolution traffic data but traffic data available to most management centers which is a significant improvement over the existing models.

As freeways are subject to demand and supply uncertainties and the measured traffic data is noisy, extending the model to consider these random elements would be an interesting future work. Developing a comprehensive calibration procedure considering the stochastic elements, e.g. the data noise and model uncertainty, is also an important future study. As lane changes affect bottleneck behavior, to apply the model to investigate the benefit of controlling lane changes (e.g. by forbidding lane changes within the bottleneck and/or posting speed advisories at key locations upstream of bottlenecks) to the traffic flow is also interesting especially for the MLC maneuvers.

Acknowledgements

The work described in this paper is part of the PhD research work of the first author. This work was jointly supported by a Postgraduate Studentship and research grants from the Research Grant Council of the Hong Kong Special Administrative Region to the Hong Kong Polytechnic University (Project No. PolyU 152074/14E), the National Natural Science Foundation of China under Grant No. 51308559. The authors sincerely thank the referees for their insightful and constructive comments which have substantially improved the paper.

Key Notation list

	Calibrated parameters:
$v_{f,\,m}$	Free-flow speed of lane m.
$w_{c,m}$	Wave-back speed of congestion of lane m.
$\rho_{c,m}$	Critical density of lane m.
$\rho_{J,m}$	Jam-density of lane m.
Q_m	Capacity of lane m.
$\alpha_0\,, \alpha_1,\ \alpha_2$	Parameters of the cumulative distribution function of MLC demand on longitudinal dimension.
c_l, c_f	Coefficients associated with the speed difference for evaluating extra leading gap and extra lag gap.
$x_c,\ x_r$	Remaining distances by which the test section is partitioned as remote, median, and close sections, respectively, with each section corresponds to a specific level of MLC urgency and minimum acceptance criterion.
$\bar{g}_{min}$	The minimal safe gap for the subject vehicle provided by the target lane.
τ	The average reaction time of lane changing.
	Variables for dynamics propagation and LC acceptance/rejection assessment:
$\rho_{i,m}(k)$	Traffic density of *cell* (*i, m*) at time step *k*.
$v_{i,m}\,(k)$	Traffic speed of *cell* (*i, m*) at time step *k*.

$s_{i,\,m}^{i+1}(k)$	The sending function that determines the flow intended to leave *cell* (*i, m*) and towards downstream *cell package* (*i*+1) involving all possible lanes (e.g. the adjacent lanes *m*-1, *m*+1 and the current lane *m*), during time interval $[kT_s, (k+1)T_s)$.
$R_{i-1}^{i,\,m}(k)$	The receiving function of *cell* (*i, m*) towards which traffic flows sent by the upstream *cell package* (*i*-1) are intended to merge during time interval $[kT_s, (k+1)T_s)$.
$U_{i-1}^{i,\,m}(k)$	The total sending function that determines the amount of flow intended to merge to *cell* (*i, m*) from upstream *cell package* (*i*-1) during time interval $[kT_s, (k+1)T_s)$.
$s_{lc,\,i,\,m}^{i+1,\,\beta}(k)$	The sending function of lane-changing flow that intends to leave *cell* (*i, m*) and towards downstream *cell* (*i*+1,β) during time interval $[kT_s, (k+1)T_s)$ with $\beta = m \pm 1$, and the lane-changing intention is of *lc* type, where *lc*=1,2 represent the MLC and DLC, respectively.
$s_{st,\,i,\,m}^{i+1,\,m}(k)$	The sending function that determines the flow intended to leave *cell* (*i, m*) and move straightly to the downstream *cell* (*i*+1, *m*) during time interval $[kT_s, (k+1)T_s)$.
$q_{lc,\,i-1,\alpha}^{i,\,m}(k)$	The lane-changing flow that left *cell* (*i*-1,α) and received by the downstream target *cell* (*i, m*) during time interval $[kT_s, (k+1)T_s)$ with $\alpha = m \pm 1$ and the lane-changing intention is of *lc* type.
$q_{st,\,i-1,m}^{i,\,m}(k)$	The flow that left *cell* $(i-1,\ m)$ and received by the downstream target *cell* (*i, m*) during time interval $[kT_s, (k+1)T_s)$.
$x(k)$	Remaining distance from the current position of the subject vehicle to its target turning point.
$\tilde{g}_{lc,\,i-1,\,\alpha}^{i,\,m}(k)$	The minimum gap acceptance criterion required by the sending flow of *cell* (*i*-1, α) towards the target *cell* (*i*-1,*m*) during time interval $[kT_s, (k+1)T_s)$ with $\alpha = m \pm 1$ and the lane-changing intention is of *lc* type.
$\tilde{\phi}_{lc,\,i-1,\,\alpha}^{i,\,m}(k)$	The minimum acceptance criterion factor normalized from $\tilde{g}_{lc,\,i-1,\,\alpha}^{i,\,m}(k)$.
$\bar{G}_{\,i,m}(k)$	The average space gap between the successive vehicles on the *cell* (*i*, *m*) at time step *k*.

	Variables for determining LC demand and execution:
$\tilde{S}_{lc,\,m}^{tm\,(r)}(t)$	The estimated total *lc* type lane changing demand from lane *m* to the terminal lane *tm* at time *t* at the r^{th} iteration.
$L_{1,\,m}^{\beta\,(r)}(t,\ x)$	The estimated cumulative MLC demand at location *x* of lane *m* to the target lane β at time *t* at the r^{th} iteration.
$\sigma_m^{\beta}(t)$	Traffic state dependent parameter of the cumulative distribution function of MLC demand on longitudinal dimension.
$I_{1,m}^{\beta\,(r)}(k \cdot T_s, x_n - i \cdot l_i)$	The estimated cumulative MLC demand (originally proposed at time step *k* and position x_n) actually executed at the downstream boundary of cell *i* at the r^{th} iteration.
N_m^{tm}	The number of lanes to be crossed from current lane *m* to the terminal target lane *tm*.
	Abbreviation list:
DLC	Discretionary lane-changing
IT	Incremental-Transfer
LFD	Lane flow distribution
MFD	Macroscopic fundamental diagram
MLC	Mandatory lane-changing
PCE	Passenger Car Equivalent

References

Ahmed, K., 1999. Modeling Drivers' Acceleration and Lane-changing Behavior. *Ph.D Thesis*, Massachusetts Institute of Technology.

Ahn, S., Cassidy, M., 2007. Freeway traffic oscillations and vehicle lane-change maneuvers. In: Proc. Int. Symp. on the Transport. and Traffic Theory, Amsterdam, 691-710.
Ben-Akiva, M., Choudhury, C., Toledo, T., 2006. Lane-changing models. In: Proc. Int. Symp. Trans. Simu., Lausanne, Switzerland.
Carter, M., Rakha, H., Aerde, M., 1999. Variability of traffic-flow measures across freeway lanes. Can. J. Civ. Eng., 26, 270-281.
Carey, M., Balijepalli, C., Waling, D., 2015. Extending the cell transmission model to multiple lanes and lane-changing. Netw. Spat. Econ., 15 (3), 507-535.
Cassidy, M., Bertini, R., 1999. Some traffic features at freeway bottlenecks. Transport. Res. Part B 33, 25-42.
Cassidy, M., Rudjanakanoknad, J., 2005. Increasing the capacity of an isolated merge by metering its on-ramp. Transport. Res. Part B, 39 (10), 896-913.
Chen, C., 2003. Freeway performance measurement system (PeMS). PhD dissertation, University of California, Berkeley.
Cheu, R., Martinez, J., Duran, C., 2009. A cell transmission model with lane-changing and vehicle tracking for port of entry simulations. Transport. Res. Rec., 2124, 241-248.
Choudhury, C., Ben-Akiva, M., Rao, A., Lee, G., Toledo, T., 2007. State dependence in lane-changing models. *In Transportation and Traffic Theory*, Allsop, R., Bell, M., and Heydecker, B., ed., Elsevier, 711-73.
Coifman, B., Mishalani, R., Wang, C., Krishnamurthy, S., 2006. Impact of lane-change maneuvers on congested freeway segment delays: pilot study. Transport. Res. Rec., 1965, 152-159.
Daganzo, C., 1994. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. Transport. Res. Part B, 28 (4), 269-287.
Duret, A., Ahn, S., Buisson, C., 2012. Lane flow distribution on a three-lane freeway: General features and the effects of traffic controls. Transport. Res. Part C, 24, 157-167.
Keyvan-Ekbatani, M., Knoop, V., Daamen, W., inpress. Categorization of the lane change decision process on freeways. Transport. Res. Part C, doi:10.1016/j.trc.2015.11.012.
Gipps, P., 1986. A model for the structure of lane-changing decisions. Transport. Res. Part B, 20 (5), 403-414.
Golob, T., Recker, W., Alvarez, V., 2004. Safety aspects of freeway weaving sections. Transport. Res. Part A, 38 (1), 35-51.
Gunay, B., 2007. Car following theory with lateral discomfort, Transport. Res. Part B, 41 (7), 722-735.
Hidas, P., 2005. Modelling vehicle interactions in microscopic simulation of merging and weaving. Transport. Res. Part C, 13 (1), 37-62.
Hou, Y., Edara, P., Sun, C., 2015. Modelling mandatory lane-changing using Bayes classifier and decision trees. IEEE Trans. Intell. Transp. Syst., 15 (2), 647-655.
Jin, W., 2010a. A kinematic wave theory of lane-changing traffic flow. Transport. Res. Part B, 44(8-9), 1001-1021.
Jin, W., 2010b Macroscopic characteristics of lane-changing traffic. Transport. Res. Rec., 2188, 55-63.
Kesting, A., Treiber, M., Helbing, D., 2007. General Lane-changing Model MOBIL for Car-Following Models. Transport. Res. Rec., 1999, 86-94.
Knoop, V., Hoogendoorn, S., Shiomi, Y., Buisson, C., 2012. Quantifying the number of lane changes in traffic: an empirical analysis. Transp. Res. Rec. No. 2278, 31-41.
Laval, J., and Daganzo, C., 2006. Lane-changing in traffic streams. Transport. Res. Part B, 40 (3), 251-264.
Laval, J., Cassidy, M., Daganzo, C., 2007. Impacts of lane changes at on-ramp bottlenecks: A theory and strategies to maximize capacity. In: Kühne, R., Poeschl, T., Schadschneider, A., Schreckenberg, M., Wolf, D. (Eds.), Traffic and Granular Flow '05'. Springer, Berlin, 577-586.
Laval, J., Leclercq, L., 2008. Microscopic modeling of the relaxation phenomenon using a macroscopic lane-changing model. Transport. Res. Part B, 42 (6), 511-522.
Laval, J., Leclercq, L., 2010. A mechanism to describe the formation and propagation of stop-and-go waves in congested freeway traffic. Phil. Trans. Roy. Soc. A, 368 (1928), 4519-4541.
Lee, J., Cassidy, M., 2009. An Empirical and Theoretical Study of Freeway Weave Bottlenecks. PATH Res. Rep., UCBITS-PRR-2009-13.
Lighthill, M., Whitham, G., 1955. On kinematic waves. I Flow movement in long rivers. II A theory of traffic flow on long crowded roads. In: Proc. Roy. Soc. 229 (A), 281-345.
Mauch, M., Cassidy, M., 2002. Freeway Traffic Oscillations: Observations and Predictions. Transportation and Traffic Theory in the 21st Century, 653-673.
Munjal, P., Pipes, L., 1971. Propagation of on-ramp density perturbations on unidirectional two- and three-lane freeways. Transport. Res. Part B, 5 (4), 241-255.
Muñoz, L., Sun, X., Horowitz, R., Alvarez, L., 2003. Traffic density estimation with the cell transmission model. In: Proc. Am. Control Conf.. Denver, Colorado, USA, 3750-3755.
Newell, G., 1993. A simplified theory of kinematic waves in highway traffic. I: general theory. II: queuing at freeway bottlenecks. III: multi-destination flows. Transport. Res. Part B, 27, 281-313.
Patire, A., Cassidy, M., 2011. Lane-changing patterns of bane and benefit: Observations of an uphill expressway. Transport. Res.

Part B, 45 (4), 656-666.
Pipes, L., 1967. Car following models and the fundamental diagram of road traffic. Transport. Res., 1, 21-29.
Richards, P., 1956. Shockwaves on the highway. Oper. Res., 4, 42-51.
Schakel, W., Knoop, V., and van Arem, B., 2012. LMRS: An integrated lane change model with relaxation and synchronization. In: Proc. Trans. Res. Board, Washington, D.C.
Sheu, J., Ritchie, S., 2001. Stochastic modeling and real-time prediction of vehicular lane-changing behavior. Transport. Res. Part B, 35 (7), 695-716.
Sheu, J., 2004. A sequential detection approach to real-time freeway incident detection and characterization. Eur. J. Opera. Res., 157 (2), 471-485.
Shiomi, Y., Taniguchi, T., Uno, N., Shimatomo, H., 2015. Multilane first-order traffic flow model with endogenous representation of lane-flow equilibrium. Transport. Res. Part C, 59, 198-215.
Shvestsov, V., Helbing, D., 1999. Macroscopic dynamics of multilane traffic. Physic. Rev. E, 59 (6), 6328-6338.
Srivastava A., Geroliminis, N., 2013. Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model. Transport. Res. Part C, 30, 161-177.
Sun, D., Elefteriadou, L., 2010. Research and implementation of lane-changing model based on driver behavior. Transport. Res. Rec., 2161, 1-10.
Sumalee, A., Zhong, R., Pan, T., and Szeto, W., 2011. Stochastic cell transmission model (SCTM): A stochastic dynamic traffic model for traffic state surveillance and assignment, Transport. Res. Part B, 45 (3), 507-533.
Szeto, W.Y., 2008. The enhanced lagged cell transmission model for dynamic traffic assignment. Transport. Res. Rec., 2085, 76-85.
Tang, T., Wong, S., Huang, H., Zhang, P., 2009. Macroscopic modeling of lane-changing for two-lane traffic flow. J. Adv. Transp. 43 (3), 245-273.
Toledo, T., Choudhury, C., Ben-Akiva, M., 2005. A lane-changing model with explicit target lane choice. Transport. Res. Rec., 1934, 157-165.
Wang, C., Coifman, B., 2008. The effect of lane-change maneuvers on a simplified car-following theory. IEEE Trans. Intell. Transp. Syst., 9 (3), 523-535.
Yang, Q., Koutsopoulos, H., 1996. A microscopic traffic simulator for evaluation of dynamic traffic management systems. Transport. Res. Part C, 4 (3), 113-129.
Zheng, Z., 2014. Recent developments and research needs in modeling lane changing. Transport. Res. Part B, 60 (1), 16-32.
Zheng, Z., Ahn, S., Chen, D., Laval, J., 2013. The effects of lane-changing on the immediate follower: Anticipation, relaxation, and change in driver characteristics. Transport. Res. Part C, 26, 367-379.
Zhong, R., Chen, C., Andy, C., Pan, T., Yuan, F., He, Z., 2015. Automatic calibration of fundamental diagram for first-order macroscopic freeway traffic models. J. Adv. Transport., 50 (3), 363-385.