Transportation Research Part C 67 (2016) 47–61

Contents lists available at ScienceDirect

Transportation Research Part C

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

A binary decision model for discretionary lane changing move based on fuzzy inference system

Esmaeil Balal [a], Ruey Long Cheu [a,*], Thompson Sarkodie-Gyan [b]

[a] Department of Civil Engineering, The University of Texas at El Paso, 500 W. University Ave, El Paso, TX 79968, USA
[b] Department of Electrical & Computer Engineering, The University of Texas at El Paso, 500 W. University Ave, El Paso, TX 79968, USA

ARTICLE INFO

Article history:
Received 23 April 2015
Received in revised form 12 February 2016
Accepted 16 February 2016

Keywords:
Lane change
Fuzzy logic
Fuzzy inference system
Gap
Distance
Time-to-collision

ABSTRACT

This paper presents a Fuzzy Inference System (FIS) which models a driver's binary decision to or not to execute a discretionary lane changing move on freeways. It answers the following question "Is it time to begin to move into the target lane?" after the driver has decided to change lane and have selected the target lane. The system uses four input variables: the gap between the subject vehicle and the preceding vehicle in the original lane, the gap between the subject vehicle and the preceding vehicle in the target lane, the gap between the subject vehicle and the following vehicle in the target lane, and the distance between the preceding and following vehicles in the target lanes. The input variables were selected based on the outcomes of a drivers survey, and can be measured by sensors instrumented in the subject vehicle. The FIS was trained with Next Generation SIMulation (NGSIM) vehicle trajectory data collected at the I-80 Freeway in Emeryville, California, and then tested with data collected at the U.S. Highway 101 in Los Angeles, California. The results of the test have shown that the system made lane change recommendations of "yes, change lane" with 82.2% accuracy, and "no, do not change lane" with 99.5% accuracy. These accuracies are better than the same performance measures given by the TRANSMODELER's gap acceptance model for discretionary lane change on freeways, which is also calibrated with NGSIM data. The developed FIS has a potential to be implemented in lane change advisory systems, in autonomous vehicles, as well as microscopic traffic simulation tools.

1. Introduction

Lane changing model is as important as car-following model that govern the second-to-second motion of vehicles in microscopic traffic simulation tools (FHWA, 1995; PTV, 2007; Quadstone, 2009; TSS, 2002; Caliper, 2011). The microscopic driving behavior is also related to macroscopic property of traffic flow (Laval and Daganzo, 2006; Zhao et al., 2013; Zheng et al., 2013). Therefore, it can be said that both lane changing and car-following models are the fundamental building blocks of traffic flow theory. In the advent of semi-autonomous and autonomous vehicles, the understanding and accurate modeling car-following and lane changing behavior, including the replication of driver's decisions, is critical to the safe operations of these vehicles and the surrounding traffic. Although car-following has been studied by researchers for more than 50 years, relatively fewer examinations on lane changing behavior have been made. This could be due to the facts that (i) a lane change involves two-dimensional motions; and (ii) there are relatively more (up to five) vehicles involved in a lane changing

* Corresponding author.
E-mail addresses: ebalalvarnosfaderani@miners.utep.edu (E. Balal), rcheu@utep.edu (R.L. Cheu), tsarkodi@utep.edu (T. Sarkodie-Gyan).

http://dx.doi.org/10.1016/j.trc.2016.02.009

event. In contrast, car-following typically involves two vehicles, one following another in the same lane. Therefore, the study of lane change is more complex and challenging than car-following.

In general, there are two types of lane change in freeways: mandatory and discretionary. Mandatory lane change is also known as forced or necessary lane change. It usually occurs, in the United States driving convention, when a vehicle is trying to move from the left or center lane to the rightmost lane in order to exit the freeway. Mandatory lane change may also happen when a vehicle has just entered the freeway from an on-ramp and is trying to move to the center or left lane to avoid a downstream exit lane. Mandatory lane change is a microscopic manifestation of the macroscopic route choice behavior. Discretionary lane change is also known as free lane change or desired lane change. It primarily occurs when a driver seeks to increase its speed or seeks a better driving environment (such as greater space ahead or behind his/her vehicle) by moving to an adjacent lane (Caliper, 2011; Zheng, 2014). Obviously, the motivations and resulting driving behavior for the two types of lane change are different. Therefore, a driver is expected have different decision rules and/or risking taking behavior for the two types of lane change. Since driving is a complex task, it is sometimes not possible to classify a lane change as mandatory or discretionary. For example, a driver may move from the left lane to the right lane on a freeway well upstream of an exit. This is a mandatory move in order for him/her to exit the freeway. However, since the exit is far ahead, he/she has time to make a discretionary move. This may be the reason that some models combine the two types of lane change. This paper makes a distinction between mandatory and discretionary lane changes, i.e., for a vehicle that is changing lane, the event is either mandatory or discretionary, not both.

A lane change might be modeled as a sequential four-step process: (1) motivation; (2) selection of target lane; (3) checking for opportunity to move; and (4) the actual move. The beginning and end of the four steps are marked by time instants t_1, t_2, t_3, t_4 and t_5, respectively, where $t_1 < t_2 < t_3 < t_4 < t_5$. At t_1, the driver begins to feel uncomfortable driving in the original lane. Between t_1 and t_2, external stimulus motivates him/her to want to change lane. At t_2, he/she has made up his mind to change lane, and begins to look for a target lane (the immediate left or immediate right lane). At t_3, the target lane is selected. From t_3 onwards, the driver actively seeks an opportunity in the target lane to make a lateral move. He/she begins the lateral move at t_4. The lateral move is completed at t_5. Keyvan-Ekbatani et al. (2015) further classifies the lane changing move (i.e., between t_4 and t_5 into four sub-steps) but this is not the focus of this research. This article focuses on the decisions at t_3.

The traditional lane changing decision models rely on deterministic mathematical equations and/or rules to replicate drivers' decisions. These models do not consider the uncertainties of drivers' perception and decisions (McDonald et al., 1997; Das and Bowles, 1999). Fuzzy logic incorporates a degree of uncertainty in the decision making process and therefore, reflects the drivers' natural or subjective perceptions of the inputs which influence their decisions. Therefore, the fuzzy logic approach is used in this research to model the lane changing decisions from t_3 to t_4.

The objective of this paper is to develop an improved binary decision tool for a discretionary lane changing move using the fuzzy logic approach. More specifically, a Fuzzy Inference System (FIS) is constructed to replicate a driver's binary decisions in the third step of the four-step lane changing process; that is, from t_3 to t_4. The FIS answers the question "Is it time to begin to move into the target lane?" It is expected that the FIS will produce a series of "no, do not change lane" recommendations from t_3 and the last and one recommendation of "yes, change lane" at t_4. Once developed, the FIS may be programmed into lane change advisory systems in actual human-driven vehicles. It also has the potential to be programmed into autonomous vehicles, and microscopic traffic simulation models. The FIS has the potential to improve freeway safety by reducing the number of crashes due to incorrect lane changing decisions.

This paper is organized as follows. After this introduction, important terms that are used in the subsequent presentation of the lane changing problem are first defined. The envisioned application of the FIS in a lane change advisory system are described, so that readers can have a proper context our developmental work for the rest of this paper. Literature reviews on conventional lane changing models, fuzzy logic and FIS, and fuzzy logic based lane changing models are made. The next section reports a survey conducted to understand drivers' lane changing behavior. This is followed by a description of the Next Generation SIMulation (NGSIM) vehicle trajectory data used in this research, and the development of the FIS. The accuracy of the developed FIS when applied to the calibration and test data sets are then reported. The performance accuracy of the FIS is also compared against an existing gap acceptance model. This article ends by highlighting the contributions of this research, limitations and future research directions.

2. Definition of terms

A lane change involves the interaction of several vehicles. This section defines the vehicles involved, potential variables that may influence the driver's decision to or not to change lane (i.e., begin a lateral move) and their mathematical symbols, so that consistent terms and symbols can be used throughout this paper.

A typical lane changing scenario is depicted in Fig. 1, which involves up to five vehicles. The subject vehicle *S* is trying to move from its original lane to the target lane. Vehicles *PA*, *FA*, *PB* and *FB* are the preceding vehicle after lane change, following vehicle after lane change, preceding vehicle before lane change and following vehicle before lane change, respectively.

The variables that describe the interactions between these vehicles may be divided into three groups: (i) distance or gap (in distance unit); (ii) headway or time-to-collision (in time unit); and (iii) speed. Ten potential variables have been identified in a focus group meeting and they are listed in Table 1. The gap, distance and time-to-collision are calculated from the

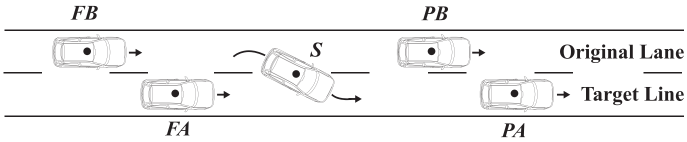

Fig. 1. Vehicles in a lane change.

Table 1
Variables that describe vehicle interactions in a lane change.

Notation	Definition	Unit	Range
G_{PB}	Gap between vehicle *S* and vehicle *PB*	m	$\geqslant 0$
G_{FB}	Gap between vehicle *S* and vehicle *FB*	m	$\geqslant 0$
G_{PA}	Gap between vehicle *S* and vehicle *PA*	m	$\geqslant 0$
G_{FA}	Gap between vehicle *S* and vehicle *FA*	m	$\geqslant 0$
D	Distance between vehicle *PA* and *FA*	m	$\geqslant 0$
T_{PB}	Time-to-collision between vehicle *S* and vehicle *PB*	s	$-\infty$ to $+\infty$
T_{FB}	Time-to-collision between vehicle *S* and vehicle *FB*	s	$-\infty$ to $+\infty$
T_{PA}	Time-to-collision between vehicle *S* and vehicle *PA*	s	$-\infty$ to $+\infty$
T_{FA}	Time-to-collision between vehicle *S* and vehicle *FA*	s	$-\infty$ to $+\infty$
V	Speed of vehicle *S*	m/s	$\geqslant 0$

rear bumper of the lead vehicle to the front bumper of the following vehicle. The gaps, and distance may be estimated by sensors embedded in the front, rear or corners of the subject vehicle *S*. The relative speeds may also be measured by these sensors. The time-to-collision may be estimated by dividing the gap or distance by relative speed.

3. Envisioned application

The FIS is designed and developed with a future application in a lane change advisory system in mind. The idea is inspired by the so-called "blind spot monitoring/warning system" that is already in existence in many luxury vehicle models. Such systems use sensors embedded the side or the rear of the subject vehicle to detect the presence of any vehicle or object in the driver's blind spots, and give visual and/or audible warnings.

The FIS is designed with the following considerations:

(a) It uses all or a subset of the variables listed in Table 1 as inputs. These variables may be measured directly and objectively by sensors instrumented in the subject vehicle, or calculated from two or more variables measured by these sensors. This avoids the need for the subject vehicle to receive data from the surrounding vehicles via vehicle-to-vehicle communication systems, and therefore making the stand-alone FIS implementable in the short term.

(b) Once the FIS has been developed and implemented as part of a lane change advisory system, the lane change advisory system will function as follows:
   - (i) The driver of the subject vehicle indicates his/her desire to change lane and the selected target lane by turning the vehicle's turn indicator (also known as turn signal).
   - (ii) The sensors in the subject vehicle estimate the subject vehicle's speed, the distances and relative speeds between itself and the surrounding vehicles, and use these values to compute the values of the input variables.
   - (iii) The input variables are fed into the FIS, and the FIS recommends a binary decision of "yes, change lane" or "no, do not change lane".
   - (iv) The recommendation is communicated to the driver by voice, audio signal and/or visual indicator in the instrument panel.
   - (v) The FIS will stop functioning when the turn indicator or turn signal is switched off by the driver. This may happen when the drive has completed the move or have changed his/her mind to remain in the original lane.

There are two approaches in expressing the outputs of a lane changing model: deterministic and probabilistic. The deterministic approach produces distinct and crisp outputs. An example of the deterministic output is the yes or no answer to our research question "Is it time to begin to move into the target lane?" Models that employ the probabilistic approach usually give a probability associated with each of the possible choices or outcomes. Examples of the probabilistic outputs are (i) the probabilities of a driver desiring to change lane, and do not wish to change lane; and (ii) the probabilities of a driver selecting the left lane and the right lane, respectively, as the target lane (Caliper, 2011). Clearly, the FIS should give a deterministic output of yes or no, so that the recommendation can be communicated clearly and quickly to the driver for him/her to make an immediate decision. It is relatively more time consuming to convey the different outcomes with the corresponding probabilities to the drivers, and even more difficult to explain their meanings to them. Therefore, for practical purpose, the FIS is designed to produce a crisp, deterministic, binary output of "yes, change lane" or "no, do not change lane".

4. Literature review

This section reviews primarily discretionary lane changing models with focuses on input variables and the decision making process. As comprehensive reviews of lane changing models have been made by Choudhury (2007), Moridpour et al. (2010) and more recently by Zheng (2014), who have devoted their entire articles for this purpose, it is not necessary to repeat the works here. Also, as stated in Zheng (2014), it is impossible to exhaustively review all the literature related to lane change. Therefore, this section only reports the works that fit the scope of this research as stated in Sections 1 and 3.

The review of lane changing models is a challenging process because in many articles, the authors do not completely describe the process, or provide the parameter values. This has made the understanding the logic difficult, and the implementation of these models impossible. A major finding of the literature is the wide variety of input variables used. These variables have been discussed and summarized in Balal et al. (2014), which led to the variables defined in Table 1 and the questions asked in the drivers survey (to be described in Section 5 of this article). In addition, lane changing models have been classified by researchers in different ways (Choudhury, 2007). Some models do not separate discretionary lane change from mandatory lane change. When describing a lane changing process, some models do not divide a lane changing event into the four-step process. Similarly, in some models it was impossible to distinguish the variables for decision to change lane, decision on target lane selection and decision to begin to move into the target lane.

In this review, discretionary lane changing models are classified into conventional models and fuzzy logic-based models. Section 4.1 reviews typical conventional models, in chronological order, that are relevant to the third step of the four-step discretionary lane changing process. Conventional lane changing models may further be classified into probabilistic and deterministic models. The discussion on these two types of models, and justification of using binary output for the FIS have already been explained in Section 3. Therefore the models are not divided into probabilistic and deterministic models in Section 4.1. Section 4.2 introduces fuzzy logic and FIS. This is followed by the review of existing fuzzy logic-based lane changing models.

4.1. Conventional lane changing models

Gipps (1986) and Hidas (2005) are two of the earliest documented and most frequently cited lane changing models. Gipps (1986) is perhaps one of the earliest to document a lane changing study in a signalized street. In Gipps' model, the decision making framework consists of the possibility, necessity and desirability to change lane. The input variables for discretionary lane change include V, G_{PA}, G_{FA}, and the relative speed between vehicles PA and FA (i.e., $V_{PA} - V_{FA}$). Gipps only demonstrated his model via computer simulation and did not explain how the variables were selected. Gipps model has been used in AIMSUN, a microscopic traffic simulation tool (TSS, 2002).

After observing video recordings of 73 lane changing maneuvers in arterials in Sydney, Australia, Hidas (2005) classified lane changes into free, forced and cooperative based on G_{PA} and G_{FA}.

Mar and Lin (2005) proposed a lane changing collision prevention system, in which the subject vehicle's controller starts to turn the steering wheel to begin a lane change if the following rule (that has four conditions) is satisfied:

$$\begin{aligned}
\text{IF}\quad & \{[(V_{PB} \leqslant V \leqslant V_{expected})] \text{ AND} \\
& [(V_{FA} \leqslant V) \text{ AND } (G_{FA} \geqslant G_{CF}^{safe})] \text{ AND} \\
& [(V_{PA} \geqslant V) \text{ AND } (G_{PA} \geqslant G_{CF}^{safe})] \text{ AND} \\
& [G_{PB} \geqslant G_{LC}^{safe}]\} \\
\text{THEN}\ & (\widetilde{C} \text{ is yes}) \\
\text{ELSE}\ & (\widetilde{C} \text{ is no})
\end{aligned}$$

In the above rule, V_{PB} is the speed of vehicle PB, $V_{expected}$ is the expected speed of the subject vehicle S, G_{PB} is the gap between vehicles PB and S, G_{CF}^{safe} is the safe car-following gap; and G_{LC}^{safe} is the safe gap for a lane change. It is possible to implement this decision rule as long as the parameter values of $V_{expected}$, G_{CF}^{safe} and G_{LC}^{safe} are known. This rule has only been tested in a simple computer simulation experiment with $V_{expected}$ = 90 km/h, $G_{CF}^{safe} = 37.6$ m and $G_{LC}^{safe} = 12.83$ m. This model did not distinguish between mandatory and discretionary lane changes.

Kesting et al. (2007) used the accelerations of vehicles S, FB, FA to form an incentive criterion for a lane change. Yeo et al. (2008) proposed a discretionary lane changing model for the subject vehicle S to increase speed or to improve its position in the traffic stream. The input variables are V, average speed of vehicles in the target lane and free-flow speed of vehicle S. Schakel et al. (2012) combined incentives to follow a route, to gain speed and to keep right into a single lane change desire value, from which three types of lane change (free, synchronized and cooperative) are distinguished. The lane changing model included seven parameters: relax headway, route desire, anticipated speed, speed desire, keep-right desire, combine desires, acceleration/deceleration and gap-acceptance. The model has been calibrated with loop detector data collected at the A20 Motorway near Rotterdam, Netherlands.

Recently, Hill and Elefteriadou (2013) studied the lane changing behavior of drivers in instrumented vehicles driving on I-4 Freeway in Orlando, Florida, and I-95 Freeway in Jacksonville, Florida. The times for lateral maneuver, desired speeds,

G_{PA} and G_{FA} were recorded for 321 discretionary lane changes and fitted with probability distributions. They have not yet developed a decision making model. This is the latest documented study which collected driver behavior data during lane changes.

Popular microscopic traffic simulation tools FRESIM (FHWA, 1995), VISSIM (PTV, 2007), PARAMICS (Quadstone, 2009), AIMSUN (TSS, 2002) and TRANSMODELER (Caliper, 2011), each has its own unique lane changing model that uses different input variables. These models have been reviewed by Balal et al. (2014).

The TRANSMODELER (Caliper, 2011), has embedded two sub-models to make discretionary lane changing decisions. The first sub-model, called "neighboring lane" model, uses the multinomial logit approach to calculate the probabilities of the subject vehicle selecting the current lane and the adjacent lanes. The second model, called gap acceptance model, is applied once a subject vehicle has selected the target lane. The gap acceptance model makes a binary decision that answers the question "is it safe to execute the lane change?" The Gap acceptance model's Recommendation, denoted by GR ={0, 1}, for {"no, do not change lane", "yes, change lane"}, is based on the comparisons of G_{FA}, and G_{PA} against their critical gaps, $G_{FA}^{\min}$ and $G_{PA}^{\min}$ respectively. The decision rule is

IF $[(G_{FA} \geqslant G_{FA}^{\min})$ AND $(G_{PA} \geqslant G_{PA}^{\min})]$
THEN (GR = 1 or "yes, change lane")
ELSE (GR = 0 or "no, do not change lane")

The critical gaps are calculated from

$$G_{PA}^{\min} = \exp[b_{0,PA} + b_{1,PA} \cdot \max(0, V - V_{PA}) + b_{2,PA} \cdot \min(0, V - V_{PA}) + b_{3,PA}V_{PA} + \alpha_{PA}\upsilon + \varepsilon_{PA}] \tag{1}$$

$$G_{FA}^{\min} = \exp[b_{0,FA} + b_{1,FA} \cdot \max(0, V_{FA} - V) + b_{2,FA}V_{FA} + \alpha_{FA}\upsilon + \varepsilon_{FA}] \tag{2}$$

in which V_{FA} and V_{PA} are the speed of vehicles FA and PA, respectively. The ε_{PA} and ε_{FA} terms are to account for the variability in aggressiveness between different drivers of the subject vehicles, with respect to the gaps between vehicles FA and PA, respectively. The model assumes that ε_{PA} and ε_{FA} follows normal distributions with mean 0 and variance σ_{PA}^2 and σ_{FA}^2 respectively, i.e., $\varepsilon_{PA} \sim N(0, \sigma_{PA}^2)$, $\varepsilon_{FA} \sim N(0, \sigma_{FA}^2)$. The υ term accounts for inconsistent decisions of the same driver at different time. It follows a truncated standard normal distribution $\upsilon \sim N(0,1)$, $-3 \leqslant \upsilon \leqslant 3$. The rest of the terms {$b_{0,PA}$, $b_{1,PA}$, $b_{2,PA}$, $b_{3,PA}$, α_{PA}, $b_{0,FA}$, $b_{1,FA}$, $b_{2,FA}$, α_{FA}} are coefficients. The above rule and equations were originally proposed and calibrated by Choudhury (2007), and later re-calibrated with NGSIM data (Caliper, 2011). The calibrated equations for the critical gaps, as coded in TRANSMODELER, are:

$$G_{PA}^{\min} = \exp[1 + 1.541 \cdot \max(0, V - V_{PA}) + 6.210 \cdot \min(0, V - V_{PA}) + 0.130V_{PA} - 0.008\upsilon + \varepsilon_{PA}] \tag{3}$$

$$G_{FA}^{\min} = \exp[1.50 + 1.426 \cdot \max(0, V_{FA} - V) + 0.640V_{FA} - 0.205\upsilon + \varepsilon_{FA}] \tag{4}$$

and $\varepsilon_{PA} \sim N(0, 0.854^2)$, $\varepsilon_{FA} \sim N(0, 0.954^2)$.

Most research developed and implemented lane changing and car-following models as two separate models. Wang et al. (2015) formulated an integrated lane changing and car-following control model using the dynamic game theory for potential implementation in autonomous vehicles and connected vehicles. In this model, the subject vehicle makes lane changing and acceleration/deceleration decisions (the control variables) based on the anticipated behavior of the surrounding vehicles (using the longitudinal positions and speeds and lateral positions of the vehicles as the input variables). The control objective is to minimize the predicted generalized cost over a horizon. They demonstrated the feasibility of the proposed model by conducting limited numerical simulation experiments for six scenarios that involved two to four vehicles in a two lane highway.

4.2. Fuzzy logic and fuzzy inference system

Fuzzy logic was introduced by Zadeh (1965) as a method to represent the imprecision in everyday life. Since then, fuzzy logic has emerged as a method for solving a wide variety of problems related to estimation, control, pattern recognition and decision making based on imprecise information. Fuzzy logic relies on several important concepts, of which fuzzy set, fuzzy membership and fuzzy rule are important components of FIS.

A fuzzy set defines several linguistic values that are used to describe a variable. For examples, the fuzzy set for G_{FA} may be defined as $\widetilde{G}_{FA}$ ={*close*, *medium*, *far*}, the fuzzy set for D may be defined as $\widetilde{D}$ ={*close*, *medium*, *far*}, and the fuzzy set for lane changing decision, C, may be defined as $\widetilde{C}$ ={*yes*, *no*}.

Fuzzy membership functions are used to map the crisp value of an input variable into the membership value (also known as degree of membership) for each linguistic value in the fuzzy set. In the above example, a crisp value of $G_{FA} = x$ is mapped by the respective membership functions, namely $\mu_{\widetilde{G}_{FA},close}(x)$, $\mu_{\widetilde{G}_{FA},medium}(x)$ and $\mu_{\widetilde{G}_{FA},far}(x)$ into [0,1]. Likewise, a crisp value of $D = y$ is mapped by $\mu_{\widetilde{D},close}(y)$, $\mu_{\widetilde{D},medium}(y)$ and $\mu_{\widetilde{D},far}(y)$ into their respective range of [0,1]. The membership functions for

$\widetilde{G}_{FA}$ and $\widetilde{D}$ may be defined as in Fig. 2. According to Fig. 2(b), for example, when $D = 20$ m, $\mu_{\widetilde{D},close}(20) = 0.5$, $\mu_{\widetilde{D},medium}(20) = 0.5$ and $\mu_{\widetilde{D},far}(20) = 0$. Fuzzy membership value is not the same as probability.

Fuzzy rules are normally expressed in the IF–THEN format. The antecedent of a rule may include more than one fuzzified variable, combined with logical operator AND or OR. A simple example of a rule which makes use of two fuzzified variables is

IF [($\widetilde{G}_{FA}$ is *close*) AND ($\widetilde{D}$ is *close*)] THEN ($\widetilde{C}$ is *no*)

This rule combines fuzzified inputs of G_{FA} and D to infer a fuzzified output of C. Mathematically, the membership values of the antecedent of the rule, $\mu_{\widetilde{G}_{FA},close}(x)$ and $\mu_{D,close}(x)$, are combined using the fuzzy set operator AND, which then fires the consequent of the rule to give an output value. There are several ways to mathematically calculate the fuzzified output of a rule. The two most commonly used methods are the Mamdani and Sugeno's fuzzy inference methods (Jang et al., 1997).

A FIS is a collection of fuzzy sets, fuzzy membership functions and fuzzy rules that are used to model a human decision making process. A typical FIS operation comprises four stages: fuzzification, inference, composition and defuzzification. The fuzzification stage converts crisp input values into membership values by means of fuzzy sets and fuzzy membership functions. The inference stage applies fuzzy rules to the fuzzified inputs and produces fuzzified outputs. In a FIS, there are multiple rules. The number of rules depends on the number of input variables and the number of linguistic values in the fuzzy set of each input variable. Using the above fuzzy rule as an example, the two variables in the antecedent have fuzzy sets of $\widetilde{G}_{FA}$ = {*close*, *medium*, *far*} and $\widetilde{D}$ ={*close*, *medium*, *far*}, respectively. Therefore nine rules are necessary to cover all the possible combinations of fuzzified values. The outcomes of the rules are combined in the composition stage, so that for the output variable, each linguistic value in its fuzzy set is assigned a membership value. Finally, in the defuzzification stage, the fuzzy set and membership values of the output variable are converted into a single crisp value. More details of FIS and its variants can be found in Jang et al. (1997).

4.3. Fuzzy logic-based lane changing models

Since 1999, there have been several applications of fuzzy logic or FIS in the modeling of lane changing decision process. These applications are reviewed in this section.

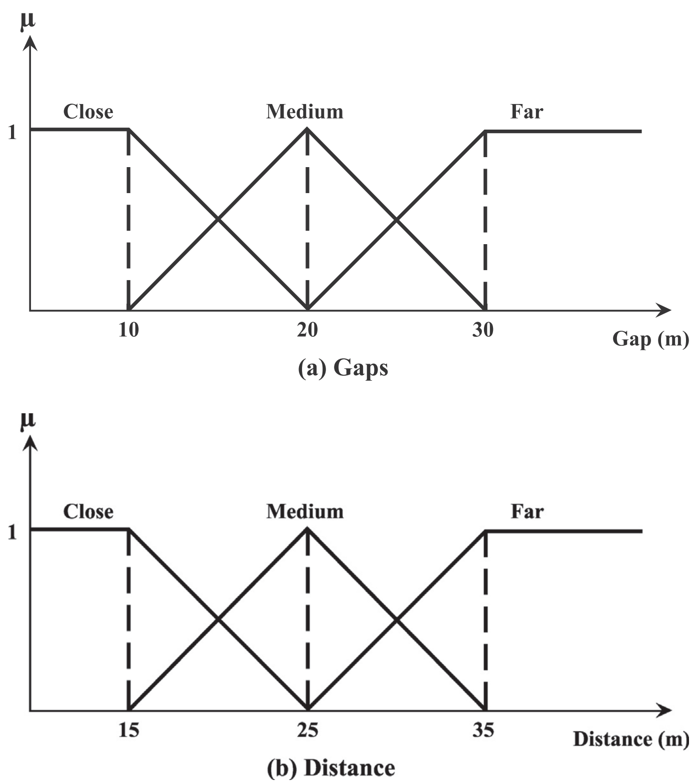

Fig. 2. Fuzzy membership functions for gap and distance.

Das and Bowles (1999) proposed a fuzzy logic-based lane changing model. The discretionary lane changing model relies on three inputs: *D*, vehicle speed in the target lane, and headway between vehicles *S* and *PB*. This model compares favorably with a set of vehicle trajectory data collected at Baltimore-Washington Parkway in 1980s, and with CORSIM simulation.

McDonald et al. (1997), Brackstone et al. (1998) and Wu et al. (2000) classified lane changes in motorways into: (i) lane changes to the near-side (shoulder) lane, mainly due to a fast-moving vehicle approaching from the rear; and (ii) lane changes to the off-side (median) lane to gain speed. For each category, they developed a fuzzy logic-based lane changing simulation model. The lane change to the near-side lane model uses two inputs: headway between vehicles *S* and *FB*; and time for the vehicle *S* to travel in its original speed in the near-side lane, without reducing its speed. The off-side lane-changing decision model has two crisp inputs: relative speed between the original lane and the target lane; and headway between vehicles *S* and *FA*. They programmed the lane changing models into a microscopic traffic simulation tool and compared the simulated lane changes against empirical data observed from four drivers. The difference between the number of lane changes predicted by their models and subject drivers was less than 11%.

Moridpour et al. (2009, 2012) proposed two fuzzy logic-based models that attempt to replicate heavy vehicle driver's motivation to change lane and selection of target lanes, i.e., from t_1 to t_3. One model was designed for lane change to slower lane while another one for lane change to faster lane. The inputs to the models were V, G_{PA}, G_{FB}, G_{FA}, relative speed between vehicles *S* and *PB*, and relative speed between vehicles *S* and *FA*. They trained and validated the models using NGSIM data collected at I-80 Freeway in Emeryville, California and U.S. Highway 101 in Los Angeles, California. The models achieved 92.8% accuracy for lane change to slower lane and up to 100% accuracy for lane change to faster lane.

More recently, Hou et al. (2012) proposed a genetic-fuzzy model for merging from on-ramp into freeway (one type of mandatory lane change). The inputs to the fuzzy model are V, G_{PA}, G_{FA}, relative speeds between vehicle *S* and *PA* and between *S* and *FA*. Genetic algorithm was used to calibrate the fuzzy membership functions. The model has 188 inference rules and produced 86.8% correct decisions when tested against on-ramp merging events in NGSIM's U.S. Highway 101 data.

4.4. Findings of literature review

The important findings from the literature review were (i) very few articles describe how the input variables for the lane changing models were selected; (ii) of the few which reported the variable selection process and the reasons of their selection, none was validated with feedback provided by drivers; and (iii) some of the variables (e.g., desired speed) cannot be estimated autonomously by sensors embedded in vehicles; they are related to the driver's psychology which render the model implementation difficult if not impossible; (iv) several models use relative speed as an input, which is not a direct indicator of risk compared to time-to-collision (which is distance divided by relative speed); (v) for most of the models, the computational steps or knowledge base necessary for the implementation are not clearly described. The challenges imposed by reasons (iii) and (v) have made the comparative evaluation of the model performances impossible, except for the gap acceptance model in TRANSMODELER (Caliper, 2011). In fact, none of the models reviewed has reported any comparative evaluation in their model development processes.

5. Drivers survey

The input variables used by the various lane changing models have been summarized and compared by Balal et al. (2014) and therefore they are not elaborated here. To make our FIS use the input variables as close to what drivers would use in real life, a questionnaire survey was conducted. The purpose of this survey was to select a few input variables most frequently used by drivers in making lane change decisions.

The survey instrument consisted of multiple choice questions concerning the respondent's motivation to make a discretionary lane change, and the variables listed in Table 1. For each of the variables, the respondent was asked to select if the variable was used all the time, most of the time, sometimes, seldom or never in making his/her lane changing decisions. Technical terms of lane change and variables are described in simple language, in both English and Spanish. The survey was administered to students, staff and faculty members on campus at The University of Texas at El Paso, drivers in local households and shopping malls from January to September 2014. A total of 443 useful responses were collected. The answers to the questions pertaining to the 10 variables were analyzed and are presented in Table 2.

Table 2 shows the percentage distribution of responses for each of the variables. The last (rightmost) column lists the percentage of the respondents who answered that they used each variable all the time or most of the time. From the tabulated results, it is obvious that gaps and distance are used more frequently than times-to-collision. This may be because it is easier for drivers to judge and estimate physical distance than time-to-collision. Of the 10 variables surveyed, G_{FA} is used all or most of the time by 94% of the respondents, followed by at D with 90%, G_{PA} at 88% and G_{PB} at 81%. These four variables (highlighted by bold in the Table 2) were therefore selected as the inputs to the FIS.

6. Vehicle trajectory data

The vehicle trajectory data used to develop the FIS was taken from the well-known NGSIM database. The NGSIM project is a data collection effort funded by Federal Highway Administration (FHWA) for the development and/or validation of new

Table 2
Results of drivers survey.

Input variables	Reported frequency of use						
	All the time (a) (%)	Most of the time (b) (%)	Sometimes (c) (%)	Seldom (d) (%)	Never (e) (%)	Total (a) + (b) + (c) + (d) + (e) (%)	All or most of the time (a) + (b) (%)
G_{PB}	56	25	13	4	2	100	**81**
G_{FB}	21	21	27	18	13	100	42
G_{PA}	61	27	9	2	1	100	**88**
G_{FA}	78	16	6	0	0	100	**94**
D	68	22	7	1	2	100	**90**
T_{PB}	7	15	29	22	27	100	22
T_{FB}	17	24	31	11	17	100	41
T_{PA}	21	28	21	11	19	100	49
T_{FA}	23	28	23	12	14	100	51
V	40	32	18	7	3	100	72

traffic models. In this research, the vehicle trajectory data collected at a segment of I-80 Freeway (Eisenhower Highway) in Emeryville, California (Cambridge, 2005a) and a segment of U.S. Highway 101 (Hollywood Freeway) in Los Angeles, California (Cambridge, 2005b) was used. The I-80 Freeway and U.S. Highway 101 sites and dates of data collection are the same as the NGSIM data used by Zheng et al. (2013).

The I-80 Freeway data was collected over a 1650 ft segment, in the northbound direction between the Powell Street on-ramp and Ashby Street off-ramp. This segment of the freeway has six lanes between the ramps. The available data was collected on April 13, 2005 from 4:00 to 4:15 p.m., 5:00 to 5:15 p.m. and 5:15 to 5:30 p.m. The data from 4:00 to 4:15 p.m. was used because it has the highest number of lane changes among the three 15-min periods.

The U.S. Highway 101 data was collected over a 2100 ft segment, in the southbound direction between the Ventura Boulevard on-ramp and Cahuenga Boulevard off-ramp. This segment of the freeway also has six lanes between the ramps. The available data was collected on June 15, 2005 from 7:50 to 8:05 a.m., 8:05–8:20 a.m. and 8:20–8:35 a.m. In this research, the data from 7:50 to 8:05 a.m. was used because it has the highest number of lane changes among the three 15–min periods.

For each of the freeway segments, vehicle motions were captured by several video cameras placed on top of a tall building. The video images were post-processed by the NGSIM project team to extract vehicle trajectory data at 0.1 s intervals, and make available to researches via the NGSIM website (Cambridge 2005a, 2005b).

The vehicle trajectory data was processed as follows:

- Only passenger cars were selected as the subject vehicles. Trucks and motorcycles, which were expected to have different lane changing behavior, and also have smaller sample sizes, were not considered.
- Only the subject vehicles originally traveled in lanes 2, 3 and 4 were used. Vehicles in lanes 5 and 6 were not considered so as to eliminate the possibility of drivers executing mandatory lane changes after entering from the upstream on-ramp or to exit at the downstream off-ramp. Similarly, subject vehicles in lane 1 were not considered as it is a high occupancy vehicle lane.
- Vehicles making multiple lane changes were excluded. This was because any lateral movement of more than one lane is more likely a mandatory move.
- For each identified S, the time t_4 was taken as the first instance when the front center of the S had lateral velocity of at least 0.2 m/s. This lateral velocity criteria is taken from Wang et al. (2014).
- Once t_4 has been determined, the positions of vehicles PB, FB, PA, FA that surrounded S were identified, and the input variables were calculated at $t_4 - 0.4$, $t_4 - 0.3$, $t_4 - 0.2$, $t_4 - 0.1$ and t_4 seconds respectively, according to the procedure recommended by Punzo et al. (2011). The average parameter values from $t_4 - 0.4$ to t_4 (five 0.1 s intervals) were used as the values perceived by the driver at t_4. The reasons for taking the average value over 0.5 s are (i) to reduce the error caused by using instantaneous values in the NGSIM data; (ii) to be more consistent with driver's perception time; and (iii) to be consistent with other research that has used NGSIM data, for example Siuhi and Kaseko (2010).
- The method of averaging data was repeated at 0.5 s intervals at, before and after t_4. Therefore, every S has multiple input vectors for the FIS at 0.5 s intervals.
- The Observed Maneuver (OM) was coded as $OM = 1$ for lane change at t_4, and $OM = 0$ for all other vectors. The observed maneuvers of $OM = \{0, 1\}$ were used as the "ground truth" to compare with FIS's recommendations when evaluating the FIS's performance.
- The above steps were repeated for passenger cars in lanes 2, 3 and 4 that did not change lane.

The statistics of the processed data are summarized in Table 3. In Dataset A, approximately 70% of the vectors from vehicles selected at random were used as calibration data. The remaining vectors in Dataset A was set aside as calibration-test data. Dataset B was reserved as test data.

7. Model development

This section describes the development of the FIS. The FIS uses the four input variables identified from the drivers survey. The fuzzy sets, membership functions and fuzzy rules are first described, followed by methods of compositions and defuzzification.

7.1. Fuzzy sets

From the drivers survey, the four decision variables selected were G_{FA}, D, G_{PA} and G_{PB}. The number of linguistic values in the fuzzy set for each variable affects the number of fuzzy rules in the FIS. To keep the number of fuzzy rules to a manageable level, a decision was made to have a fuzzy set of three linguistic values for each of the input variables, i.e., {*close*, *medium*, *far*}.

The FIS has only one output variable for lane change, denoted as C. The fuzzy set for the output variable is $\widetilde{C}$ = {*yes*, *no*}.

7.2. Fuzzy membership functions

Since there are four input variables and each of them has a fuzzy set of three linguistic values, $4 \times 3 = 12$ membership functions were necessary. The most popular triangular function was used for *medium*, while the trapezoidal functional form was used for *close* and *far*. The membership functions for G_{FA}, G_{PA} and G_{PB} are shown in Fig. 2(a) while those for D are shown in Fig. 2(b). The base and tip of the triangles and trapezoid were arbitrarily set at multiples of 5 m so as to approximate integer number of car length.

7.3. Fuzzy rules

Fuzzy rules were next written to infer fuzzified inputs to a fuzzified output. Given that each rule has four input variables, each variable has three linguistic values, the maximum number of rules was $3^4 = 81$. Two examples of the rules are:

IF [($\widetilde{G}_{FA}$ is *close*) AND ($\widetilde{G}_{PA}$ is *close*) AND ($\widetilde{D}$ is *close*) AND ($\widetilde{G}_{PB}$ is *close*)]
  THEN ($\widetilde{C}$ is *no*)
IF [($\widetilde{G}_{FA}$ is *close*) AND ($\widetilde{G}_{PA}$ is *far*) AND ($\widetilde{D}$ is *far*) AND ($\widetilde{G}_{PB}$ is *close*)]
  THEN ($\widetilde{C}$ is *yes*)

The numerical output of each rule is assign a binary value of C = [0, 1].

It has been mentioned that there could be up to 81 fuzzy rules. However, certain combinations of fuzzified inputs are infeasible. For example, it is impossible to have

[($\widetilde{G}_{FA}$ is *close*) AND ($\widetilde{G}_{PA}$ is *close*) AND ($\widetilde{D}$ is *far*) AND ($\widetilde{G}_{PB}$ is *close*)]

Because, based on the ranges of input values defined in the membership functions (see Fig. 2), when ($\widetilde{G}_{FA}$ is *close*) AND ($\widetilde{G}_{PA}$ is *close*), $\widetilde{D}$ cannot be *far*. After removing the infeasible rules, only 51 rules remained in the rule base. The 51 fuzzy rules are listed in Table 4.

7.4. Composition

Since there are 51 valid rules, and each rule is expected to produce a binary output of C = [0, 1], the purpose of composition is to combine the 51 binary output values into a single C^* value. The Mamdani fuzzy model first assigns the minimum membership values of the four fuzzified inputs as the output C for each rule. It then extract the maximum C values among the 51 rules and assign it to C^*. This is also known as the Mamdani composition method (Jang et al., 2007).

Table 3
Summary of Datasets A and B.

Dataset	A			B		
Source	I-80 Freeway April 13, 2005, 4:00–4:15 p.m.			U.S. Highway 101 June 15, 2005, 7:50–8:05 a.m.		
	Lane change	No lane change	Total	Lane change	No lane change	Total
No. of vehicles	**163**	**3202**	**3365**	**171**	**2612**	**2783**
Total no. of vectors	163	232,493	232,656	171	209,681	209,852
No. of vectors set aside for calibration	139	162,721	162,860	145	146,777	146,922

Table 4
Fuzzy rules.

Rule no.	IF				THEN
	$\widetilde{G}_{FA}$	$\widetilde{G}_{PA}$	$\widetilde{D}$	$\widetilde{G}_{PB}$	$\widetilde{C}$
1	Close	Close	Close	Close	No
2	Close	Medium	Close	Close	Yes
3	Close	Close	Medium	Close	Yes
4	Close	Close	Far	Close	Yes
5	Close	Close	Close	Medium	No
6	Close	Close	Close	Far	No
7	Close	Medium	Medium	Medium	Yes
8	Close	Far	Far	Far	Yes
9	Close	Close	Medium	Medium	No
10	Close	Close	Medium	Far	Yes
11	Close	Close	Far	Medium	Yes
12	Close	Medium	Close	Medium	No
13	Close	Medium	Far	Close	No
14	Close	Medium	Far	Medium	Yes
15	Close	Medium	Far	Far	Yes
16	Close	Medium	Medium	Close	No
17	Close	Far	Far	Close	Yes
18	Close	Medium	Medium	Far	No
19	Close	Medium	Close	Far	No
20	Close	Medium	Medium	Far	Yes
21	Close	Far	Far	Medium	Yes
22	Close	Far	Medium	Far	Yes
23	Medium	Medium	Medium	Medium	Yes
24	Medium	Close	Medium	Medium	Yes
25	Medium	Medium	Far	Medium	Yes
26	Medium	Medium	Medium	Far	Yes
27	Medium	Medium	Medium	Close	Yes
28	Medium	Far	Far	Far	Yes
29	Medium	Medium	Far	Far	Yes
30	Medium	Close	Medium	Close	Yes
31	Medium	Far	Far	Medium	Yes
32	Medium	Medium	Far	Close	Yes
33	Medium	Close	Far	Close	No
34	Medium	Close	Far	Medium	Yes
35	Medium	Close	Medium	Far	Yes
36	Medium	Close	Close	Far	No
37	Medium	Far	Far	Close	Yes
38	Medium	Close	Close	Close	No
39	Far	Far	Far	Far	Yes
40	Far	Medium	Medium	Medium	Yes
41	Far	Close	Far	Far	Yes
42	Far	Medium	Far	Far	Yes
43	Far	Far	Medium	Far	Yes
44	Far	Far	Far	Close	Yes
45	Far	Far	Far	Medium	Yes
46	Far	Far	Medium	Medium	Yes
47	Far	Close	Far	Close	Yes
48	Far	Medium	Far	Medium	Yes
49	Far	Medium	Far	Close	Yes
50	Far	Close	Medium	Close	Yes
51	Far	Close	Medium	Medium	Yes

7.5. Defuzzification

The purpose of defuzzification is to convert C^* to a binary decision or recommendation of "yes, change lane" or "no, do not change lane". This is achieved by comparing C^* against a threshold value τ, to come out with a FIS's Recommendation *FR* which has a crisp binary value of {0, 1}:

$$FR = \begin{cases} 1 & \text{for "yes, change lane"} \quad \text{if } C^* \geqslant \tau \\ 0 & \text{for "no, do not change lane"} \quad \text{if } C^* < \tau \end{cases} \tag{5}$$

8. Calibration

The proposed FIS was implemented in MATLAB's Fuzzy Logic Designer app (MathWorks, 2014). The FIS was initially calibrated with the calibration vectors in Dataset A to determine an appropriate τ value.

To help to select the τ value, the calibration vectors of Dataset A was presented to the FIS. Fig. 3(a) plots the frequency of $F(C^* > \tau|OM = 0)$, while Fig. 3(a) plots the frequency of $F(C^* < \tau|OM = 1)$. $F(C^* > \tau|OM = 0)$ is the number of vectors which have no observed lane change, but the FIS (with the given τ value) recommends a lane change ($FR = 1$). On the other hand, $F(C^* < \tau|OM = 1)$ is the number of training vectors which have observed lane changes, but the FIS (with the given τ value) did not recommend a lane change ($FR = 0$). Both situations represent errors made by the FIS. The optimal τ value should ideally minimize the total number of errors, i.e., minimize $F(C^* > \tau|OM = 0)+F(C^* < \tau|OM = 1)$. An alternative is to use the objective function:

$$\text{Minimize } \omega_1 F(C^* > \tau|OM = 0) + \omega_2 F(C^* < \tau|OM = 1) \tag{6}$$

where ω_1 and ω_2 are the expected cost of committing each type of error, respectively. Ideally, ω_1 is the probability of a collision (when the FIS recommends a lane change when it is not supposed to) multiplied by the average cost of a collision. ω_2 is the average delay cost of not changing lane in the next 0.5 s. ω_1 is expected to be very high compared to ω_2. However, from our data sets and from Fig. 3(a) and (b), it can be observed that $F(C^* > \tau|OM = 0)$ occurs much less frequently than $F(C^* < \tau|OM = 1)$. This may have compensating effect that makes $\omega_1 F(C^* > \tau|OM = 0)$ in approximately the same order of magnitude as $\omega_2 F(C^* < \tau|OM = 1)$. In the absence of the average costs of error, we arbitrarily relied on $F(C^* > \tau|OM = 0) + F(C^* < \tau|OM = 1)$, i.e., $\omega_1 = \omega_2 = 1$ to make a decision on the τ value. However, from Fig. 3(a) and (b), it is clear that the minimum $F(C^* > \tau|OM = 0) + F(C^* < \tau|OM = 1)$ occurs when $\tau \approx 0.1$, which is too small and therefore in practice undesirable. An alternate heuristic was employed to decide the τ value. First, we set $0.5 \leqslant \tau < 1$ because higher τ value will reduce the error of $FR = 1$ when in fact $OM = 0$. However, increase τ beyond 0.5 will not reduce $F(C^* > \tau|OM = 0)$ significantly (see Fig. 3(a)) but instead will increase $F(C^* < \tau|OM = 1)$ (see Fig. 3(b)). It was therefore decided that $\tau = 0.5$ be used for the purpose of subsequent test. If this FIS is eventually implemented in practice, the designer may select to set a different τ value.

9. Performance evaluation

Once the τ value was selected, the FIS was evaluated with the entire Dataset A. The FIS with the same τ value was then tested with Dataset B. The test with Dataset B serves as a transferability test, to see if the internal setting (including the τ

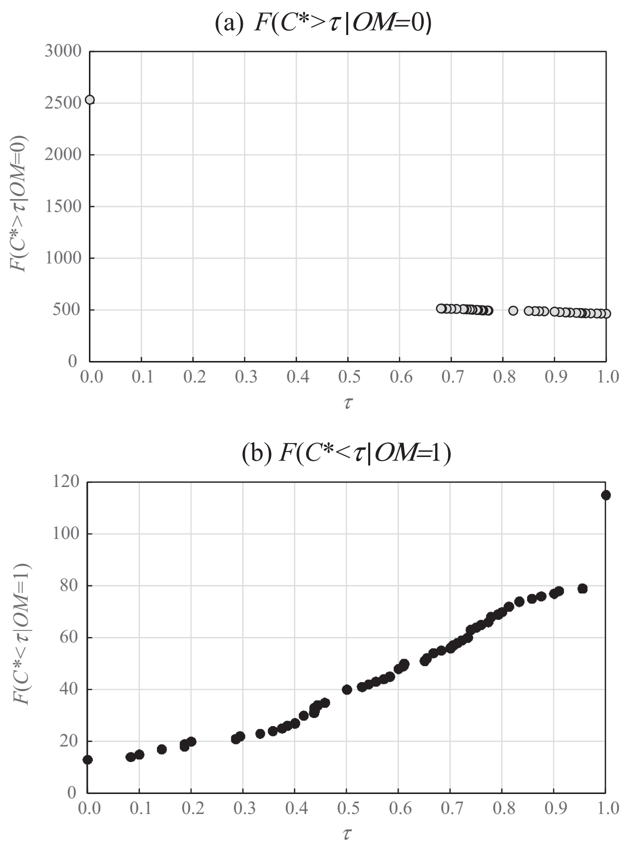

Fig. 3. Cumulative frequency distributions of C^* from training vectors.

value) of the FIS is sensitive to the driving behavior in a different city. The performances of the FIS and the gap acceptance model in TRANSMODELER was also compared to illustrate the superiority of this FIS.

9.1. Evaluation with Dataset A

The FIS, with $\tau = 0.5$, was evaluation using the entire Dataset A. The format of classification matrix in Moridpour et al. (2012) is modified to present the results in Table 4. In the matrix, the cell with [$FR = 1$, $OM = 1$] has the number of "yes, change lane" recommendations made by the FIS that also resulted in observed lane changes. The FIS made correct recommendations to 134 out of the 163 vectors at the onsets of lane changing moves, equivalent to an accuracy of 82.2%. The cell with [$FR = 0$, $OM = 0$] represents the number of "no, do not change lane" recommendations made by the FIS which coincided with no action by the drivers of the subject vehicles. Out of the 232,493 vectors which have $OM = 0$, the FIS made correct recommendations for 227,869 of them, or 98.0%. Two other cells, i.e., [$FR = 0$, $OM = 1$] and [$FR = 1$, $OM = 0$] are vectors with FIS outputs inconsistent with the observed behavior.

Despite the apparently high 98.0% accuracy for the vectors with $OM = 0$, there were still 4624 instances when the FIS wrongly recommended $FR = 1$. If the driver follow the recommendation, the lane change maneuver may potentially lead to a collision. Upon careful examination on these 4624 vectors, it was found that most of them happened at a fraction to a few seconds before t_4, the instant of an observed lane change. When Dataset A was set up, for each subject vehicle, OM was labeled as 1 only once at t_4 while the rest of the vectors for this vehicle had $OM = 0$. It was possible that the opportunity for a lane change presented itself 0.5 to a few seconds before t_4. However, due to perhaps the perception-reaction delay, or the conservative behavior of the driver (e.g., took time to double check the surrounding vehicles), there was no observable lateral move until t_4. Therefore, for the 163 subject vehicles that had changed lane, those vectors before t_4 which were labeled $OM = 0$ but the FIS recommended $FR = 1$ were considered as "no error" and retagged as $FR = 0$. The FR values, instead of OM values, were adjusted because it was not possible to change the observed maneuver. This resulted in 3955 vectors be moved from [$FR = 1$, $OM = 0$] cell to [$FR = 0$, $OM = 0$] cell. Table 6 presents the matrix after this reclassification of the FIS's outputs.

Comparing Table 5 with Table 6, the overall accuracy of FIS recommendations for $OM = 0$ has improved to 231,958 out of 232,656 vectors, or 99.5%. In summary, for Dataset A, the accuracy of FIS's recommendations for vectors at the onset of the observed lane changing move ($OM = 1$) is 82.2%. For vectors that belong to no lane change ($OM = 0$), the FIS's accuracy is between 98.0% and 99.5%. The actual accuracy may be somewhere in between these two values.

9.2. Evaluation with Dataset B

The FIS was then tested with Dataset B. Unlike Dataset A which was collected at I-80 Freeway in Emeryville, California, Dataset B was collected at U.S. Highway 101 in Los Angeles, California. Thus, the test with Dataset B served as a transferability test of the FIS, developed using one city's data, to another.

Prior to this test, the process of calibration was repeated for Dataset B (with 70% of randomly selected vectors, see the last row of Table 3) to see if it would give a different τ value, that is, the τ value is site dependent and there is a need to retrain τ. Graphs similar to Fig. 3(a) and (b) were plotted with the results of the FIS's application to the calibration vectors in Dataset B. Both curves showed similar trends and it was determined that $\tau = 0.5$ was still suitable. Therefore, it can be said that the FIS trained with Dataset A is transferable to Dataset B. The evaluation result with the entire Dataset B is presented in Table 7.

The initial classification outcomes with Dataset B resulted in an accuracy of 82.5% for vectors which belong to $OM = 1$. This accuracy is almost the same as that obtained with Dataset A (see Table 5). For vectors which belong to $OM = 0$, the initial classification accuracy was 97.0% which is marginally lower than the results given by Dataset A. After some of the vectors that belong to $OM = 0$ but the FIS's initial recommendations of $FR = 1$ immediately before t_4 were reclassified as $FR = 0$ (for the reason stated in Section 9.1), the revised result is presented in Table 7. The accuracy of FIS for vectors which belong to $OM = 0$ has improved to 99.5%, which is almost the same as the 99.7% obtained with Dataset A (see Table 5).

Comparing the results in Table 5 with Table 7, and Table 6 with Table 8, it can be concluded that the FIS is transferable between the two NGSIM data collection sites. Therefore, the FIS has the potential to achieve good performance in other cities.

Table 5
Initial classification matrix for Dataset A.

		FIS Recommendation FR			
		Yes, change lane $FR = 1$	No, do not change lane $FR = 0$	Total	Accuracy (%)
Observed Maneuver OM	Changed lane $OM = 1$	134	29	163	82.2
	Did not change lane $OM = 0$	4624	227,869	232,493	98.0
	Total	4758	227,898	232,656	

Table 6
Revised classification matrix for Dataset A.

		FIS Recommendation *FR*			
		Yes, change lane *FR* = 1	No, do not change lane *FR* = 0	Total	Accuracy (%)
Observed Maneuver *OM*	Changed lane *OM* = 1	134	29	163	82.2
	Did not change lane *OM* = 0	669	231,824	232,493	99.7
	Total	803	231,853	232,656	

Table 7
Initial classification matrix for Dataset B.

		FIS Recommendation *FR*			
		Yes, change lane *FR* = 1	No, do not change lane *FR* = 0	Total	Accuracy (%)
Observed Maneuver *OM*	Changed lane *OM* = 1	141	30	171	82.5
	Did not change lane *OM* = 0	6285	203,396	209,681	97.0
	Total	6426	203,426	209,852	

9.3. Comparative performance

This section compares the performance of the FIS, in terms of classification accuracy, against the performance of the gap acceptance model in TRANSMODELER. This gap acceptance model was selected for comparison because the decision rule and equations are clearly described, and the coefficients have been calibrated with NGSIM data (Caliper, 2011). The model proposed by Mar and Lin (2005) was not implemented because it was impossible to estimate the values of $V_{expected}$, G_{CF}^{safe} and G_{LC}^{safe} for individual subject vehicle's driver from the NGSIM data.

The gap acceptance model in TRANSMODELER has been described in Section 4. The details of the model calibration has not been documented in detail by Caliper (2011). However, Caliper (2011) states that the calibration process used the NGSIM data. These coefficients, in Eqs. (3) and (4), are very similar to the coefficients of the gap acceptance model calibrated by Choudhury (2007) that used the I-395 Freeway data in Arlington, Virginia. They are different from the coefficients calibrated with NGSIM's arterial data collected at Lankershim Boulevard in Los Angeles, California (Choudhury, 2007). The authors have confirmed with the original model developer (Choudhury, 2007) that the gap acceptance model in TRANSMODELER was initially calibrated using I-395 Freeway data and subsequently re-calibrated with NGSIM's I-80 Freeway data. Therefore, the decision rule, and Eqs. (3) and (4) are applied directly to Dataset B without further calibration. The gap acceptance model is also a binary decision model. Its recommendation is either "yes, change lane" ($GR = 1$) or "no, do not change lane" ($GR = 0$). Based on the classification outcomes, some vectors with $OM = 0$ but the gap acceptance model's initial recommendations of $GR = 1$ immediately before t_4 were reclassified as $GR = 0$, in the same fashion as described in Section 9.1. The results after reclassification are reported in Table 9. The accuracy of the gap acceptance model for vectors which belong to $OM = 1$ is only 58.5%, while that for vectors which belong to $OM = 0$ is only 66.7%. Compare to the results in Table 8 (which has 82.5% and 99.5% respectively), the FIS has much better accuracy. This means that the FIS makes recommendations on the lane changing move much closer to what is observed in Dataset B.

10. Summary, contributions, limitations and future research

This research has developed a FIS to advise the driver if he/she can begin a discretionary lane changing move to the adjacent target lane. The FIS was trained with Dataset A (collected at I-80 Freeway in Emeryville, California), and then tested with Dataset B (collected at U.S. Highway 101 in Los Angeles, California). The results obtained with both data sets are almost the same. This indicates that the FIS is transferable between the two cities. Based on the results of our test with Dataset B:

- At onsets of observed lane changing moves, the FIS made correct recommendations of "yes, change lane" among 82.2% of the test vectors.

Table 8
Revised classification matrix for Dataset B.

		FIS Recommendation *FR*			
		Yes, change lane *FR* = 1	No, do not change lane *FR* = 0	Total	Accuracy (%)
Observed Maneuver *OM*	Changed lane *OM* = 1	141	30	171	82.5
	Did not change lane *OM* = 0	1020	208,661	209,681	99.5
	Total	1161	208,691	209,852	

Table 9
Revised classification matrix for Dataset B, from the gap acceptance model.

		Gap acceptance model recommendation *GR*			
		Yes, change lane *GR* = 1	No, do not change lane *GR* = 0	Total	Accuracy (%)
Observed	Changed lane *OM* = 1	100	71	171	58.5
Maneuver *OM*	Did not change lane *OM* = 0	69,810	139,871	209,681	66.7
	Total	69,910	139,942	209,852	

- When the vehicles have no observable lateral movement, the FIS made correct recommendations of "no, do not change lane" between 97.0% and 99.5% of the test vectors.

The FIS model has achieved very encouraging results in the transferability test using Dataset B.

In addition, the FIS outperformed the existing TRANSMODELER's gap acceptance model (which is developed for discretionary lane change, and calibrated with NGSIM's freeway data). The FIS has better accuracies than this competitor in making "yes, change lane" and "no, do not change lane" recommendations.

This FIS takes four inputs variables most frequently used by surveyed drivers in making lane changing decisions. These variables may be estimated by sensors instrumented in the subject vehicle, avoiding the necessity of vehicle-to-vehicle communications. This FIS can be programmed as part of a stand-alone lane change advisory system. The U.S. Federal Highway Administration has estimated that between 8.4% and 13.7% of vehicle-to-vehicle collisions on highways occurred during merging or changing lanes (FHWA, 1996). The occurrence of collisions during lane changes may be reduced with the implementation of lane change advisory systems embedded with this FIS. This is the potential benefit of this research. This work also contributes to the field of lane changing research by:

- Surveyed drivers to ascertain the key variables they use in making decisions on lane changing move.
- Clearly explained the step by step process of implementing the FIS, so that this FIS may be used as a benchmark for comparison with new lane changing decision tools in the future.
- Demonstrated that our FIS has superior performance than the gap acceptance model in TRANSMODELER.

Although promising, there exist several limitations in the FIS which should be addressed in future research:

- The FIS developed so far is for passenger cars as the subject vehicles. Similar models, with different fuzzy membership functions and τ values may be developed for trucks, for mandatory lane change, and for arterial roads. For the distinctions between mandatory and discretionary lane changes, and between different road types, integrating the FIS with the vehicle's map matching/navigation system is necessary.
- Although the current version of FIS has very high accuracy, the model may further be improved by adjusting the bases and tips of the triangular and trapezoidal fuzzy membership functions, and/or by assigning different weights to the fuzzy rules.
- A more objective way could be developed to determine the τ value. One possibility is to estimate and include the ω_1 and ω_2 values in the objective function.
- Due to budget constraint, the drivers survey was conducted only in El Paso, Texas. The results of the survey may be biased toward the local behavior. In future, the survey should be expanded to cover other cities, especially the cities where the vehicle trajectory data was collected and used to calibrate and test the FIS.
- The FIS model was developed and tested with NGSIM data which is from moderate to heavy volume (1200–1600 vphpl, and 15–30 mph). The performance of the FIS in low volume, high speed traffic is not tested as data is not yet available.

The real test of the FIS is user acceptance of its recommendations, and the resulting safe maneuver during actual freeway driving. Therefore, conducting laboratory test (using a driving simulator) and field test (with an instrumented vehicle) with a sample of drivers should be two of the major tasks in future research.

Acknowledgements

The authors thank Dr. Qi Yang and Dr. Charisma Choudhury for their assistance in clarifying the calibration process of the TRANASMODELER's gap acceptance model that used NGSIM data.

References

Balal, E., Cheu, R.L., Gyan-Sarkodie, T., Miramontes, J., 2014. Analysis of discretionary lane changing parameters on freeways. Int. J. Transp. Sci. Technol. 3 (3), 277–296.

Brackstone, M., McDonald, M., Wu, J., 1998. Lane changing in the motorway: factors affecting its occurrence, and their implications. Road Transport Information and Control Conference, Publication No. 454, IEE, U.K., pp. 160–164.

Caliper, 2011. TransModeler User's Guide. Version 2.6, Caliper Corp. Newton, MA.

Cambridge, 2005a. NGSIM I-80 Data Analysis (4:00 p.m. to 4:15 p.m.) Summary Report. Cambridge Systematic Inc, Oakland, CA.
Cambridge, 2005b. NGSIM U.S. 101 Data Analysis (7:50 a.m. to 8:05 a.m.) Summary Report. Cambridge Systematic Inc, Oakland, CA.
Choudhury, C., 2007. Modeling Driving Decisions with Latent Plans. Ph.D. Thesis. Massachusetts Institute of Technology.
Das, S., Bowles, B.A., 1999. Simulation of highway chaos using fuzzy logic. In: Proceedings of the 18th International Conference of the North American Fuzzy Information Society (NAFIPS), New York, NY, pp. 130–133.
FHWA, 1995. The FRESIM User's Guide. Office of Traffic Safety and Operations R&D, Federal Highway Administration.
FHWA, 1996. Accident Analysis of Older Drivers on Freeways. Pub. No. FHWA-RD-96-035 <http://www.fhwa.dot.gov/publications/research/safety/humanfac/96035/> (accessed April 15, 2015).
Gipps, P.A., 1986. A model for the structure of lane-changing decisions. Transp. Res. 20B, 403–414.
Hidas, P., 2005. Modelling vehicle interactions in microscopic simulation of merging and weaving. Transp. Res. 13C, 37–63.
Hill, C., Elefteriadou, L., 2013. An exploration of lane changing behavior on freeways. Paper 13-2199, The 92nd Annual Meeting of the Transportation Research Board Compendium of Papers.
Hou, Y., Edara, P., Sun, C., 2012. A genetic fuzzy system for modeling mandatory lane changing. In: Proceedings of the 15th International IEEE Conference on Intelligent Transportation Systems, Anchorage, Alaska, pp. 1044–1048.
Jang, J.-S.R., Sun, C.-T., Mizutani, E., 1997. Neuro-Fuzzy and Soft Computing. Prentice Hall, Upper Saddle River, NJ.
Keyvan-Ekbatani, M., Knoop, V.L., Daamen, W., 2015. Categorization of the lane change decision process on freeways. Transp. Res. C (online first).
Kesting, A., Treiber, M., Helbing, D., 2007. General lane-changing model MOBIL for car-following models. Transp. Res. Rec. 1999, 86–94.
Laval, J., Daganzo, C., 2006. Lane-changing in traffic streams. Transp. Res. 40B, 251–264.
Mar, J., Lin, T.-H., 2005. The car-following and lane changing collision prevention system based on the cascaded fuzzy inference system. IEEE Trans. Veh. Technol. 54 (3), 910–924.
MathWorks, 2014. MATLAB Fuzzy Logic Toolbox Use'r Guide. Release 2014b, MathWorks, Inc., Natick, MA.
McDonald, M., Wu, J., Brackstone, M., 1997. Development of a fuzzy logic based microscopic motorway simulation model. In: Proceedings of the IEEE Conference on Intelligent Transportation Systems, Boston, MA, pp. 82–87.
Moridpour, S., Rose, G., Sarvi, M., 2009. Modelling the heavy vehicle drivers' lane changing decision under heavy traffic conditions. J. Road Transport Res. 18 (4), 49–57.
Moridpour, S., Sarvi, M., Rose, G., 2010. Lane changing models: a critical review. Transp. Lett. 2, 157–173.
Moridpour, S., Sarvi, M., Rose, G., Mazloumi, E., 2012. Land-changing decision model for heavy vehicle drivers. J. Intell. Transp. Syst. 16 (1), 24–35.
PTV, 2007. VISSIM 5.00 User Manual. PTV AG, Karlsruhe, Germany.
Punzo, V., Boizacchiello, M.T., Ciuffo, B., 2011. On the assessment of vehicle trajectory data accuracy and application to the Next Generation Simulation (NGSIM) program data. Transp. Res. 19C, 1243–1262.
Quadstone, 2009. Paramics User and Reference Manual, Version 6.6.1, Quadstone Ltd., Edinburgh.
Schakel, W.J., Knoop, V.L., van Arem, B., 2012. LMRS: an integrated lane change model with relaxation and synchronization. The 91st Annual Meeting of the Transportation Research Board Compendium of Papers.
Siuhi, S., Kaseko, M., 2010. Parametric study of stimulus-response behavior for car-following models. Paper 10-1179, The 89th Annual Meeting of the Transportation Research Board Compendium of Papers.
TSS, 2002. AIMSUN 4.1 User Manual. TSS Transport Simulation System, Barcelona, Spain.
Wang, M., Hoogendoorn, S., Daamen, W., van Arem, B., Happee, R., 2015. Game theoretic approach for predictive lane-changing and car-following control. Transp. Res. 58C, 73–93.
Wang, Q., Li, Z., Li, L., 2014. Investigation of discretionary lane-change characteristics using next-generation simulation data sets. J. Intell. Transp. Syst. 18, 246–253.
Wu, J., Brackstone, M., McDonald, M., 2000. Fuzzy sets and systems for a motorway microscopic simulation model. Fuzzy Sets Syst. 116, 65–76.
Yeo, H., Skabardonis, A., Halkis, J., Colyar, J., Alexiadis, V., 2008. Oversaturated freeway flow algorithm for use in Next Generation Simulation. Transp. Res. Rec. 2088, 68–79.
Zadeh, L.A., 1965. Fuzzy sets. Information. Control 8, 338–353.
Zhao, L., Sun, J., Zhang, H.M., 2013. Observations and analysis of multistep approaching lane changing behavior at two expressway merge bottlenecks in Shanghai. The 92nd Annual Meeting of the Transportation Research Board Compendium of Papers.
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. Transp. Res. 26C, 367–379.
Zheng, Z., 2014. Recent developments and research needs in modeling lane changing. Transp. Res. 60B, 16–32.