PAPER: 196

STATE DEPENDENCE IN LANE CHANGING MODELS

Charisma Choudhury[1], Moshe E. Ben-Akiva[1], Tomer Toledo[2], Anita Rao[1] and Gunwoo Lee[1]
[1]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology
[2]Faculty of Civil and Environmental Engineering, Technion - Israel Institute of Technology

Published in Transportation and Traffic Theory 2007, Allsop R., Bell M.G.H. and Heydecker B., eds., Elsevier Science, pp. 711-734, 2007

INTRODUCTION

Driving behaviour models are used within microscopic traffic simulations to predict driving manoeuvres. With the increasing popularity of such tools, there has been extensive research in improving the key driving behaviour models: acceleration, lane changing and route choice. Existing models usually assume that drivers react to current traffic conditions and make instantaneous decisions. However, in reality, drivers may plan a set of actions based on previous, current and anticipated future conditions and make a sequence of choices to execute the chosen plan. For example, a driver who has decided to change lanes but cannot do that immediately may continue to attempt to change lanes by selecting a target gap and adapting his acceleration to facilitate lane changing into that gap. The actions of the driver are thus implementations of the prior decision to change lanes and the decision tree is state dependent. However, in most cases the decision state of the driver (e.g. the decision to attempt to change lanes) is unobserved and only lane action and acceleration manoeuvres of the driver are observed.

In most of the existing driving behaviour models, the drivers are assumed to be myopic (Gipps 1986, Benekohal and Treiterer 1988, Yang and Koutsopoulos 1996, Zhang *et al.* 1998, Ahmed 1999, Choudhury 2005). A 'partial short-term plan' based decision framework for lane changing and acceleration was proposed by Toledo (2003), where the effects of a driver's short term plan to execute a lane change through a chosen gap is reflected on his acceleration decisions. However, state dependency has been ignored in this model and it is assumed that the driver revaluates his short term plan at every instant regardless of his current or previous state. Wang *et al*. (2005) tested the sensitivity of model parameters for a similar partial short-term based gap selection and acceleration model for freeway merging situation within a simulation framework.

The above mentioned state dependency is thus not captured in existing lane changing models. As a consequence, application of these models in micro-simulation tools often result in unrealistic traffic flow characteristics in congested and incident situations where the decisions of the driver involve significant planning, cooperation and risk taking.

This paper presents a framework for modelling state dependency in lane changing behaviour of drivers and demonstrates it through an on-ramp merging model for congested freeway situations. The proposed model explicitly considers the anticipation of future conditions as a basis of decision-making and incorporates state dependence to capture the effects of past decisions the driver has made on his current decision-making process. The paper is structured as follows: the structure of the state dependent merging model is described first. The estimation data and the estimation methodology are presented next followed by the estimation results. The improvements in the proposed model are demonstrated by statistical comparisons of the model against an instantaneous model that is estimated with the same dataset ignoring state dependency.

MODELING STATE DEPENDENCE IN FREEWAY MERGES

Model Framework

In congested situations, acceptable gaps are often not available and more complex merging phenomena are observed. For example, drivers may merge through courtesy of the lag driver in the target lane or become impatient and decide to force in, compelling the lag driver to slow down. The execution of all types of merges involve acceptance of available gaps. The definition of acceptable gaps may depend on the merging mechanism.

Normal merge occurs when the available adjacent gaps are immediately acceptable and is therefore an instantaneous process. However, in case of courtesy lane change and forced merge, even after the driver has initiated the merge, the actual lane change may not be possible immediately. A driver who has initiated a forced (or courtesy) merge remains in the initiated forced (or courtesy) merge state and continues to evaluate the adjacent gaps for the chosen merging mechanism until they are acceptable. Thus the gap acceptance decisions the driver makes at any instant depend on his state.

The deicision to select the merging mechanism is a sequential process. The decision framework of the driver is summarized in Figure 1. The model hypothesizes four levels of decision-making: normal gap acceptance, gap anticipation and aniticipated gap acceptance (or decision to initiate a courtesy merging), decision whether to initiate a forced merging or not, and gap acceptance for courtesy/forced merging. The decision process is latent and only the end action of the driver (lane change to the target lane) is observed. Latent choices are shown in ovals and observed actions are shown in rectangles.

The merging driver first compares the available lead and lag gaps in the mainline to the corresponding minimum acceptable gaps (critical gaps) for normal gap acceptance. Critical gaps are modeled as random variables, their means being functions of explanatory variables. If both the lead and the lag gaps are greater than the critical gaps, a lane change can be executed using the existing gaps.

If the gaps are not acceptable, the merging vehicle evaluates the speed, acceleration and relative position of the through vehicles and tries to evaluate whether or not the lag driver is providing courtesy to him. The courtesy or discourtesy of the lag driver is reflected on the anticipated gap. If the lag driver has decided to provide courtesy to a merging vehicle and has started to decelerate, the anticipated gap increases. The anticipated gap of a particular driver also depends on the length of the time horizon over which the gap is estimated. Differences in perception and planning abilities among drivers are captured by the distribution of the length of the time horizon. If the anticipated gap is acceptable, the merging driver perceives that he is receiving courtesy from the lag driver and initiates a courtesy merge.

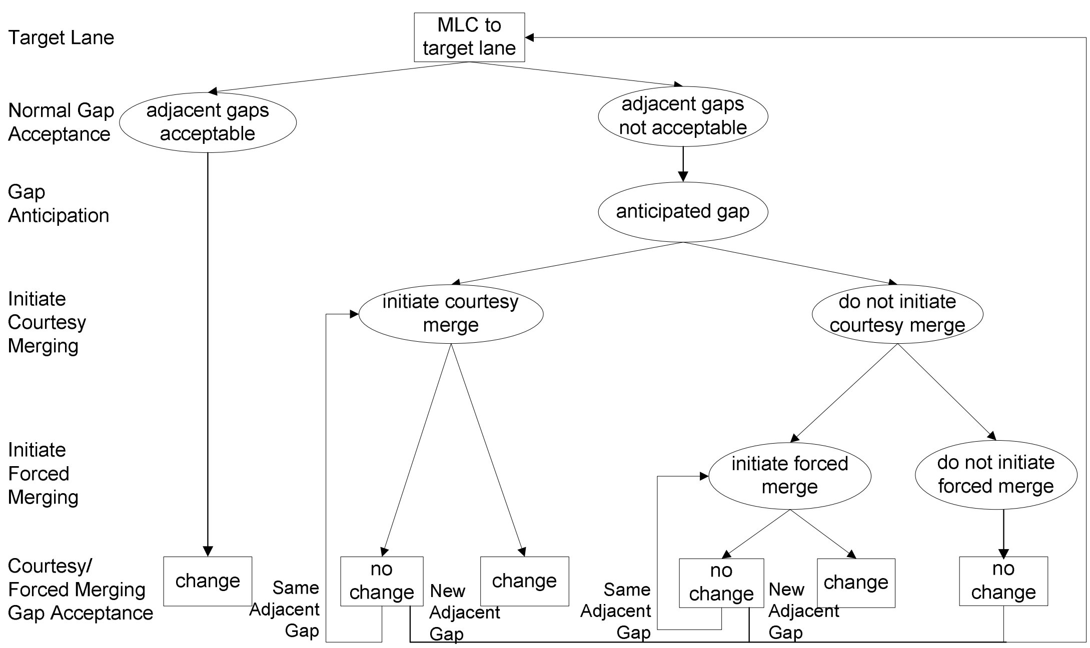


Figure 1: Framework of the merging model

If the anticipated gap is unacceptable, the driver decides whether to force the lag driver to slow down or not by nosing in. This decision can depend on the urgency of the merge, driver characteristics (e.g. risk averseness) and traffic conditions. If the driver does not initiate a courtesy or forced merge, he remains in the normal merging state and the entire decision process is repeated in the next time step.

A driver who has initiated a courtesy lane changing, compares the adjacent gaps against the courtesy merging critical gaps and makes the lane-change once these gaps are acceptable. The driver remains in the initiated courtesy merge state until the lane change is executed or the driver is adjacent to a new gap. Similarly, a driver who has initiated a forced merge remains in initiated forced merge state until the adjacent gaps are acceptable to execute the forced merge or the adjacent gap changes.

Therefore, the observed lane change (or no change) action at any instant is state dependent and can be the outcome of many possible decision sequences. Both the state of the driver and the decision sequence that led to the state are however unobserved/latent.

This paper focuses on formulation of the decision framework of the merging driver. The decisions of other drivers (e.g. decisions made by the lag driver whether or not to provide courtesy) are treated as external/observed variables in the model.

Model Components

Normal gap acceptance

Normal gap acceptance model indicates whether a lane change is possible or not using the existing gaps. The definition of related variables is presented in Figure 2. An available gap is acceptable if it is greater than the critical gap. Critical gaps are assumed to follow lognormal distributions, the mean gap being a function of explanatory variables. This can be expressed as follows:

$$\ln\left(G_{nt}^{Mg}\right) = \beta^{Mg^T} X_{nt} + \alpha^{Mg} \upsilon_n + \varepsilon_{nt}^{Mg} \quad g \in \{lead, lag\} \tag{1}$$

Where G_{nt}^{Mg} is the critical gap g of individual n at time t for normal gap acceptance (M), $g \in \{lead, lag\}$, X_{nt} are explanatory variables, β^{Mg} is the corresponding vector of parameters for normal gap acceptance, υ_n is the individual specific random effect: $\upsilon_n \sim N(0,1)$ and α^{Mg} is the coefficient of the individual specific random term for normal gap acceptance, ε_{nt}^{Mg} is the random term for normal gap acceptance of individual n at time t: $\varepsilon_{nt}^{Mg} \sim N\left(0, \sigma_{Mg}^2\right)$.

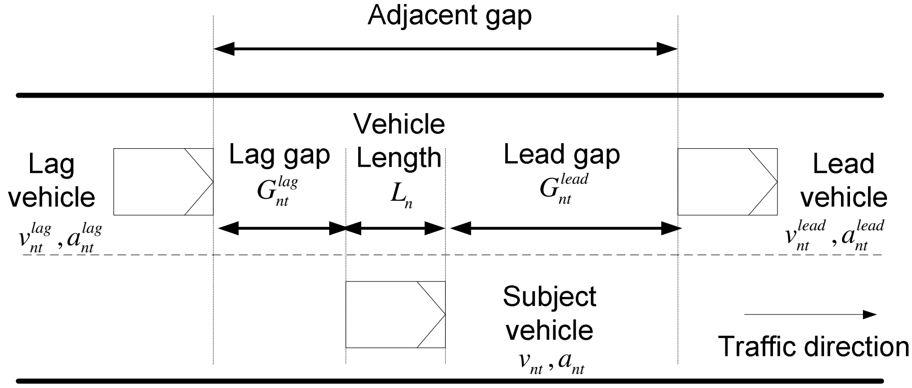


Figure 2: Vehicle relationships in a merging situation

The gap acceptance model assumes that the driver must accept both the lead and the lag gap to change lanes. If a merging vehicle is in normal state ($s_{t-1} = M$), i.e., he has not initiated a courtesy or forced merge, the probability of a lane change through normal gap acceptance, conditional on the individual specific term υ_n can be expressed as follows:

$$
\begin{aligned}
&P_n\left(l_t = 1 \middle| s_{t-1} = M, \upsilon_n\right) = \\
&P_n\left(accept\ lead\ gap \middle| s_{t-1} = M, \upsilon_n\right) P_n\left(accept\ lag\ gap | s_{t-1} = M, \upsilon_n\right) \\
&= P_n\left(G_{nt}^{lead} > G_{nt}^{M\,lead} \middle| s_{t-1} = M, \upsilon_n\right) P_n\left(G_{nt}^{lag} > G_{nt}^{M\,lag} \mid s_{t-1} = M, \upsilon_n\right)
\end{aligned}
\qquad (2)
$$

Where, for driver n at time t, l_t is the lane changing indicator, 1 if a lane-change is performed, 0 otherwise. s_t denotes state of the driver (M=normal, C=courtesy, F=forced), G_{nt}^{lead} and G_{nt}^{lag} are the available lead and lag gaps respectively.

Assuming that critical gaps follow lognormal distributions, the conditional probabilities that gap $g \in \{lead, lag\}$ is acceptable can be expressed as follows:

$$
\begin{aligned}
&P_n\left(l_t = 1 \middle| s_{t-1} = M, \upsilon_n\right) \\
&= P\left(\ln\left(G_{nt}^{g}\right) > \ln\left(G_{nt}^{Mg}\right) \mid s_{t-1} = M, \upsilon_n\right) \\
&= \Phi\left[\frac{\ln\left(G_{nt}^{g}\right) - \left(\beta^{Mg^T} X_{nt} + \alpha^{Mg}\upsilon_n\right)}{\sigma_{Mg}}\right]
\end{aligned}
\qquad (3)
$$

$\Phi[\cdot]$ denotes the cumulative standard normal distribution.

If a driver has already initiated a courtesy or forced merge in a previous time step, he cannot decide to merge to the same adjacent gap under normal gap acceptance. Therefore, if a merging vehicle is in initiated courtesy/forced merging state at time *(t-1)*, the probability of a lane change through normal gap acceptance at *t* is zero, unless there is a new adjacent gap.

Anticipated gaps and decision to initiate courtesy yielding

If the adjacent gaps are not acceptable to make a normal merge, the merging vehicle evaluates the speed, acceleration and relative position of the through vehicles and approximates an expected/anticipated gap that is going to open up after time τ_n. Because of the difference in perception among individuals, the anticipation time τ_n may vary among individuals.

The anticipated/expected gap for individual *n* at time *t* can be expressed as follows:

$$\overline{G}_{nt}(\tau_n) = G_{nt}^{lead} + G_{nt}^{lag} + L_n + \tau_n (v_{nt}^{lead} - v_{nt}^{lag}) + \frac{1}{2}\tau_n^2 (a_{nt}^{lead} - a_{nt}^{lag}) \tag{4}$$

Where, for individual *n* at time *t*, $\overline{G}_{nt}$ is the anticipated gap, L_n is the length of the vehicle, v_{nt}^{lead} and v_{nt}^{lag} are the speeds of the lead and lag vehicles*,* a_{nt}^{lead} and a_{nt}^{lag} are the acceleration of the lead and lag vehicles respectively (Figure 2).

If this anticipated gap is acceptable, the driver decides to initiate a courtesy merge. The critical gap of the driver for the anticipated gap acceptance is assumed to follow a lognormal distribution and can be expressed as follows:

$$\ln\left(G_{nt}^{A}\right) = \beta^{A^T} X_{nt} + \alpha^{A} \upsilon_n + \varepsilon_{nt}^{A} \tag{5}$$

Where, individual *n* at time *t*, G_{nt}^{A} is the critical gap for anticipated gap acceptance, β^{A} is the corresponding vector of parameters, ε_{nt}^{A} is the random term for anticipated gap acceptance: $\varepsilon_{nt}^{A} \sim N\left(0, \sigma_A^2\right)$.

If the driver has already initiated a courtesy merge in a previous time step and the adjacent gap has not changed, the probability of being in initiated courtesy merge state is 1. If the driver has already initiated a forced merge to the same gap in a previous time step, the probability of being in initiated courtesy merge state at current time step is 0. However, if the driver cannot complete the inititated courtesy merging within the time he is adjacent to the same gap and is adjacent to a new gap, the state of the driver is reset to the normal (not initiated courtesy or forced merging) state. This can be expressed as follows:

$$\begin{aligned}
&P_n\left(s_t = C \mid s_{t-1} = C, \upsilon_n, \tau_n\right) = \delta_{nt} + P_n\left(s_t = C \mid s_{t-1} = M, \upsilon_n, \tau_n\right)(1-\delta_{nt}) \\
&P_n\left(s_t = C \mid s_{t-1} = M, \upsilon_n, \tau_n\right) = P_n\left(\overline{G}_{nt} > G_{nt}^{A} \mid s_{t-1} = M, \upsilon_n, \tau_n\right)\left[1 - P_n\left(l_t = 1 \middle| s_{t-1} = M, \upsilon_n\right)\right] \\
&P_n\left(s_t = C \mid s_{t-1} = F, \upsilon_n, \tau_n\right) = 0
\end{aligned} \tag{6}$$

Where δ_{nt} =1 if driver *n* is adjacent to the same gap at time *(t-1)* and *t*, 0 otherwise.

The anticipation time is assumed to be truncated normally distributed with truncation on both sides. The distribution is given by:

$$f(\tau_n) = \begin{cases} \dfrac{\dfrac{1}{\sigma_\tau}\phi\left(\dfrac{\tau_n - \mu_\tau}{\sigma_\tau}\right)}{\Phi\left(\dfrac{\tau_{\max} - \mu_\tau}{\sigma_\tau}\right) - \Phi\left(\dfrac{\tau_{\min} - \mu_\tau}{\sigma_\tau}\right)} & \text{if } \tau_{\min} \le \tau_{\mathrm{n}} \le \tau_{\max} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

Where, μ_τ, σ_τ are the constant mean and standard deviations of the untruncated distribution, $\tau_{\min}$ and $\tau_{\max}$ are the minimum and maximum values of τ_n respectively. $\phi(\,)$ is the probability density function of a standard normal random variable and $\Phi(\,)$ is the cumulative distribution function of a standard normal random variable.

The advantage of using a truncated normal distribution is that it is not restricted to be skewed to a particular direction. This ensures that no a priori assumption is made on the probability of a driver being myopic or not.

Decision to initiate a forced merge

If the normal gaps are not acceptable and the driver perceives that he cannot merge through courtesy yielding (anticipated gap is not acceptable), he considers the decision whether to initiate forced merge $(s_t = F)$ or not $(s_t = M)$.

By initiating a forced merge, the merging driver takes a risk and imposes a deceleration on the lag vehicle in the mainline. The utility of initiating a forced merge can be expressed as follows:

$$U_{nt}^F = \beta^{F^T} X_{nt} + \alpha^F \upsilon_n + \varepsilon_{nt}^F \tag{8}$$

Where, for individual *n* at time *t*, U_{nt}^F is the utility of initiating a forced merge, β^F is the corresponding vector of parameters, ε_{nt}^F is the random term for initiating forced merging, α^F is the coefficient of the driver specific random term for forced merging.

By assuming that the random error terms ε_{nt}^F are i.i.d. Gumbel distributed, this decision can be modelled as a logit model.

Similar to the initiation of the courtesy merge, the probability of the driver being in initiated forced merge state is conditional on his previous state: the probability being 1 if the driver had already initiated a forced merge to the same gap in a previous time step and 0 if the driver had already initiated a courtesy merge to the same gap in a previous time step. However, if the driver cannot finish the initiated forced merging within the time he is adjacent to the same gap and is adjacent to a new gap, the state of the driver is reset to the normal (not initiated courtesy or forced merging) state.

$$P_n\left(s_t = F \mid s_{t-1} = F, \upsilon_n, \tau_n\right) = \delta_{nt} + P_n\left(s_t = F \mid s_{t-1} = M, \upsilon_n, \tau_n\right)\left(1-\delta_{nt}\right)$$
$$P_n\left(s_t = F \mid s_{t-1} = M, \upsilon_n, \tau_n\right)$$
$$= \left[\frac{1}{1+\exp\left(\left(-\beta^{F^T} X_{nt} - \alpha^F \upsilon_n\right) \mid \upsilon_n, \tau_n\right)}\right]\left[1 - P_n\left(s_t = C \middle| s_{t-1} = M, \upsilon_n\right)\right]\left[1 - P_n\left(l_t = 1 \middle| s_{t-1} = M, \upsilon_n\right)\right] \quad (9)$$
$$P_n\left(s_t = F \mid s_{t-1} = C, \upsilon_n, \tau_n\right) = 0$$

Where δ_{nt} =1 if the driver is adjacent to the same gap at time (t-1) and t, 0 otherwise.

Decision to make a courtesy/forced lane change

Even though a driver decides to initiate a courtesy/forced merge, the completion of the merge may take some time. That is, the actual merge is executed only when the available gaps are acceptable in comparison with the critical gaps for the respective merge. From the moment a driver initiates a forced merge up to T_n (the last time step the vehicle is observed as a merging vehicle), he is considered to be in initiated courtesy/forced merging state.

The functional form and variables influencing the critical gaps for courtesy and forced merging are assumed to be the same as in merging under normal gap acceptance, but the parameters are likely to be different.

State Transitions

At time *t* given an adjacent gap, driver *n*, can be in any one of the following states:

- Initiated courtesy merging ($s_t = C$)
- Initiated forced merging ($s_t = F$)
- Have not initiated courtesy/forced merging: normal ($s_t = M$)

Once a driver has initiated forced merging to an adjacent gap, he does not consider courtesy merging or normal gap acceptance in the subsequent time steps unless the gap changes. The decision in the subsequent time steps is only to decide whether or not to complete the forced merge in that time step. Thus once a transition is made from normal to forced merging state, the state cannot go back to normal and it cannot change to the initiated courtesy merging state unless the gap changes. Similarly, for a particular adjacent gap, once a transition is made from normal to initiated courtesy merging state, the state cannot change to initiated forced merging or normal. When the driver moves to a new adjacent gap, the state is reset to normal.

The possible decision state sequences are illustrated in Table 1 with two examples.

Table 1: Possible Decision State Sequences

Case 1: Same Adjacent Gap

Time Period	**Observed Lane**	**Lane Action**	**State Sequences**												
			1	2	3	…	T_n-1	T_n	T_n+1	T_n+2	T_n+3		$2T_n$-1	$2T_n$	$2T_n$+1
1	**CL**	**0**	C	M	M	…	M	M	F	M	M	…	M	M	M
2	**CL**	**0**	C	C	M	…	M	M	F	F	M	…	M	M	M
3	**CL**	**0**	C	C	C	…	M	M	F	F	F	…	M	M	M
⋮			⋮	⋮	⋮		⋮	⋮	⋮	⋮	⋮		⋮	⋮	⋮
T_n -1	**CL**	**0**	C	C	C	…	C	M	F	F	F	…	F	M	M
T_n	**CL**	**1**	C	C	C	…	C	C	F	F	F	…	F	F	M
T_n +1	**TL**														

Case 2: Two Adjacent Gaps

Gap	**Time Period**	**Observed Lane**	**Lane ACTION**	**State Sequences**										
				1	2	3	…	T_n-1	T_n	T_n+1	T_n+2	T_n+3		$2T_n$-1	$2T_n$
	1	**CL**	0	C	M	M	…	M	M	F	M	M	…	M	M
	2	**CL**	0	C	C	M	…	M	M	F	F	M	…	M	M
1	**3**	**CL**	0	C	C	C	…	M	M	F	F	F	…	M	M
	⋮			⋮	⋮	⋮		⋮	⋮	⋮	⋮	⋮		⋮	⋮
	T_n^1-1	**CL**	0	C	C	C	…	C	M	F	F	F	…	F	M
	T_n^1	**CL**	0	C	C	C	…	C	C	F	F	F	…	F	F
	1	**CL**	0	C	M	M	…	M	M	F	M	M	…	M	M
	2	**CL**	0	C	C	M	…	M	M	F	F	M	…	M	M
2	⋮			⋮	⋮	⋮		⋮	⋮	⋮	⋮	⋮		⋮	⋮
	T_n^2	**CL**	0	C	C	C	…	C	C	F	F	F	…	F	F
	T_n+1	**TL**													

C= Initiated courtesy merge $s_{nt} = C$, F= Initiated forced merge $s_{nt} = F$, M= Normal (Had not initiated a courtesy or forced merge), $s_{nt} = M$,CL= Current Lane, TL= Target Lane, 0=No change, 1=Change, P_n= Total number of adjacent gaps of individual n (2 in this case), T_n= Time individual n is observed as a merging vehicle, T_n^p = Time individual n is adjacent to gap p, $T_n = T_n^1 + T_n^2$ in this case.

As observed in the table, when the driver is adjacent to the same gap in two subsequent time instants, the following state transitions are possible:

- Normal to Normal ($s_t = M \mid s_{t-1} = M$)
- Normal to Courtesy ($s_t = C \mid s_{t-1} = M$)
- Normal to Forced ($s_t = F \mid s_{t-1} = M$)
- Courtesy to Courtesy ($s_t = C \mid s_{t-1} = C$)
- Forced to Forced ($s_t = F \mid s_{t-1} = F$)

When the driver is adjacent to a new gap, the following transitions are possible.

- Normal to Normal ($s_t = M \mid s_{t-1} = M$)
- Courtesy to Normal ($s_t = M \mid s_{t-1} = C$)
- Forced to Normal ($s_t = M \mid s_{t-1} = F$)

The probabilities of each of these transitions can be calculated using equations (6) and (9).

MODEL ESTIMATION

Data

The disaggregate data used for estimating the merging model was collected from the northbound direction of Interstate-80 (I-80) in Emeryville, California (Figure 3). The data was collected and processed as part of the Federal Highway Administration's Next Generation Simulation (NGSIM) project. Vehicles were tracked over a length of 503 meters (merging needs to be completed by 200 meters). The vehicle trajectory data containing the coordinates of the various vehicles in the section were used to derive the required variables for estimation. The merging drivers entering from the on-ramp to the rightmost lane of the mainline were used for estimation. The resulting dataset included 17352 observations at a 1 second time resolution of 540 vehicles.

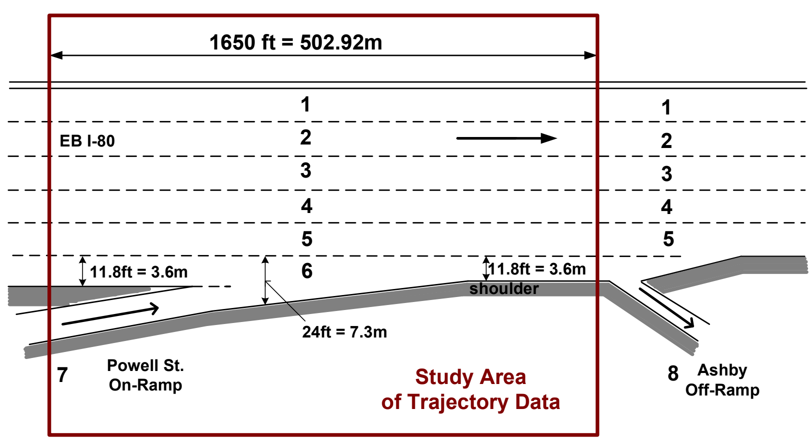


Figure 3: Data collection site

It may be noted that it was not possible to uniquely identify the state of the driver from the estimation data. For example, if there is an observation involving gap creation through deceleration of the lag vehicle, it is not difficult to determine whether it is the result of courtesy by the lag or the response to the merging vehicle forcing its way in. This motivated the latent choice formulation that has the flexibility to account for the various merge mechanisms without explicit knowledge of the mechanism that the driver has used.

Detailed analyses of the data and data processing methodology are presented in Cambridge Systematics (2005) and Choudhurry et al. (2006a).

Likelihood of the Trajectory

All model parameters were estimated jointly using a maximum likelihood technique. The likelihood function that was maximized is presented in this section.

At any time *t*, an individual can be in courtesy merging ($s_t = C$), forced merging ($s_t = F$) or normal ($s_t = M$) state. The lane changing decision of the driver depends on his state. The state of the driver at any instant depends on his previous state and the lane changing decision at that state.

According to the first-order Markov assumption: the probability of individual *n* being in a particular decision state *j* at time *t* only depends on his decision state at time *(t-1)*.

Therefore, the fact that a person is in state *j* at time *t*, where $t<T_n$, indicates the following:

- He has made a transition to state *j* from state *i* at t^{th} time step , where $i, j \in M, C, F$
- He was at state *i* at time *t-1*
- He has not made any lane change when he was at state *i* at time *t-1* (since the observation for an individual ends when he makes a lane change)

The probability of being in state $s_t = j$ is therefore the product of probability of a transition from state *i* to state *j* at time *t*, the probability of being in state *i* at time *(t-1)* and is conditional on the lane actions at previous time periods. This can be expressed as follows:

$$P_n(s_t = j \mid l_{t-1}, \upsilon_n, \tau_n) = \sum_i [P_n(s_t = j \mid s_{t-1} = i, l_{t-1}, \upsilon_n, \tau_n) P_n(s_{t-1} = i \mid l_{t-2}, \upsilon_n, \tau_n)] \quad (10)$$
$$i, j \in M, C, F$$

It may be noted that $P_n(s_t = j)$ is thus the sum of probabilities of all possible paths to $s_t = j$ (Figure 4).

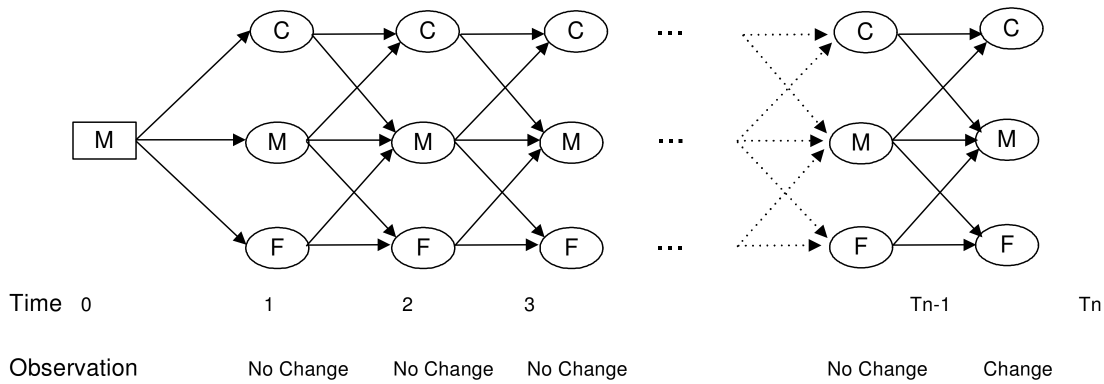


Figure 4: Decision state sequences

The state of the driver can thus be calculated recursively, the state of the driver at time t = 0 (when the driver first approaches the merging section) being normal.

Probability that at time *t* driver *n* executes lane changing decision l_t at state s_t is given by:

$$P_n(l_t, s_t \mid l_{t-1}, \upsilon_n, \tau_n) = P_n(l_t \mid s_t, \upsilon_n, \tau_n) P_n(s_t \mid l_{t-1}, \upsilon_n, \tau_n) \tag{11}$$

The state of the driver is not observed and only the lane changing actions are observed. Therefore, probability that driver *n* executes lane changing decision l_t at time *t* is given by:

$$P_n(l_t \mid l_{t-1}, \upsilon_n, \tau_n) = \sum_j P_n(l_t, s_t = j \mid l_{t-1}, \upsilon_n, \tau_n) \tag{12}$$

If driver *n* is observed over a sequence of T_n consecutive time intervals, the probability of observing his entire trajectory is the product of the probabilities given in equation (12) and can be expressed as:

$$P_n(\mathbf{l} \mid \upsilon_n, \tau_n) = \prod_{t=1}^{T_n} P_n(l_t \mid l_{t-1}, \upsilon_n, \tau_n) \tag{13}$$

The unconditional individual likelihood is given by:

$$L_n = \int_\upsilon \int_\tau P_n(\mathbf{l} \mid \upsilon_n, \tau_n) f(\upsilon) f(\tau)\, d\upsilon\, d\tau \tag{14}$$

Where,

$f(\upsilon)$ is the standard normal probability density function, $f(\tau)$ is the probability density function of a doubly truncated normal distribution with mean μ_τ and variance σ_τ^2.

Maximum likelihood estimators of the model parameters can be found by maximizing this function.

Estimation Results

The estimation results estimated using the statistical estimation software Gauss7.0 are summarized in Table 2. The final log-likelihood is -1609.65 and the adjusted rho-bar square is 0.88.

Table 2. Estimation Results of State Dependence Model

Variable		Parameter Value	t-statistic
Normal Lead Gap			
Normal lead constant		-0.230	-0.33
*Relative average speed (positive) (m/sec)		0.521	0.81
*Relative lead speed (m/sec)		-0.505	-3.13
*Remaining distance function	Relative lead speed (negative) (m/sec)	1.32	3.64
	Remaining distance to MLC point (10 m)	0.420	0.89
	Remaining distance constant	0.355	1.68
σ^{MLead}		3.42	9.67
α^{MLead}		-0.819	-3.12
Normal Lag Gap			
Normal lag constant		0.198	2.87
*Relative lag speed (positive) (m/sec)		0.208	1.78
*Relative lag speed (negative) (m/sec)		0.184	1.63
*Remaining distance function	Remaining distance to MLC point (10 m)	0.439	5.09
	Remaining distance constant	0.0242	0.03
	$\alpha^{\mathrm{Re}mDisLag}$	0.000180	0.03
*Lag acceleration (positive) (m/sec^2)		0.0545	0.61
σ^{MLag}		0.840	3.03
α^{MLag}		-0.0000776	-0.01
Initiate Courtesy Merge			
Anticipated gap constant		1.82	1.00
Relative average speed (positive) (m/sec)		1.82	2.13
Relative lead speed (m/sec)		-0.153	-0.97
Remaining distance function	Distance to MLC point (10 m)	0.244	1.50
	Constant	0.449	0.49
	$\alpha^{\mathrm{Re}mDisA}$	0.360	0.18
σ_A		0.0106	0.07

Variable		Parameter Value	t-statistic
α^{A}		-0.231	-1.90
μ_{τ}		1.87	9.51
σ_{τ}		1.44	17.71
Courtesy Lead Gap			
Courtesy lead constant		-0.582	-0.20
*Relative average speed (positive) (m/sec)		0.521	0.81
*Relative lead speed (negative) (m/sec)		-0.505	-3.13
*Remaining distance function	Distance to MLC point (10 m)	1.32	3.64
	Constant	0.420	0.89
	$\alpha^{\mathrm{Re}mDisLead}$	0.355	1.68
σ^{CLead}		0.0109	0.08
α^{CLead}		-0.0540	-0.03
Courtesy Lag Gap			
Courtesy lag constant		-1.23	-0.07
*Relative lag speed (positive) (m/sec)		0.208	1.78
*Relative lag speed (negative) (m/sec)		0.184	1.63
*Remaining distance function	Distance to MLC point (10 m)	0.439	5.09
	Constant	0.0242	0.03
	$\alpha^{\mathrm{Re}mDisLag}$	0.000180	0.03
*Lag acceleration (positive) (m/sec^2)		0.0545	0.61
σ_{CLag}		0.554	0.05
α^{CLag}		-0.0226	-0.04
Initiate Forced Merge			
Initiate force constant		-6.41	-4.63
Heavy lag vehicle dummy		-1.25	-0.63
α^{F}		5.43	3.26
Forced Lead Gap			
Forced lead constant		3.11	2.11
*Relative average speed (positive) (m/sec)		0.521	0.81
*Relative lead speed (m/sec)		-0.505	-3.13
*Remaining distance function	Distance to MLC point (10 m)	1.32	3.64
	Constant	0.420	0.89

Variable		Parameter Value	t-statistic
	$\alpha^{\mathrm{Re}mDisLead}$	0.355	1.68
σ^{FLead}		7.95	5.82
α^{FLead}		-0.0401	-0.07
Forced Lag Gap			
Forced lag constant		-2.53	-3.42
*Relative lag speed (positive) (m/sec)		0.208	1.78
*Relative lag speed (negative) (m/sec)		0.184	1.63
*Remaining distance function	Distance to MLC point (10 m)	0.439	5.09
	Constant	0.0242	0.03
	$\alpha^{\mathrm{Re}mDisLag}$	0.000180	0.03
*Lag acceleration (positive) (m/sec^2)		0.0545	0.61
σ^{FLag}		0.465	2.49
α^{FLag}		-0.0239	-0.19

* same coefficients in normal, courtesy and forced gap acceptance levels

The lead critical gap is a function of the average speed in the mainline relative to the subject vehicle's speed, the relative speed of the lead with respect to the subject and the remaining distance to the mandatory lane changing point. The lag critical gap is a function of the subject relative speed with respect to the lag vehicle, the remaining distance to the mandatory lane changing point and the acceleration of the lag vehicle.

The estimated lead and lag critical gaps for the normal gap acceptance are given by:

$$\begin{aligned} G_{nt}^{MLead} &= \exp\left(-0.230 + 0.521V_{nt}^{'} - 0.505Min\left(0, \Delta V_{nt}^{lead}\right) + \frac{1.32}{1+\exp(0.420+0.355\upsilon_n)} d_{nt} - 0.819\upsilon_n + \varepsilon_{nt}^{Mlead}\right) \\ G_{nt}^{MLag} &= \exp\left(\begin{aligned} &0.198 + 0.208Max\left(0, \Delta V_{nt}^{lag}\right) + 0.184Min\left(0, \Delta V_{nt}^{lag}\right) + \frac{0.439}{1+\exp(0.0242+0.00018\upsilon_n)} d_{nt} \\ &+0.0545Max\left(0, a_{nt}^{lag}\right) - 0.840\upsilon_n + \varepsilon_{nt}^{Mlag} \end{aligned}\right) \end{aligned} \tag{15}$$

Where, G_{nt}^{Mlead} is the lead critical gap for the normal gap acceptance level (m), G_{nt}^{Mlag} lag critical gap for the normal gap acceptance level (m), $V_{nt}^{'}$ is the relative average speed factor (m/sec), ΔV_{nt}^{lead} relative speed of the lead vehicle with respect to the subject (m/sec), d_{nt} is the remaining distance to the mandatory lane changing point (10m), ΔV_{nt}^{lag} relative speed of the lag vehicle with respect to the subject (m/sec), a_{nt}^{lag} acceleration of the lag vehicle, ε_{nt}^{Mlead} *and* ε_{nt}^{Mlag} are random error terms with $\varepsilon_{nt}^{Mlead} \sim N\left(0, 3.83^2\right)$ and $\varepsilon_{nt}^{Mlag} \sim N\left(0, 0.532^2\right)$.

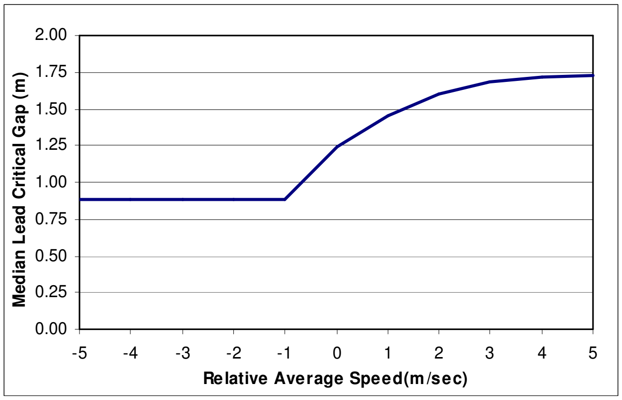
The lead critical gap increases with the increase in average speed of the mainline. As the mainline average speed increases, the driver needs larger critical gaps to adjust his speed to the speed of the mainstream. However, critical gap does not increase linearly with increasing average speeds in the mainline (Figure 5a), it rather increases as a diminishing function $\beta^{Mavg} V_{nt}^{'}$, where, $V_{nt}^{'} = \left(1 + \frac{1}{1 + \exp\left(-Max\left(0, \Delta V_{nt}^{avg}\right)\right)}\right)$, ΔV_{nt}^{avg} being the relative speed of the average mainline speed with respect to the subject (m/sec).

The lead critical gap is larger when the lead vehicle is moving slower than the subject since the driver perceives an increased risk when the lead is slowing down and he gets closer to the lead vehicle (Figure 5b).

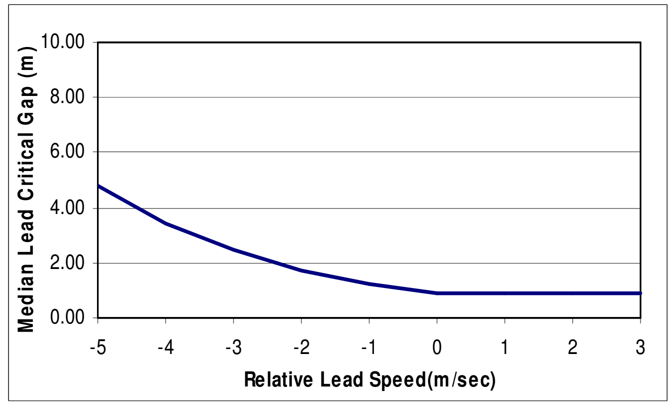
The lag critical gap increases with the relative lag speed: the faster the lag vehicle is relative to the subject, the larger the critical gap is (Figure 5c).The lag critical gap increases as the acceleration of the lag vehicle increases (Figure 5d), due to the higher perceived risk into merging onto the mainstream when the lag vehicle is accelerating.

Both the lead and lag critical gaps decrease as the remaining distance to the mandatory lane changing point decreases. This is because as the driver approaches the point where the ramp ends, his urgency to make the merge increases and he is willing to accept lower gaps to merge in to. To capture drivers' heterogeneity, an individual specific random term has been introduced in the coefficient of the remaining distance. Aggressive and timid drivers can thus have different critical gaps, the remaining distance being equal. The aggressiveness/timidness of the driver basically captures the heterogeneity among the driver population and is assumed to have a continuous distribution (truncated normal in this case) rather than discrete having a discrete class membership. For example, all other variables having no effect, the lead and lag critical gaps as a function of remaining distance for the aggressive drivers are much smaller than the gaps of timid drivers. Thus, aggressive drivers can find lead and lag gaps to be acceptable even when they are far from the MLC point. On the other hand, timid drivers have large critical gaps till they reach the end of the ramp, implying that they do not consider lane changes in the beginning of the on-ramp. The sensitivity of the lead and lag critical gaps as a function of the remaining distance according to the individual characteristics of the driver is shown in Figure 5e and Figure 5f respectively. The t-statistics for the linear part of the coefficient of remaining distance is found to be very significant both for lead and lag gaps.

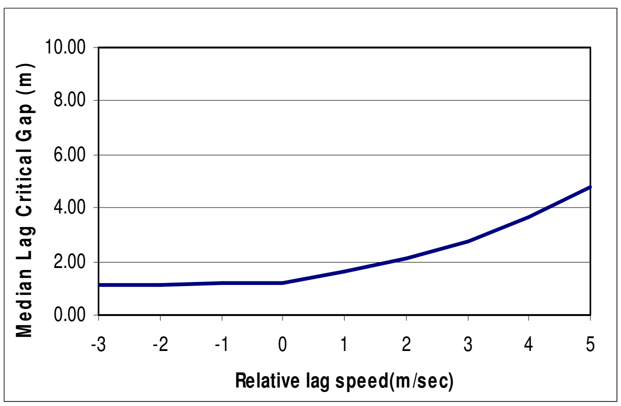
Estimated coefficients of the unobserved driver characteristics (υ_n) are negative for both the lead and lag critical gaps. This implies that an aggressive driver requires smaller gaps for lane changing as compared to a timid driver.



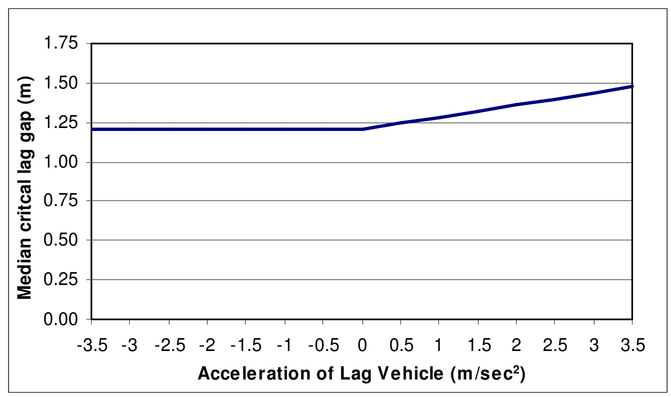
a. Lead critical gap as a function of relative average speed



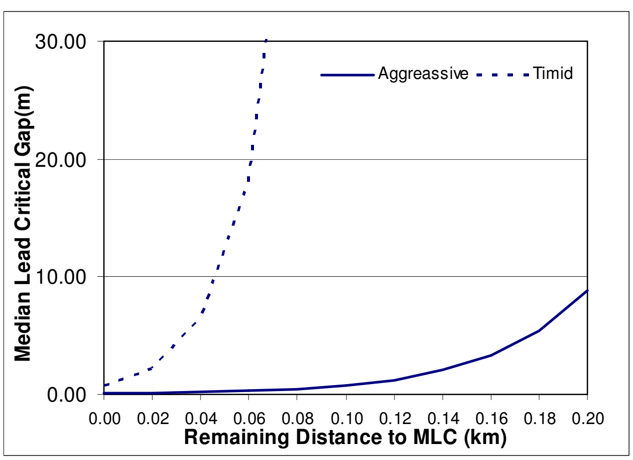
b. Lead critical gap as a function of relative lead speed



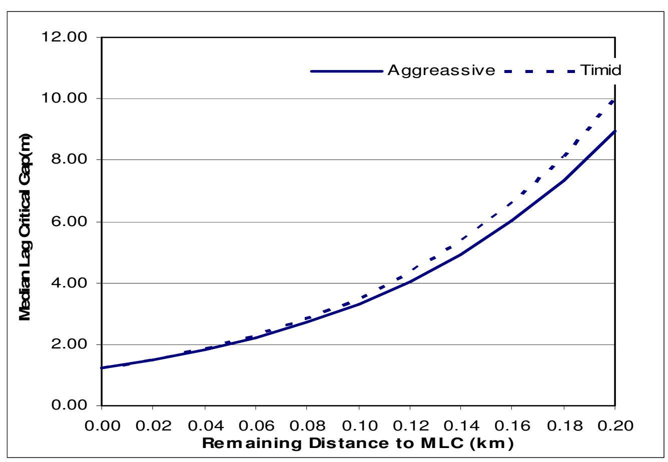
c. Lag critical gap as a function of relative lag speed



d. Lag critical gap as a function of lag vehicle acceleration



e. Lead critical gap as a function of remaining distance to MLC point



f. Lag critical gap as a function of remaining distance to MLC point

Figure 5- Median Lead and Lag gap variations

The anticipated gap acceptance (initiating courtesy) depends on lag speed, remaining distance and density of the traffic stream. The estimated critical anticipated gap is given by:

$$G_{nt}^{A} = \exp\left(1.82 + 1.81 Max\left(0, \Delta V_{nt}^{lag}\right) - 0.153\rho_{nt} + \frac{0.244}{1+\exp(0.449+0.360\upsilon_{n}}d_{nt} - 0.213\upsilon_{n} + \varepsilon_{nt}^{A}\right) \quad (16)$$

Where, G_{nt}^{A} is the critical anticipated gap for initiating courtesy merge (m), ρ_{nt} is the density in the rightmost lane of the mainline (veh/10m), $\varepsilon_{nt}^{A} \sim N\left(0, 0.0106^{2}\right)$

Similar to normal critical gaps, the critical anticipated gap is higher at higher lag speeds. It decreases as the remaining distance decreases and it is smaller for aggressive drivers as compared to timid drivers. Courtesy yielding/merging more commonly occurs in dense traffic conditions and hence the probability to merge through courtesy increases as the density of traffic in the mainline increases. The critical anticipated gap therefore reduces with density of traffic in the rightmost lane of the mainline. Median critical anticipated gap as a function of density is presented in Figure 6.

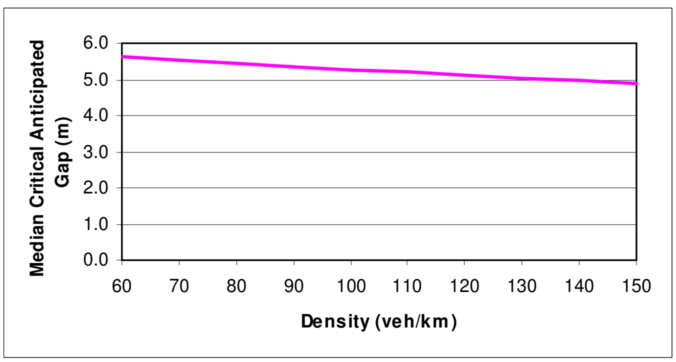


Figure 6 - Median critical anticipated gap as a function of density in target lane

The decision to initiate a forced merge is dependent on whether the lag vehicle is a heavy vehicle or not. If the lag is a heavy vehicle, the probability of initiating a forced merge decreases, as the driver perceives a higher risk in undertaking such a manoeuvre.

The probability of initiating a forced merge is given by the following equation:

$$P_{nt}^{F} = \frac{1}{1+\exp\left(6.41 + 1.25\delta_{nt}^{hv} - 5.43\upsilon_{n}\right)} \quad (17)$$

Where, δ_{nt}^{hv} is the heavy lag vehicle dummy, 1 if the lag vehicle is a heavy vehicle, 0 otherwise. It may be noted that the coefficient of aggressiveness has a significant impact on the decision to initiate a forced merge.

On initiating a courtesy/forced merge, the driver decides whether to complete the merge by accepting the available gap or not based on his respective lead and lag critical gaps. For identification purposes, except for the constant and the unobserved driver characteristics, the coefficients of variables in these levels are restricted to be the same as for the normal gap acceptance level.

Thus, the estimated lead and lag critical gaps can be given by the following equation:

$$G_{nt}^{CLead} = \exp\left(-0.582 + 0.521V_{nt}^{'} - 0.505Min\left(0, \Delta V_{nt}^{lead}\right) + \frac{1.32}{1+\exp(0.420+0.355\upsilon_n)} d_{nt} - 0.054\upsilon_n + \varepsilon_{nt}^{Clead}\right)$$
$$G_{nt}^{CLag} = \exp\left(\begin{array}{l} -1.23 + 0.208Max\left(0, \Delta V_{nt}^{lag}\right) + 0.184Min\left(0, \Delta V_{nt}^{lag}\right) \\ + \frac{0.439}{1+\exp(0.0242+0.00018\upsilon_n)} d_{nt} + 0.0545Max\left(0, a_{nt}^{lag}\right) - 0.554\upsilon_n + \varepsilon_{nt}^{Clag} \end{array}\right) \tag{18}$$

$$G_{nt}^{FLead} = \exp\left(3.11 + 0.521V_{nt}^{'} - 0.505Min\left(0, \Delta V_{nt}^{lead}\right) + \frac{1.32}{1+\exp(0.420+0.355\upsilon_n)} d_{nt} - 0.0401\upsilon_n + \varepsilon_{nt}^{Flead}\right)$$
$$G_{nt}^{FLag} = \exp\left(\begin{array}{l} -2.53 + 0.208Max\left(0, \Delta V_{nt}^{lag}\right) + 0.184Min\left(0, \Delta V_{nt}^{lag}\right) \\ + \frac{0.439}{1+\exp(0.0242+0.00018\upsilon_n)} d_{nt} + 0.0545Max\left(0, a_{nt}^{lag}\right) - 0.0239\upsilon_n + \varepsilon_{nt}^{Flag} \end{array}\right) \tag{19}$$

Where, G_{nt}^{Clead} and G_{nt}^{Flead} are lead critical gaps for the courtesy and forced gap acceptance levels (m) respectively, G_{nt}^{Clag} and G_{nt}^{Flag} are lag critical gaps for the courtesy and forced gap acceptance levels (m) respectively, $\varepsilon_{nt}^{Clead}, \varepsilon_{nt}^{Clag}, \varepsilon_{nt}^{Flead}$ *and* ε_{nt}^{Flag} are random error terms: $\varepsilon_{nt}^{Clead} \sim N\left(0, 0.0109^2\right)$ and $\varepsilon_{nt}^{Clag} \sim N\left(0, 0.554^2\right)$, $\varepsilon_{nt}^{Flead} \sim N\left(0, 7.95^2\right)$ and $\varepsilon_{nt}^{Flag} \sim N\left(0, 0.465^2\right)$.

The estimation results showed that all other things held constant, a driver is more willing to accept smaller lead and lag gaps when he is in the courtesy merging state than in normal or forced merging state. This is intuitive since in case of courtesy merging, the lag vehicle is slowing down and therefore, a smaller buffer space is sufficient.

The constant term for the lag critical gap for forced merging is the smallest. However, the lead critical gap for the forced merging case is relatively large reflecting the fact that once the driver has initiated a forced merge (pushed his front bumper establishing his right of way), the merge is completed only when the lead gap is sufficiently large since the manoeuvre involves significantly higher risk as compared to the normal gap acceptance.

The anticipation time is normally distributed within 0 to 4 sec.[1] The estimated distribution of anticipation time is

[1] Different values between 0 to 6 sec were tested as the upper limit of anticipation time and the selected value (4 sec) provided the best goodness-of-fit.

$$f(\tau_n) = \begin{cases} \frac{1}{0.833}\phi\left(\frac{\tau_n - 1.87}{1.44}\right) & \text{if } 0 \le \tau_n \le 4 \\ 0 & \text{otherwise} \end{cases} \tag{20}$$

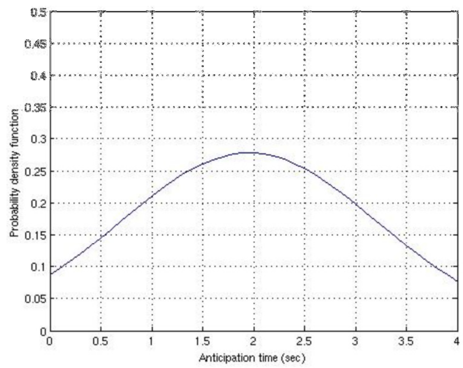


Figure 7- The distribution of anticipation time

MODEL COMPARISON

The state dependent merging model is compared against a simpler instantaneous model (Lee 2006) that does not capture the persistent behaviour of drivers and ignores state dependency. The instantaneous model aims at capturing the normal, forced and courtesy behaviour of drivers through a single gap acceptance level by including variables relevant to all three types of merges in a single critical gap function. The model structure is shown in Figure 8. The model is estimated with the same trajectory data.

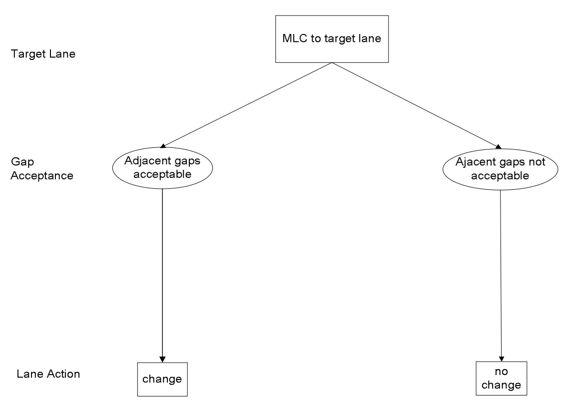


Figure 8-Framework of single level/instantaneous merging model (Lee 2006)

The state dependent model is an extension of the instantaneous model. The summary statistics of the estimation results for the two models, presented in Table 3, show a significant improvement in the fit of the model, even when accounting for the larger number of parameters in the state dependent model.

Table 3 – Model Comparison

Model	**Likelihood Function Value**	**Number of Parameters**
Instantaneous	-1639.69	17
State dependent model	-1609.65	42
$LR = 2\left[L(U) - L(R)\right] \sim \chi^2_{1-\alpha, k_U - k_R}$ $LR = 60.08 > \chi^2_{(0.95,25)} = 37.65$		

A likelihood ratio test was performed to select between the two alternative models. The likelihood ratio test results, also presented in Table 3, indicate that the unrestricted (U) state dependent model is significantly better than the restricted (R) instantaneous model. Therefore, the instantaneous model can be rejected as incorrect at 95% confidence interval.

The simulation capability of the state dependent model was compared with the performance of the instantaneous model within the microscopic traffic simulator MITSIMLab (Yang and Koutsopoulos 1996). Both models were implemented in MITSIMLab and the same merging section used for the model estimation (Interstate 80, California) was simulated. The comparison of the distribution of the actual travel time in the section and MITSIMLab simulations using each of the models are presented in Figure 9.

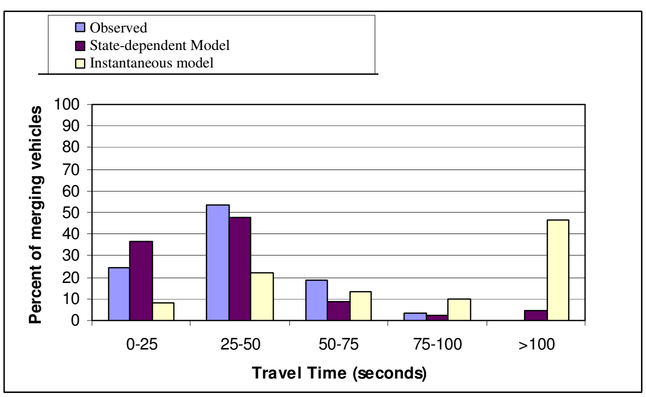


Figure 9- Observed and simulated travel times in the Interstate-80

As observed in Figure 9, the instantaneous model over predicts congestion in the merging section while the state dependent model has a much better replication of the reality. An extensive validation study to compare the simulation capability of the state dependent model using aggregate trajectory data collected from another site with a different ramp configuration is presented in Choudhury *et al.* (2006b).

CONCLUSIONS

In this paper, a methodology to model state dependency in lane changing behaviour has been demonstrated by applying it to model the merging behaviour of drivers in a congested freeway. The model has explicit normal, courtesy and forced merging components sequenced in a single decision framework. The decision to initiate a merge and the acceptance of gaps to complete the merge are affected by the decision state of the driver as well as neighbourhood variables and driver characteristics (agent effect). The model parameters for state-transition are estimated simultaneously with the parameters of the gap acceptance models with detailed vehicle trajectory data using maximum likelihood estimation technique.

The statistical model selection criteria using the estimation results showed that the proposed state dependent merging model is superior to a single level instantaneous model estimated with the same data ignoring state dependency. This result was further strengthened by a validation case study, which compared the results obtained from simulation runs from each of the model implementations in the microscopic traffic simulator MITSIMLab.

In the current model, only lateral decisions involved with the merging decision was modelled. The extent of the improvements obtained by incorporation of state dependency in the structure indicates the possibility of further enhancements through extension of the model to explicitly capture the state dependency between lane changing, target gap choice and acceleration decisions of the driver.

It may be noted that the methodology presented in this paper to model state dependency in merging behaviour can be extended to other driving behaviour models as well and this will be explored in future research.

ACKNOWLEDGEMENT

This paper is based upon work supported by the Federal Highway Administration under contract number DTFH61-02-C-00036. Any opinions, findings and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the Federal Highway Administration.

REFERENCES

Ahmed K.I. (1999), Modeling drivers' acceleration and lane changing behaviors, PhD thesis, Department of Civil and Environmental Engineering, MIT.

Benekohal R. and Treiterer J. (1988), Carsim: Car following model for simulation of traffic in normal and stop-and-go conditions. Transportation Research Record 1194, pp. 99-111

Cambridge Systematics Inc. (2005), NGSIM I-80 Data Analysis Report, prepared for Federal Highway Administration.

Choudhury C.F. (2005), Modeling Lane-changing Behavior in Presence of Exclusive Lanes, M.S thesis, Department of Civil and Environmental Engineering, MIT.

Choudhury C.F., Ben-Akiva M., Toledo T., Rao A., Lee G. (2006a), Cooperative Lane changing and Forced Merging Model: NGSIM Final Report, Federal Highway Administration.

Choudhury C.F., Ben-Akiva M., Toledo T. , Lee G. and Rao A. (2006b), Modeling Cooperative Lane Changing and Forced Merging Behaviour, Transportation Research Board, 86th Annual Meeting.

Gipps P.G. (1986), A Model for the Structure of Lane-changing Decisions, Transportation Research, 20B, pp. 403-414.

Lee G., Modelling gap acceptance in freeway merges (2006), M.S thesis, Department of Civil and Environmental Engineering, MIT.

Toledo T. (2003), Integrated Driving Behaviour Modelling, PhD thesis, Department of Civil and Environmental Engineering, MIT.

Wang J., Liu, R., Montgomery F. (2005), A simulation Laboratory for Motorway Merging Behaviour. In: Mahmassani, H (ed.) Transportation and Traffic Theory: Flows, Dynamics and Human Interaction, Elsevier, 5(3), pp.127-140.

Yang Q. and Koutsopoulos H.N. (1996), A microscopic traffic simulator for evaluation of dynamic traffic management systems. Transportation Research 4C, pp. 113-129.

Zhang Y., Owen L.E. and Clark J.E. (1998), A multi-regime approach for microscopic traffic simulation. Transportation Research Board, 77th Annual Meeting.