

2
3
4 **From Behavioral Psychology to Acceleration Modeling: Calibration,**
5 **Validation, and Exploration of Drivers' Cognitive and Safety**
6 **Parameters in a Risk-Taking Environment**
7
8
9

10
11 **Samer H. Hamdar***

12 hamdar@gwu.edu

13 Department of Civil and Environmental Engineering

14 **The George Washington University**

15 800 22nd Street, NW

16 Science and Engineering Hall #3810

17 Washington, DC 20052

18 Phone: (202) 994-6652

19 Fax: (202) 994-0127

20
21
22
23
24 **Hani S. Mahmassani**

25 masmah@northwestern.edu

26 Department of Civil and Environmental Engineering

27 **Northwestern University**

28 The Transportation Center

29 Chambers Hall, 600 Foster Street

30 Evanston, IL 60208

31 Phone: (847) 491-5017

32 Fax: (847) 491-3090

33
34
35
36
37 **Martin Treiber**

38 treiber@vwi.tu-dresden.de

39 Institute for Transport and Economics

40 **Technische Universität Dresden**

41 Andreas-Schubert-Straße 23

42 D-01062, Germany

43 Phone: +49-351-463-36794

44 Fax: +49-351-463-36809

45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
February 2015

Submitted for publication in Transportation Research Part-B: Methodological

From Behavioral Psychology to Acceleration Modeling: Calibration, Validation, and Exploration of Drivers' Cognitive and Safety Parameters in a Risk-Taking Environment

Samer H. Hamdar* (hamdar@gwu.edu), Hani S. Mahmassani, and Martin Treiber

ABSTRACT

We investigate a utility-based approach for driver car-following behavioral modeling while analyzing different aspects of the model characteristics especially in terms of capturing different fundamental diagram regions and safety proxy indices. The adopted model came from an elementary thought where drivers associate subjective utilities for accelerations (i.e. gain in travel times) and subjective dis-utilities for decelerations (i.e. loss in travel time) with a perceived probability of being involved in rear-end collision crashes. Following the testing of the model general structure, the authors translate the corresponding behavioral psychology theory - prospect theory - into an efficient microscopic traffic modeling with more elaborate stochastic characteristics considered in a risk-taking environment.

After model formulation, we explore different model disaggregate and aggregate characteristics making sure that fidelity is kept in terms of equilibrium properties. Significant effort is then dedicated to calibrating and validating the model using microscopic trajectory data. A modified genetic algorithm is adopted for this purpose while focusing on capturing inter-driver heterogeneity for each of the parameters. Using the calibration exercise as a starting point, simulation sensitivity analysis is performed to reproduce different fundamental diagram regions and to explore rear-end collisions related properties. In terms of fundamental diagram regions, the model in hand is able to capture traffic breakdowns and different instabilities in the congested region represented by flow-density data points scattering. In terms of incident related measures, the effect of heterogeneity in both psychological factors and execution/perception errors on the accidents number and their distribution is studied. Through sensitivity analysis, correlations between the crash-penalty, the negative coefficient associated with losses in speed, the positive coefficient associated with gains in speed, the driver's uncertainty, the anticipation time and the reaction time are retrieved. The formulated model offers a better understanding of driving behavior, particularly under extreme/incident conditions.

Keywords:

Car-Following, Congestion, Driver Behavior, Heterogeneity, Prospect Theory, Risk, Uncertainty

1. INTRODUCTION

In 2008, an acceleration-based car-following model was proposed that incorporates the risk-taking attitudes of drivers and uses prospect theory to evaluate the perceived consequences of applying different acceleration rates, a probability of collision and a crash penalty term are explicitly introduced in the formulation (Hamdar et al., 2008). This paper builds on this approach while changing the fundamental equations previously suggested (i.e. subjective utility function) for exploring the characteristics of the formulated car-following model in terms of its ability to capture congestion regions, equilibrium characteristics, inter-driver heterogeneity and collective accident-prone behaviors on a freeway section. Being calibrated against real-life trajectory data (FHWA – 2004 a,b,c), different bottleneck and incident scenarios are modeled: bottleneck scenarios are tested via deceleration exerted by the leader and on-ramp merging; incident scenarios are tested via rear-end collision and fixed object crashes. Special interest is given to studying the resulting fundamental diagram especially traffic breakdown and the congestion disturbances. The effect of both psychological factors and execution/perception errors on the accidents number and their distribution along a freeway length is also studied. Through sensitivity analysis, insights into the relationships between the crash-penalty, the negative coefficient associated with losses in speed, the positive coefficient associated with gains in speed, the driver's uncertainty, the anticipation time and the reaction time are provided.

The objective of this research is to offer a stochastic fully dynamical acceleration model incorporating a) traffic flow theory fundamentals (i.e. fundamental diagram), (b) findings of prospect theory on the form of the perceived (subjective, generalized) utility, (c) risk taking, (d) perception limits and subjective behavioral fluctuations. The model is supported by decision-making theories' concepts and safe/unsafe driving maneuvers are created as an inherent result of the utility function without imposing unrealistic safety constraints.

The first contribution of this research is capturing driver behavior in congested and situations while incorporating drivers' risk-taking attitude in the model equations. The model formulated in this paper does not exogenously impose safety constraints to prevent accidents. Models used in practice typically preclude accidents, contrary to real-life situations. One more implication of this contribution is capturing that drivers do not perfectly register existing stimuli without subjectively weighing different alternatives based on their personality (aggressive versus conservative drivers). This allows risky behavior as an inherent result of the model. Moreover, the corresponding acceleration choice emerges as a probabilistic decision making process facing uncertainty; the method by which the resulting accident causing behavior is weighed can be calibrated based on recorded traffic data. It should be mentioned that in the first and only published work on the related formulation (i.e. Hamdar et al., 2008), the formulation was not developed enough neither to derive the corresponding homogenous fundamental diagram nor to reproduce real-world trajectory data in different traffic conditions. In other words, the authors did not look beyond the acceleration probability density functions in different local scenarios (i.e. for specific relative velocities, spacing and initial velocities with non-

calibrated model parameters). The resulting probability density functions were feasible and acceptable as drivers accelerated when they were supposed to accelerate and decelerated when they were supposed to decelerate. The logical steps beyond such preliminary findings and that are performed in this research work are to:

- a- Expand the formulation and further test the corresponding car-following model in terms of the flow-density-speed macroscopic properties.
- b- Utilize a suitable calibration methodology and make sure that the suggested model reproduces rear-world trajectories especially during congestion conditions.
- c- Transform the model into an efficient simulation tool (possibly utilized for prediction and real-time evaluation purposes) and analyze the corresponding non-homogenous fundamental diagram (flow-density data points scattering).
- d- Make sure that the resulting model does not produce the same non-realistic crash patterns observed in existing acceleration models if the corresponding safety constraints were relaxed. These non-realistic patterns were already analyzed in Hamdar and Mahmassani, 2008.

The second and the third steps mentioned above are essential to be realized so the model may be useful for the traffic flow theory and traffic engineering communities. In other words, from a practitioner stand point, the main challenge in realizing the stated paper's objective while incorporating the corresponding parameters is the degree of complexity that would be added to the eventual model, which would preclude its usefulness in actual practice. Accordingly, another contribution of this research is to put forward a "logic" that is robust enough to advance the state of knowledge related to the driving task but simple and fast enough so that it can be readily implemented, calibrated and validated. The resulting model is intended to provide a competitive stochastic alternative to existing simpler models that lack cognitive dimensions.

In summary, the main challenge faced is translating the behavioral psychology prospect theory into a concise acceleration formulation given the importance of such structure for the calibration and the simulation exercise; this challenge is faced through the use of a Genetic Algorithm (GA) heuristic that allows calibrating the model for each "feasible" vehicle and attempting to capture a heterogeneity pattern. The structure of this paper will then follow; a background review on incidents and pertinent car-following models is presented in the following section. The framework of the work is shown in the third section where the corresponding car-following model is presented. The review and the framework will motivate testing the model in terms of equilibrium conditions. After calibrating the model, the fifth section includes the simulation results and the corresponding data analysis before concluding with some future research needs.

2. BACKGROUND REVIEW

In the year 2000, the monetary cost related to traffic accidents reached 230.6 billion USD (U. S. Dollars) in the U.S.A., only (NHTSA, 2007). Newer financial figures on this financial topic were not available. However, since the reduction of fatal US accidents in the last decade essentially is compensated for by inflation, it is safe to

assume that the present figure is not significantly different in terms of accidents' costs. Based on the National Highway Traffic Safety Agency (NHTSA) studies, 5 accident types of interest can be identified: 1) rear impacts (29.6% of US accidents), 2) angle or side impacts (28.6 % of US accidents), 3) fixed object crashes (16.1 % of US accidents), rollovers (2.3% of US accidents), head-on collisions (2 % of US accidents) and collision with pedestrians/bicyclists (1.8 % of US accidents) (3). In car-following, the focus is on the tailgating behavior that may lead to rear-end collisions (Type 1).

Based on the objective stated in the previous section, there are three main attributes in a car-following model that are desired: a) the model needs to be consistent with traffic flow theory fundamentals (i.e. fundamental diagram and traffic flow instabilities), (b) the model needs to have a cognitive dimension allowing for "utility" maximization, (c) risk taking and perception limits to be considered, and d) the model allows for incident prone conditions and errors to be made so that collisions may be created as an inherent result of the model structure without imposing unrealistic safety constraints.

Regarding the first attribute, there are several models that accurately capture traffic trajectory data, such as Gipps's model (Gipps, 1981) or the Intelligent Driver Model (IDM) (Treiber et al., 2000). Moreover, other models were built so that the perception limits are captured (Wiedemann and Reiter, 1992). Finally, some models were formulated using a utilitarian approach with a specific utility to be maximized by the drivers (Ahmed, 1999). All these models were tested in terms of the corresponding calibration efforts and resulting accuracy (Brockfeld et al., 2005). So one may ask what is the added contribution to be made?

Regarding the first attribute (i.e. attribute a: capturing all traffic flow fundamentals especially in relation to the congestion formation), the main assumption in "standard" car-following models is that the behavior of the following vehicle (e.g. change in acceleration) is directly related to a stimulus observed/perceived by the driver, defined relative to the lead vehicle (e.g. difference in speeds, headways etc.). This idea was adopted in the car-following models of Chandler, Gazis and Herman (Chandler et al., 1958, Gazis et al., 1959 and Herman et al., 1959), known as the General Motor (GM) models. These first models are not complete in the sense that they are not applicable to all traffic situations including, e.g., free traffic or approaching standing vehicles or obstacles. Later investigations proposed improved models by introducing a "safe" time gap and a desired speed. The Gipps model (Gipps, 1981), and the intelligent-driver model (IDM) (Treiber et al., 2000) contain intuitive parameters that can be related to the driving style such as desired accelerations, comfortable decelerations, and a desired "safe" time gap. Furthermore, they include braking strategies that prevent accidents under a given heuristic. With such safety constraints imposed (i.e. ignoring the importance of attribute d stated earlier: allowing incident formation as an inherent result of the driver decision making process), these models were able to capture the different traffic flow dynamics in congested regimes. However, these models lacked the cognitive dimension desired to further understand driver behavior (i.e. attribute b stated earlier). Accordingly, subsequent studies have extended these models, by introducing additional parameters

intended to capture dimensions such as anticipation, learning, and response to several vehicles ahead. Other models such as the human driver model (HDM) (Treiber et al., 2006) also model human deficiencies, including variable reaction times and the size and persistency of estimation errors of the input stimuli depending on the traffic situation. The Wiedemann model captures the indifference of the drivers to small changes in the stimuli. It also allows different execution modes including emergency braking (Wiedemann and Reiter, 1992). The MITSIM model (Ahmed, 1999), adopted a more utility-based approach to look into the fundamentals of decision making processes.

An open problem is that there is no car-following model having all the attributes stated in the objective. More precisely, the car-following models cited earlier are all designed to be accident free and therefore are, by definition, unsuited to capture driver behavior during incident scenarios (FHWA, 2004; Hamdar and Mahmassani, 2008); Moreover, a limited amount of research focused on the cognitive and risk-taking attitudes in driver behavior including the heterogeneity aspect that leads to scattering of flow-density data points and a more favorable environment for incidents (Lu et al., 2013).

One additional challenge to be faced when formulating new car-following models is the corresponding ease of calibration. Calibrating the above models need different levels of effort based on data availability, the number of parameters to calibrate, the calibration method and the model structure. For example, calibrating the Wiedemann Model requires estimating 18 parameters found in 17 different equations. On the other hand, in the IDM model, drivers behavior is captured by one equation with 5 parameters to estimate.

Before recent developments in collecting microscopic data (allowing the NGSIM research effort, FHWA 2004, 2005 a, b and c), a rare amount of data was available to calibrate the existing microscopic car-following models stated above. One of these data sets was collected by wire-linked vehicles on a test track at the General Motors Technical Center (Hamdar et al., 2009). Another technique was by using a camera attached to a helicopter. The gathered pictures were input to a time consuming manual processing system (Ossen and Hoogendoorn, 2007). Lately, image processing software and Differential Global Positions System (DGPS) have become available. This gave new tools to researcher to collect more accurate and detailed individual driver information.

Once the data is available, different calibrations techniques can be applied. In the “traditional” model calibration process, the “car-following” model parameters need to be adjusted until an acceptable (qualitative and quantitative) match is found between the simulated model dynamics and the observed drivers' behavior. Engineering judgment and trial-and-error methods are still widely used especially in the industry (Chu et al., 2004). More systematic approaches including the gradient method (Hourdakis et al., 2002) and Genetic Algorithm (Cheu et al., 1998) address the model calibration procedure as an optimization problem: a combination of parameter values are searched so an objective function (error term) is minimized. Lately, most research is oriented to capture intra and inter driver heterogeneity and time correlation in the parameters estimates (Ossen and Hoogendoorn, 2007). Given the presented literature, the formulation to be presented in

this paper need to be transformed in order to calibrate and validate the suggested model's "cognitive-based" parameters while leading to realistic traffic flow dynamics.

3. FRAMEWORK: the Car-Following Model

In this section, the general structure of the stochastic acceleration model is introduced. The implementation details of this model can be seen in (Hamdar et al., 2008). Some analytical and numerical derivations are not presented in this paper for conciseness.

In the free-flow regime, the main factor governing the acceleration behavior of a driver is his or her desired speed v_0 (Gipps, 1981). The acceleration applied by a driver to reach this speed starting at a speed v and having a maximum possible acceleration value a_{\max} is given by

$$\dot{v}_{free} = a_{\max} \left(1 - \frac{v}{v_0} \right). \quad \text{Equation 1}$$

In other words, the acceleration is always to be restricted by a free-flow acceleration function where $\frac{dv}{dt} = \dot{v}_{free} + \dot{v}_{int}$ and \dot{v}_{int} is the acceleration adopted when interactions between vehicles is present.

The acceleration in dense or congested traffic is mainly controlled by interactions with the leading vehicle. In this model, the acceleration is modeled by a stochastic process that is characterized by the following:

- 1- The expected acceleration value $\langle \hat{a} \rangle(t)$
- 2- the variance $\sigma_a^2(t)$
- 3- and by the correlation time τ_{corr}

At each given time, the stochastic acceleration \dot{v}_{int} is distributed according to a continuous logit model whose distribution function is conditioned to the actual

speed $v(t)$, the space gap $s(t)$ to the leader, and the relative speed $\Delta v(t)$ to the leader

($\Delta v > 0$ when approaching): $\dot{v}_{int} \sim \text{LOGIT}(s, v, \Delta v)$.

The conditional probability density $f_{int}(a | s, v, \Delta v)$ of the Logit model is given by

$$f_{int}(a | s, v, \Delta v) = \frac{e^{\beta U(a; s, v, \Delta v)}}{\int e^{\beta U(a'; s, v, \Delta v)} da'}. \quad \text{Equation 2}$$

The (generalized) utility U of the model is composed of the generalized (or perceived) prospect-theoretic acceleration utility $U_{PT}(a)$ whose form is derived by the prospect theory, and a penalty $U_{crash}(a)$ for the risk of accidents:

$$U(a; s, v, \Delta v) = U_{PT}(a) + U_{crash}(a; s, v, \Delta v). \quad \text{Equation 3}$$

We specify the utility component by

$$U_{PT}(x) = x * [0.5 * w_m + 0.5 * (1 - w_m) * (\tanh(x) + 1)] * (1 + x^2)^{0.5 * (\gamma - 1)}, \quad \text{Equation 3.1}$$

$$\text{where } x = \frac{a}{a_0},$$

and

$$U_{crash}(a; s, v, \Delta v) = p_c w_c. \quad \text{Equation 3.2}$$

In the prospect theoretic utility U_{PT} , a weighing function is adopted to evaluate the subjective utilities of different accelerations (Tversky and Kahnemann, 1986). The gains and losses are expressed as a function of acceleration, or, equivalently, in terms of expected speed gains and losses over a specific period of time (Figure 1). The non-varied model parameter a_0 indicates the subjective scale of the acceleration: accelerations $|\dot{v}_{int}| < a_0$ are considered to be “near the reference point” leading to increased sensitivity (Figure 1). In fact, even though a_0 is a fixed number for a given simulation (e.g., 1 m/s²), this referencing parameter may have two simultaneous meanings: (i) it may be interpreted as the scaling unit of the acceleration to be used inside exponentials or non-integer powers (such as $U_{PT}(a)$) requiring dimensionless arguments; (ii) moreover, this parameter denotes the range of accelerations where the prospect-theoretic subjective utility is significantly influenced by the reference point $a=0$ (hence the wording “range” in Table 1). Other parameters of interest in the corresponding value function are the weight associated with negative acceleration (w^- or w_m) and the nonlinear sensitivity component γ . The weight associated with the gains (w^+) is assumed to be one, so is the weighting of losses relative to that of gains (relative measure between w^+ and w_m). In other words, for $\gamma = 1$ (no increased sensitivity at the reference acceleration), the function expressed in Equation 3.1 has linear asymptotes just retaining the different positive and negative weighing, and a smooth transition of width a_0 around zero. Additionally, when $w_m = 1$, the utility becomes linear as $U_{PT}(x) = x = a / a_0$.

The second term on the right hand side of Equation 3 denotes the crash-related utility. In contrast to $U_{PT}(a)$ which is monotonously increasing with the acceleration a , U_{crash} is monotonously decreasing with a since a higher acceleration, and the ensuing higher

future speed, increases the risk of rear-end collisions. The utility U_{crash} consists of the estimated probability of a crash p_c and a crash weight w_c . The gradient of the crash utility is given explicitly by Equation 13. The estimated crash probability p_c is the probability of a rear-end collision within the time horizon τ assuming that i) the chosen acceleration a of the follower remains unchanged within this interval, ii) the speed of the leader is constant, and iii) this speed is only known imprecisely in terms of an unbiased Gaussian distribution of relative error (variation coefficient) α . In other words, when estimating the crash probability, drivers are assumed to predict the future position of a leader where the variation of this speed is dictated by an estimation uncertainty $\sigma(v_l) = \alpha v_l$ of the speed of the leader v_l . Regarding w_c , a higher w_c corresponds to conservative individuals while a lower corresponds to drivers willing to take a higher risk.

Notice that we assume the utility to be dimensionless. Furthermore, its derivative $U' \equiv \frac{dU}{da}$ with respect to acceleration is of the order of $1/a_0$ where $a_0 = 1\text{m/s}^2$. Consequently, $\frac{1}{\beta}$ has the order of magnitude of the intra-driver uncertainties of the acceleration.

3.1. Expectation Value

For sufficiently high values of β (which we will assume henceforth), the expectation value $\langle a \rangle$ of the distribution for \dot{v}_{int} can be approximated by its mode value, i.e., by the acceleration at the maximum of its probability density:

$$\langle a \rangle(s, v, \Delta v) = \int a f_{int}(a | s, v, \Delta v) da \approx a^*(s, v, \Delta v) = \arg \max U(a; s, v, \Delta v). \quad \text{Equation 4}$$

As usual, the value $a^*(t)$ for maximum utility can be determined by the condition:

$$U'(a^*; s(t), v(t), \Delta v(t)) = 0. \quad \text{Equation 5}$$

The dependencies between $U_{PT}(a)$ and $U_{crash}(a; s, v, \Delta v)$ usually lead to a unique maximum of the generalized utility at some acceleration a^* . The condition in Equation 5, however, generally will be satisfied for two values of the acceleration, where the higher one pertains to a minimum of the utility (unsafe driving mode).

Since, in general, the maximum of the total utility $U(a)$ can not be calculated analytically, a good initial guess for a^* is essential to allow for a robust and effective numerical approximation. Such a guess is provided by the special case $\gamma = w_m = 1$, where $U'_{PT}(a) = 1/a_0$. The ensuing condition $U'_{crash}(a^*) + 1/a_0 = 0$, with $U'_{crash}(a)$ given by Equation 13, can be analytically solved for a^* . This results in following initial guess (for details, see Hamdar et al., 2008):

$$a^* = \frac{2}{\tau} \left(\frac{\max(s - s_0, 0)}{\tau} - \Delta v + \alpha v z^* \right)$$

where

$$z^* = -\sqrt{2 \ln \left(\frac{\alpha_0 w_c \tau}{2\sqrt{2\pi\alpha v}} \right)}$$

3.2. Variance and Correlation Time

Assuming again a sufficiently large value of the Logit uncertainty parameter β , the variance of the distribution characterized by the probability density in Equation 2 can be calculated with the method of the asymptotic expansion. The result is:

$$\sigma_a^2(t) = \frac{-1}{\beta U''(a^*(t), s(t), v(t), \Delta v(t))}, \quad \text{Equation 6}$$

where the second derivative $U''(a^*, s, v, \Delta v)$ can be calculated analytically. The correlation time is given directly by the model parameter τ_{corr} .

The total number of parameters that need to be calibrated is seven. These parameters are presented in Table 1. The corresponding prospect theoretic utility function U_{PT} may be observed below:

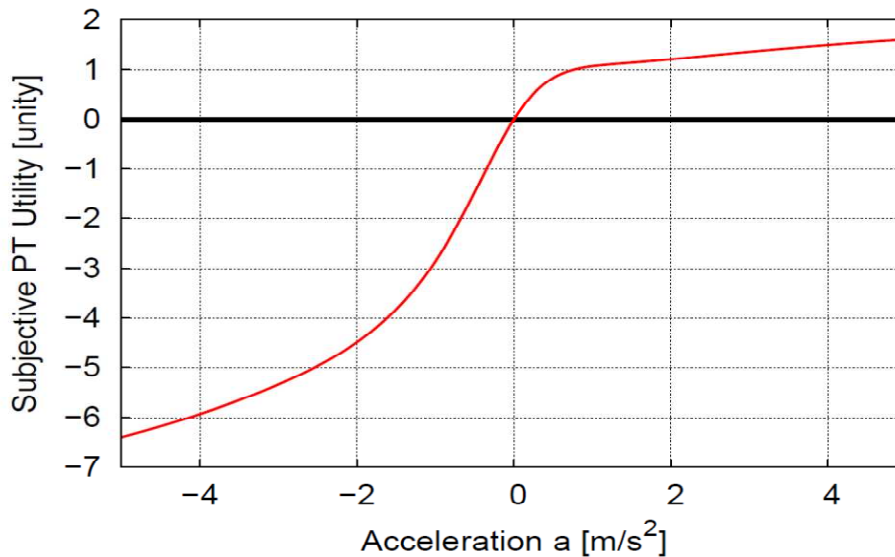


Figure 1: Prospect theoretic utility function U_{PT} for the proposed car-following model with the parameters from Table 1.

It should be noted that the logit expression (corresponding to Gumbel distributed acceleration estimation errors) expressed in Equation 2 was first chosen since there is no closed exact form maximizing the realizations of the total subjective utility in the presence of more than two alternatives for any other distribution. However, in the following, the approximation of asymptotic expansion leading to Equation 6 have led to the approximation of the Gumbel distribution by a Gaussian distributions, i.e., a second-order Gaussian stochastic process is produced for the acceleration errors which is uniquely characterized by β and τ_{corr} .

Given the above, for Equation 6 to be valid, $U''(a)$ must not vary significantly in $[a^* - 1/\beta; a^* + 1/\beta]$. Assuming a constant $U''(a)$ near the maximum a^* , i.e., $U(a) = U(a^*) + \frac{1}{2}U''(a^*)(a - a^*)^2$ and inserting this into f_{int} automatically leads to f_{int} proportional to $\exp[\frac{1}{2}U''(a^*)(a - a^*)^2]$, i.e., to Gaussian distributed acceleration uncertainties whose variance is given by Equation 6.

4. FUNDAMENTAL DIAGRAM AND EQUILIBRIUM CONDITIONS

The fundamental diagram, i.e., the steady-state relation for speed or flow as a function of density (or spatial gap), is given by the full model with following equilibrium conditions:

- The speeds of all vehicles are the same, and constant over time: i.e.,
 - acceleration $\dot{v}_0 = 0$,
 - speed difference (approaching rate to the leader) $\Delta v = 0$,
- no stochasticity is allowed, i.e., $\beta \rightarrow \infty$.

4.1. Microscopic Relations

The equilibrium relation is formulated as a relation between the gap s and the speed v , e.g., $v = v_e(s)$ or $s = s_e(v)$. Assuming $\frac{\partial \dot{v}}{\partial v} \leq 0$ (which should be satisfied for all sensible micro-models) (Treiber and Kesting 2013), the above equilibrium condition leads to:

$$v_e(s) = \min(v_0, v(s)), \quad \text{Equation 7}$$

where the steady-state speed $v(s)$ in the interacting range is defined by

$$\dot{v}_{int}(s, v, \Delta v = 0, \beta \rightarrow \infty) = 0. \quad \text{Equation 8}$$

Table 1: Car-Following Model Parameters and Corresponding Symbols for the Simulation Exercise: the parameters in the top part (above the horizontal line) are the actual model parameters to be calibrated; the parameters in the lower part are secondary parameters that are not subject to calibration.

Parameter	Symbols and Initial Values
Sensitivity Exponents of the Generalized Utility	$\gamma = 0.3$
Asymmetry Factor for Negative Utilities	$w_m = 4$
Speed Uncertainty Variation Coefficient	$\alpha = 0.08$
Weighing Factor for Accidents	$w_c = 100000$
Maximum Anticipation Time Horizon	$\tau = 5s$
Logit Uncertainty Parameter (Intra-Driver Variability)	$\beta = 5$
Correlation Time of Intra-Driver Variability	$\tau_{corr} = 20s$
Maximum Acceleration	$a_{max} = 1.5 \text{ m/s}^2$
Desired Speed	$v_0 = 30 \text{ m/s}$
Minimum Gap	$s_0 = 3 \text{ m}$
Acceleration Range Considered Near Interaction Point	$a_0 = 1 \text{ m/s}^2$

For the deterministic limit $\beta \rightarrow \infty$, the interaction acceleration reads

$$\dot{v}_{int} = \arg(\max(U(a; s, v, 0))), \quad \text{Equation 9}$$

so the steady-state relation can be expressed by

$$U'(a; s, v, 0)|_{a=0} = 0, \quad \text{Equation 10}$$

where the generalized utility $U(a; s, v, \Delta v)$ is understood as a function of the acceleration a .

Setting the seriousness term of the crash utility $\sigma(v, \Delta v) = 1$, one obtains from Equation 3 the condition:

$$U'_{PT}(0) + U'_{crash}(0) = U'_{PT}(0) + w_c p'_c(0; s, v, 0) = 0. \quad \text{Equation 11}$$

The gradient of the PT utility at $a = 0$ is obtained according to:

$$U'_{PT}(0) = \frac{\left(\frac{1 + w_m}{2}\right)}{a_0}. \quad \text{Equation 12}$$

As stated in Section 3 (see Hamdar et al., 2008), the gradient of the utility due to the crash risk is given by:

$$U'_{crash}(a) = -w_c f_N \left(\frac{\Delta v + \frac{1}{2} a \tau - \frac{s - s_0}{\tau}}{\alpha v} \right) \frac{\tau}{2\alpha v}, \quad \text{Equation 13}$$

or, after inserting the steady-state conditions $a = 0$ and $\Delta v = 0$:

$$U'_{crash}(0) = -\frac{w_c \tau}{2\alpha v} f_N \left(\frac{-(s - s_0)}{\alpha v \tau} \right) = -\frac{w_c \tau}{2\sqrt{2\pi}\alpha v} e^{-\left(\frac{s - s_0}{\sqrt{2}\alpha v \tau}\right)^2}. \quad \text{Equation 14}$$

Inserting Equations (14) and (12) into (11) and solving for the steady-state gap s results in:

$$s(v) = s_0 + \sqrt{2} v \alpha \tau \sqrt{\ln\left(\frac{a_0 \tau}{v}\right) + \ln\left(\frac{w_c}{2\sqrt{2\pi}\alpha\left(\frac{1 + w_m}{2}\right)}\right)}. \quad \text{Equation 15}$$

4.2. Macroscopic Relations

The congested (interacting) branch of the macroscopic fundamental diagram is obtained by applying the usual micro-macro relations:

$$\rho(v) = \frac{1}{l_{veh} + s(v)}, \quad \text{Equation 16}$$

and

$$Q(v) = v * \rho(v). \quad \text{Equation 17}$$

It should be noted that based on the initial parameter testing, the first log term in the square root term of Equation 15 is of the order of unity (unless the speed is very small), while the second constant log term is of the order of ten (w_c is in the range of 100 000).

Therefore the steady-state time gap $T_e(v) = \frac{s(v)}{v}$ is nearly constant (Figure 2) resulting in an approximately triangular fundamental diagram (Figure 3).

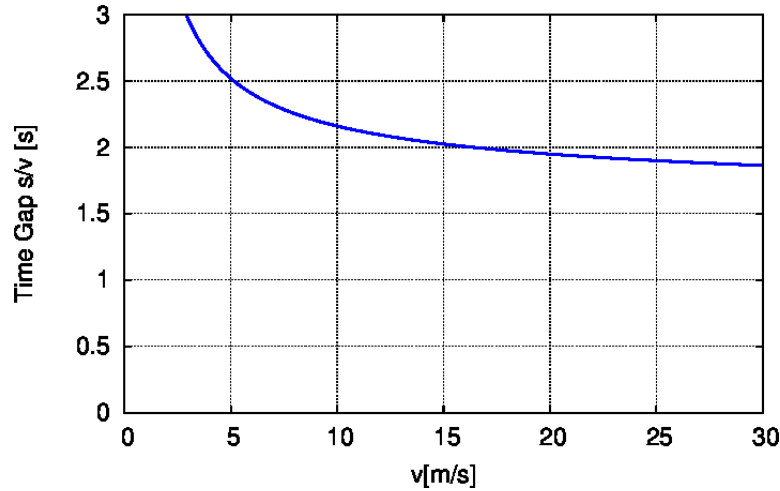


Figure 2: Steady-state time gap $T_e(v) = s(v)/v$ for the proposed car-following model with the parameters from Table 1.

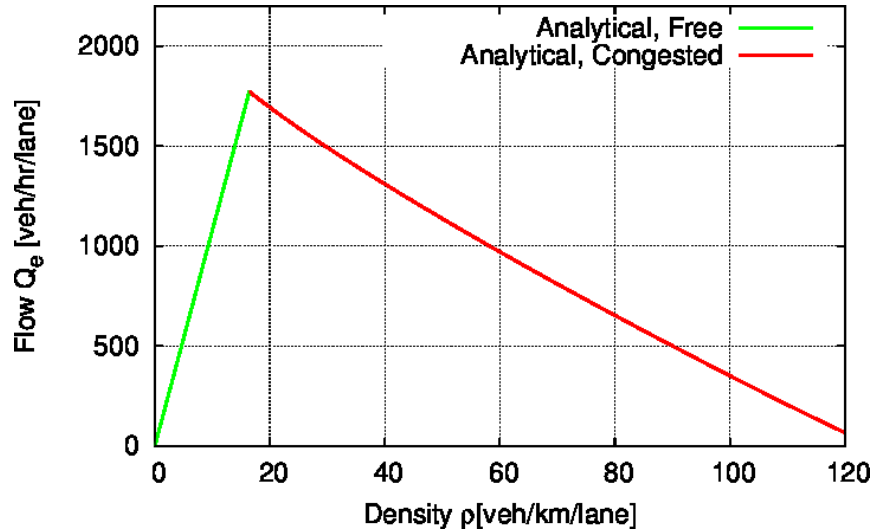


Figure 3: Fundamental diagram for the acceleration model with the parameters from Table 1.

As in other triangular-shaped fundamental diagrams, the capacity is mainly determined by the inverse of the time gap. The only parameter in the logarithmic term of Equation 15 influencing $s(v)$ (and thus the capacity), and where a variation by several orders of magnitude is probable, is the crash weighing factor w_c : all other logarithmic expressions are essentially zero compared to $\ln(w_c)$. Therefore, the effective time gap can be approximated by

$$T_{eff} := \frac{s_e - s_0}{v} \approx \alpha \tau_{max} \sqrt{2 \ln(w_c)}. \quad \text{Equation 18}$$

Figure 4 plots the exact analytic value for T_{eff} calculated based on Equation 15 together with the approximation obtained from Equation 18– both as a function of the speed. One sees that the approximation breaks down only for very low values of v . By varying the model parameters while keeping the product defined by Equation 18 constant, once can change the dynamical model properties (e.g., string stability, sensitivities, and accelerations) independently of the static properties which are essentially defined by the capacity.

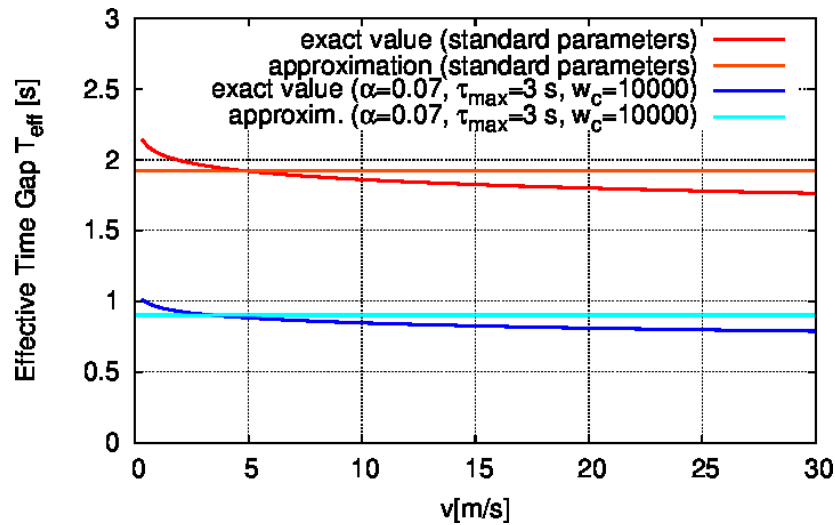


Figure 4: Effective time gap $T_{eff} = (s_e(v) - s_0) / v$ as a function of the speed in the congested regime. Compared is the exact analytic expression based on Equation 15 with the constant approximation of Equation 18.

After studying the suggested car-following model in terms of derivation and micro/macro properties, a detailed numerical analysis is presented in the following section.

5. NUMERICAL RESULTS: Calibration and Data Analysis

The authors recognize the complexity involved when calibrating a car-following model using trajectory data, if the model's acceleration function is not given explicitly. Multiple alternatives exist when choosing the optimization algorithm, the measures of performance, and the goodness of fit function (Punzo et al., 2012). Even though calibrating the model does not constitute the focus of this study, this section presents a thorough calibration exercise highlighting the car-following model properties.

5.1. Data Description and Calibration

To calibrate the model, we have used trajectory data of the Federal Highway Administration's (FHWA) Next Generation Simulation (NGSIM) project (FHWA, 2005 a, b and c); Data for 5678 vehicles have been collected on the 13th of April, 2005, on a segment of the Interstate I-80 in Emeryville, San Francisco, USA. The vehicles considered were traveling North-Bound and were tracked using video cameras mounted on the Pacific Park Plaza (a 30 story building located on 6363 Christine Avenue).

The videos were recorded using 7 video cameras (cameras 1 through 7). Camera 1 recorded the southernmost section of the I-80 segment included in the study area and Camera 7 recorded the northernmost section.

Finally, the filtered trajectory points were grouped into three 15 minutes' intervals:

- 1- Data Set 1: collected from 4:00 PM to 4:15 PM (2052 vehicles)
- 2- Data Set 2: collected from 5:00 PM to 5:15 PM (1836 vehicles)
- 3- Data Set 3: Collected from 5:15 PM to 5:30 PM (1790 vehicles)

Data were recorded every 1/10 second. The area covered in these data sets includes an on-ramp but does not include an off-ramp and has a length of 1650 feet. The focus of this paper is on Data Sets 1 and 2 (FHWA, 2005 a and b).

5.2. Model Calibration with Heterogeneity

Since the model relies on utility maximization technique with a stochastic choice between different acceleration alternatives, the corresponding equations contain stochastic elements themselves and therefore are analytically intractable. Moreover, this also implies that the objective function (the sum of squared errors) is not smooth as a function of the parameters. For that reason, we have calibrated the model using a nonlinear optimization procedure that is based on a genetic algorithm (Hamdar et al., 2009). The objective is to minimize the deviations between the observed and simulated trajectories when following the same designated leader, and avoiding local minima.

Based on the stimuli considered in the acceleration model, the required trajectory data should include speeds for both the leading and the following vehicles of interest. Accordingly, a direct comparison between the measured driver behavior and the trajectories simulated by the car-following model - with the leading vehicle serving as externally controlled input - is possible. In the simulation set-up, the calibration is performed by taking different leader-follower pairs and comparing their driving dynamics with the behavior obtained from the simulated car-following model. The simulated relative speed and the distance gap are initialized to the empirically given relative speed and gap (prescribed values: Treiber and Kesting, 2013):

$$\Delta v^{sim}(t=0) = \Delta v^{data}(0), \quad \text{Equation 19}$$

$$s^{sim}(t=0) = s^{data}(0). \quad \text{Equation 20}$$

The microscopic acceleration model is used to compute the acceleration and thus the trajectory of the following vehicle. The gap to the leading vehicle is computed as the difference between the simulated trajectory $s^{sim}(t)$ (front bumper position) and the recorded position of the rear-bumper of the leading vehicle $s_{lead}^{data}(t)$:

$$s^{sim}(t) = x_{leader}^{data}(t) - x^{sim}(t). \quad \text{Equation 21}$$

The above measure can be directly compared to the gap $s^{data}(t)$ provided by the data. It should be noted that the rear-end and the front-end bumper positions of the leaders and the followers can be extracted since the NGSIM data contains the corresponding vehicle lengths.

In the calibration process, the difference between the observed driving behavior and the driving behavior obtained by the simulated car-following model should be minimized by choosing a set of “optimal” model parameters. Different error measures based on speed, relative speed, or the space gap, can be used. Normally, the error in the space gap s is adopted: when optimizing with respect to s , the average relative speed errors are automatically reduced. In contrast, when optimizing with respect to the relative speeds Δv , the error in the distance gap may incrementally grow (Punzo et al., 2012).

In this study, due to errors in recording the space gaps in the NGSIM data (Thiemann et al., 2008), the optimization procedure is performed with respect to the speed v ; since the NGSIM data are collected during the peak-hour PM congested period, the image processing of the recorded videos resulted in transforming some space gaps into negative space gaps after subtracting the vehicle lengths. On the other hand, since the form of the objective function has a direct impact on the calibration results, three different error measures can be considered. The relative error is defined as a function of the empirical and the simulated time series ($v^{sim}(t)$ and $v^{data}(t)$ respectively):

$$F_{rel}[v^{sim}] = \sqrt{\left\langle \left(\frac{v^{sim} - v^{data}}{v^{data}} \right)^2 \right\rangle}, \quad \text{Equation 22}$$

where $\langle \cdot \rangle$ refers to the temporal average of a time series of duration ΔT :

$$\langle z \rangle = \frac{1}{\Delta T} \int_0^{\Delta T} z(t) dt. \quad \text{Equation 23}$$

The relative error is more sensitive to small speeds v than to large speeds. The main reason behind such a sensitivity is that the measure is weighted by the inverse recorded speed v^{data} .

The second measure considered is the absolute error:

$$F_{abs}[v^{sim}] = \sqrt{\frac{\langle (v^{sim} - v^{data})^2 \rangle}{\langle v^{data} \rangle^2}}. \quad \text{Equation 24}$$

Since the denominator is averaged over the whole trajectory interval, the absolute error $F_{abs}[v^{sim}]$ is less sensitive to small deviations from the empirical data than the relative error $F_{rel}[v^{sim}]$. On the other hand, the absolute error is more sensitive to large differences

in the numerator (large speeds reflecting large gaps). It should be noted that the error measures are normalized so they are independent from the duration ΔT .

Since the absolute error systematically overestimates errors for large gaps (at high speeds) while the relative error systematically overestimates deviations of the observed gap in the low speed range, a mixed error measure will be used as the objective function in this paper:

$$F_{mix}[v^{sim}] = \sqrt{\frac{1}{\langle v^{data} \rangle} \left\langle \frac{(v^{sim} - v^{data})^2}{|v^{data}|} \right\rangle}. \quad \text{Equation 25}$$

Once the objective function to be minimized is defined, the genetic algorithm is applied as a search heuristic to find an approximate solution to the nonlinear optimization problem:

- i- A “chromosome” represents a parameter set of the car-following model introduced earlier in this work (notice that, in biology, the genome, i.e., the complete genetic information, generally includes several chromosomes; here we just consider one chromosome). A population consists of N_{GA} such chromosomes.
- ii- In each population generation, the fitness of each chromosome is determined via the objective function defined in Equation 25.
- iii- Pairs of chromosomes are exclusively generated from the current population and recombined to generate a new population of chromosomes.
- iv- The cross-over point where two chromosomes are combined is randomly selected. It should be noted that we simply adopt a one-point cross-over strategy at this stage.
- v- Except for the chromosome with the best fitness score, all the genes (model parameters) are mutated (varied randomly) following a given probability. The resulting chromosomes (new generation) are used in the next iteration.
- vi- Initially, a fixed number of generations is evaluated. The evolution is then terminated when the best-of-generation score does not improve from one iteration to another for a given number of generations.

In this research, the initial set of 10 parents is initiated where the parameters are given values in the proximity of those provided and tested in Table 1. At each iteration, these parents produce 90 children chromosomes where the best 10 candidates of the N_{GA} population ($N_{GA} = 90 + 10$) are kept to the next iteration. The calibration process continues until no improvement of more than 0.01 is observed for 20 consecutive iterations, or when the error reaches the threshold of 15%. It should be noted that a mutation rate of 10% is applied in all iterations.

As seen in Table 1, there are seven main parameters to be calibrated. The additional parameter incorporated in the process is the reaction time. The reaction time is normally incorporated in an acceleration model using one of two methods. In the first method, it is associated with the update time of the numerical integration scheme. If accelerations are

1
2
3
4 kept constant during each update step (“ballistic update”) as is done here, this means the
5 driver is not reacting to his/her surrounding during the update time or, in other words, the
6 driver reacts to his surrounding every reaction time in a discrete. On the other hand, the
7 reaction time may be implemented on the model’s level as a explicit delay leading to a
8 delay-differential equation which is subsequently solved with small time steps (2nd
9 implementation method). The authors chose the first implementation method when
10 calibrating the corresponding reaction time. It should be noted that the effect of a given
11 update time is essentially that of an explicit delay of half that duration (Treiber et. al.,
12 2006).
13
14

15 Focusing on the car-following instances in the offered trajectory data, the related
16 summary results of the calibration exercise using data sets 1 and 2 are shown in Tables 2
17 and 3 (i.e; all vehicle trajectories). The presented parameters are the calibrated parameters
18 that led to errors below a 30% threshold. For illustration reasons, sample simulated versus
19 trajectory data are shown in Figure 5. Notice that towards the end of the calibration
20 simulation (around time-step 1000 in Figure 5), the simulated speeds and the simulated
21 space gaps drop to zero. The time of this drop corresponds to the time when gap data
22 ceases to exist in the NGSIM data sets (i.e., the cameras cannot detect anymore the
23 corresponding lead vehicle). When such lead-vehicle data are not available (space gaps
24 and relative speeds), the calibration exercise is terminated and no further contribution to
25 the mixed error term is recorded.
26
27
28

29 The first interesting finding is the important level of inter-driver heterogeneity
30 although the distributions of parameters values are not clear Gaussian distributions. A
31 clear peak appears in all distributions but with a vast range of parameters values. When
32 examining the average values, the cognitive nature of parameters allows interesting
33 interpretations; for example, drivers seem to put 4 times the negative weight ($W_m \sim 4$) on
34 losses in speed than on gains (the corresponding weight is assumed to be equal to 1).
35 Moreover, even though crashes are not avoided through the use of safety constraints, the
36 calibrated high value of the crash weight W_c (~ 100000) reduces the possibility of
37 accidents in this simulation exercise. On the other hand, notice that the mean and the
38 standard deviation for the *Gamma* and the reaction time R_t parameters are close in value.
39 This may suggest a possible exponential probability density function. Also notice that the
40 correlation matrix shows mainly low correlation values between different parameters.
41 This may indicate a high level of independence between parameters which is a desirable
42 property. The corresponding parametric correlation is a subject of future research.
43
44
45
46

47 The distributions of the different calibrated parameters values are illustrated in the
48 figures of Appendix A; one may point that the parameters have a non-Gaussian shaped
49 distribution functions with a governing peak value.
50
51
52
53
54
55
56
57
58
59

Table 2: Summary Statistics for Calibrated Parameter Values Using Genetic Algorithm (GA) – Data Set 1.

Parameter	Units	Mean	Std. Dev.	Minimum	Maximum
<i>Gamma</i> (γ)	-	0.333991	0.339046	0	1.9
<i>Wm</i> (w_m)	-	3.97208	2.6452	0.2	9.8
<i>Wc</i> (w_c)	-	97077.4	21143.6	50000	149000
<i>T</i> (τ)	seconds	5.08938	1.98571	1	10.9
<i>Alpha</i> (α)	-	7.74E-02	3.93E-02	1.00E-02	0.46
<i>Beta</i> (β)	-	5.32671	2.097	1	10.9
<i>Tcorr</i> (τ_{corr})	seconds	19.9833	4.53906	10	29
Reaction Time (<i>Rt</i>)	seconds	0.587102	0.688398	0.1	3.2

Table 3: Correlation Matrix for Calibrated Parameters Using GA– Data Set 1.

Correlation	<i>Gamma</i>	<i>Wm</i>	<i>Wc</i>	<i>Tmax</i>	<i>Alpha</i>	<i>Beta</i>	<i>Tcorr</i>	<i>Rt</i>
<i>Gamma</i>	1	0.26	-0.07	0.4	-0.14	0.2	0.01	-0.009
<i>Wm</i>	0.26	1	-0.02	0.15	-0.09	0.11	0.05	-0.04
<i>Wc</i>	-0.07	-0.02	1	-0.07	0.001	-0.03	0.01	-0.002
<i>T</i>	0.4	0.15	-0.07	1	-0.18	0.04	-0.06	0.19
<i>Alpha</i>	-0.14	-0.09	0.001	-0.18	1	-0.08	-0.00007	-0.02
<i>Beta</i>	0.2	0.11	-0.03	0.04	-0.08	1	-0.03	0.07
<i>Tcorr</i>	0.01	0.05	0.01	-0.06	-0.00007	-0.03	1	0.04
<i>Rt</i>	-0.009	-0.04	-0.002	0.19	-0.02	0.07	0.04	1

Table 4: Summary Statistics for Calibrated Parameter Values Using Genetic Algorithm (GA) – Data Set 2.

Parameter	Units	Mean	Std. Dev.	Minimum	Maximum
<i>Gamma</i> (γ)	-	0.31619	0.345064	0	1.9
<i>Wm</i> (w_m)	-	3.85276	2.62247	0.2	9.9
<i>Wc</i> (w_c)	-	97556.2	22059.5	50000	149000
<i>T</i> (τ)	seconds	5.04476	1.95507	1	10.9
<i>Alpha</i> (α)	-	7.72E-02	3.54E-02	1.00E-02	0.45
<i>Beta</i> (β)	-	5.37314	2.33261	1	10.8
<i>Tcorr</i> (τ_{corr})	seconds	20	4.45203	10	29
Reaction Time (<i>Rt</i>)	seconds	0.658857	0.726583	0	2.9

Table 5: Correlation Matrix for Calibrated Parameters Using GA– Data Set 2.

Correlation	<i>Gamma</i>	<i>Wm</i>	<i>Wc</i>	<i>Tmax</i>	<i>Alpha</i>	<i>Beta</i>	<i>Tcorr</i>	<i>Rt</i>
<i>Gamma</i>	1	0.23	0.04	0.34	-0.11	0.16	-0.03	-0.02
<i>Wm</i>	0.23	1	0	0.19	-0.11	0.13	0.03	-0.05
<i>Wc</i>	0.04	0	1	0.02	0.03	0.02	-0.07	0.06
<i>T</i>	0.34	0.19	0.02	1	-0.24	0.13	0.05	0.16
<i>Alpha</i>	-0.11	-0.11	0.03	-0.24	1	-0.07	-0.06	0.02
<i>Beta</i>	0.16	0.13	0.02	0.13	-0.07	1	0.00	0.12
<i>Tcorr</i>	-0.03	0.03	-0.07	0.05	-0.06	0.00	1	0
<i>Rt</i>	-0.02	-0.05	0.06	0.16	0.02	0.12	0	1

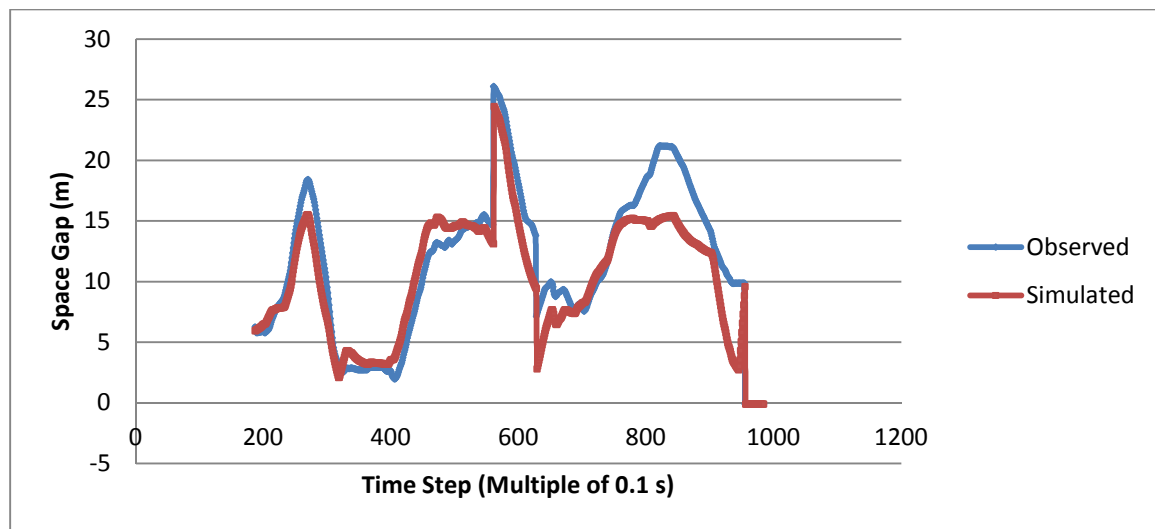
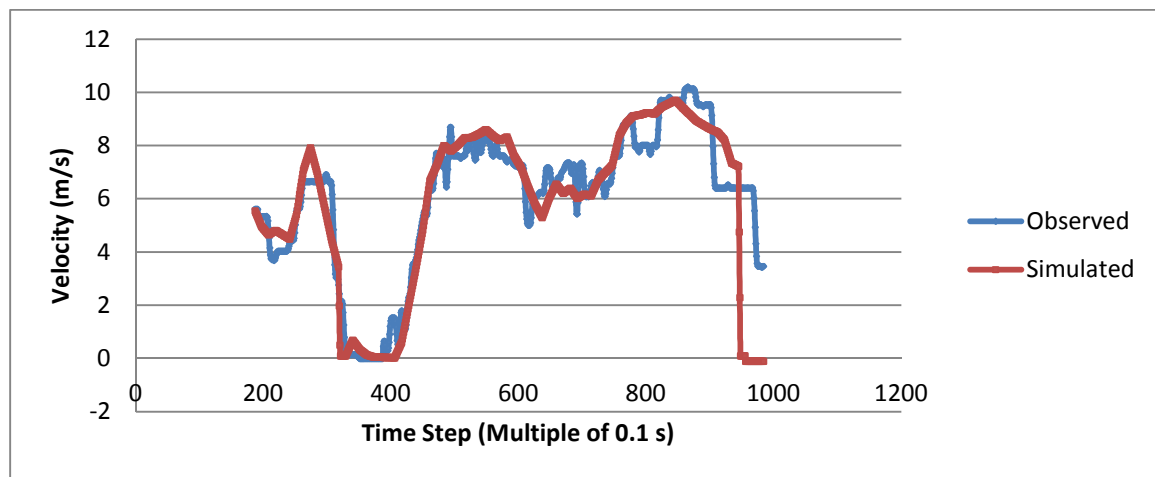


Figure 5: Simulated versus observed speeds (upper graph – mixed error = 0.12) and space gaps (lower graph – mixed error = 0.29) for Vehicle 32 from data set 1

However, when trying to estimate the corresponding distributions, no statistically significant function was found. In other words, at this stage, due to the lack of data,

conclusive results on the distribution followed by each parameter values could not be reached (null hypothesis on the corresponding distributions tested for acceptance or rejection); In addition to the lack of data, the main reason behind such result is that, even though a significant heterogeneity (spread of parameter values across drivers) exists, the concentration of parameter values around one peak is too high for existing parametric distribution functions to capture.

5.3. Model Validation

In this section, we perform a simple validation to assess the robustness of the parameter values calibrated in the previous section. Since the acceleration model parameters (Table 1) are driver specific, and data sets 1 and 2 have a different number of vehicles, the parameter values corresponding to the peak values found in the calibration process are applied in the validation process. In other words, for each vehicle in data set 2, the longitudinal trajectory points (x^{data} and v^{data}) obtained from NGSIM are compared to the trajectory points (x^{sim} and v^{sim}) obtained if applying the peak-parameter values obtained through the calibration exercise performed on data set 1 (i.e. Appendix A: Figure A.1 through Figure A.8). For each vehicle, the mixed error velocity term is recorded. The same exercise is repeated on data set 1 using the calibrated peak-parameters of data set 2 (i.e. Appendix A: Figure A.9 through Figure A.16). The corresponding error terms are presented in Table 7.

Table 7: Validation Error Terms

Validation Error	Mean Mixed Error	Std of Mixed Error	Minimum	Maximum
Data Set 1	0.3258	0.353026	0.036213	1.965293
Data Set 2	0.3107	0.291286	0.048659	2.191862

For both data sets 1 and 2, the mean mixed error term is equal to ~0.3.

Some of the validation errors reach values close to 200%. To examine this phenomenon, the distribution of the error term is plotted for data sets 1 and 2 in Figure 6; around 50% of the errors have a value less than 30% (~50% of the error values less than 30%). This error threshold is comparable to that found in existing calibration studies (Kesting and Treiber, 2008) and is considered reasonable. For the remaining 50%, the main problem is the deterministic nature of assigning calibrated parameter values (peak values) to different drivers irrespective of their behavioral nature found in the calibration process (their initial calibrated parameters). In addition, inter-driver variations pertain to the discussed distributions of the parameters; an upper limit of the intra-driver variations can be assessed by the residual sum of squared errors (SSE). In terms of variances and assuming independence between inter- and intra-drive variations, the cross-calibration/validation variance (or SSE_{cross}) can be used to estimate the relative contributions:

$$SSE_{cross} = SSE_{intra} + SSE_{inter} \quad \text{Equation 26}$$

Where SSE_{intra} (the average SSE of all trajectories on calibration) characterizes the upper limit of the intra-driver variations, and SSE_{inter} represents inter-driver variations. After filtering the trajectories that led to determinate error values, SSE_{cross} and SSE_{intra} were computed for data sets 1 and 2 and the following was found:

Data Set 1: $SSE_{cross} = 548.7836$ and $SSE_{intra} = 118.3996$

Data Set 2: $SSE_{cross} = 390.0792$ and $SSE_{intra} = 102.7188$

Such results indicate the considerable contribution of inter-driver heterogeneity to the cross-validation error variance.

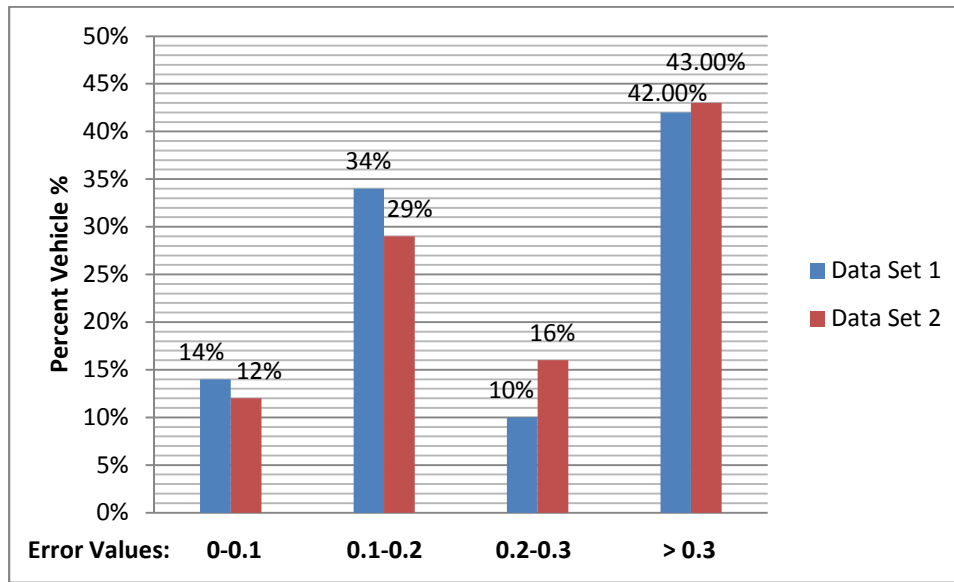


Figure 6: Error distribution across vehicles when using parameters of data set 2 on data set 1 (Data Set 1), and using parameters of data set 1 on data set 2 (data set 2).

Finally, for added verification, the same calibration method (i.e., the Genetic Algorithm – GA – Method) was used to calibrate a well-known car-following model, the Intelligent Driver Model (IDM) (Treiber et al., 2000). This model was also tested with the MOBIL lane-changing model showing that integrating an acceleration function within the lane-changing module is a possibility without reducing the significance of the contributions related to the captured car-following dynamics and the resulting congestion regime(s). For the NGSIM data set collected between 5 and 5:15 pm, the longitudinal trajectories for 2956 vehicles were compiled. With 20 iterations as a maximum allowable number of iterations, a defined calibration solution corresponding to a mixed speed error term less than 33% was found for 1779 vehicles. The median mixed speed error value is 11% and the minimum mixed speed error value is 2%. These results are comparable to that obtained for the Prospect-Theory based model. For the 1588 trajectories from the same data set having a mixed speed error threshold below 33%, the median mixed speed error value is 13% and the minimum mixed speed error value is 3%. We find that the proposed prospect-theoretic model results in slightly larger fitting errors. However, this is to be

1
2
3
4 expected since, in contrast to the IDM, our proposed model is stochastic in nature
5 simulating all the intra-driver variability. Since one cannot expect any correlation
6 between the real and the simulated intra-driver variations, the variance of the speed errors
7 is the sum of the variances originating from both sources when calibrating single
8 trajectories.
9

10 11 **6. SIMULATION AND SENSITIVITY ANALYSIS**

12 13 **6.1. Flow-Density Relation**

14
15 In this section, the car-following model is simulated using parameters values of
16 the same order of the estimates found in the previous section. The car-following model is
17 combined in this exercise with the MOBIL lane changing model for added robustness
18 (Kesting et al., 2007): the acceleration rules are used for assessing the “safe” comfortable
19 gaps to change lanes. The vehicles are “injected” on a two-lane freeway and a one-lane
20 on-ramp merging together at location $x = 10$ km. The “freeway entrance” is at $x=0$ km,
21 and $x=6$ km, 9 km, and 10 km are the positions of virtual detectors. The calibrated lane-
22 changing function is called at each simulation time step to determine the desirability of
23 changing lanes. Figure 7 illustrates the resulting fundamental diagram.
24
25

26
27 In the fundamental diagram, analytically, two main regions emerge: a free-flow
28 region (green straight line) and a congested region (red straight line). When simulating
29 the model, virtual one-minute detectors are placed to collect flow and density measures in
30 three different scenarios based on the location of the merging of the on-ramp traffic and
31 the main-stream traffic. When this location is close to the “freeway entrance” ($x = 6$ km),
32 the transition between the free-flow region and the congested region is seen through a
33 sharp traffic breakdown (sudden drop in volume – red line). This traffic breakdown is
34 followed by scattered “non-synchronized” flow-density points in the congested region.
35 As the on-ramp is further away from the free-way entrance, a smoother transition occurs
36 where synchronized flow-density points appear. This kind of traffic dynamics imitate
37 some of the observations that appear in real-life situations (Treiber et al., 2000).
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59

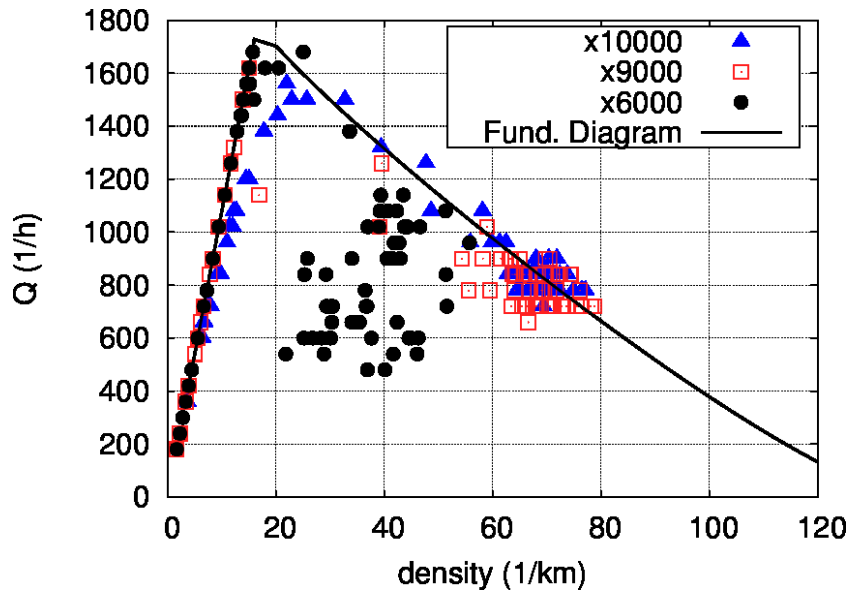


Figure 7: Fundamental diagram of the combined car-following/MOBIL model with the parameters described in Table 1. The results are based on virtual one-minute detectors at the indicated locations. The on-ramp bottleneck is at $x=10$ km.

6.2. Inter-Driver Heterogeneity

In the inter-driver heterogeneity related simulation sensitivity analysis, the base-case scenario is taken with the calibrated parameters. The main cognitive parameters of interest are the crash-penalty, the negative coefficient associated with losses in speed, the positive coefficient associated with gains in speed, the driver's uncertainty, the anticipation time and the reaction time.

The acceleration model is simulated with a simulation time-step of 0.1 second. The vehicles are "injected" into a 10 km two-lane freeway section. The initial flow-rate is controlled by an exponential inter-arrival time with a given mean. Since the interest is in capturing all the regions of the fundamental diagram (free-flow and uncongested), at different road sections, a kilometer bottleneck is created through the allowance of an unstable and abrupt vehicle deceleration. This also favor the creation of rear-end collision for testing the influence of the different model parameters. Figure 8-a shows the resulting flow-density data points and hysteresis triangle if formed accordingly.

To test the effect of inter-driver heterogeneity, two families of scenarios are offered. The first family is related to homogenous traffic where the parameters values for all vehicles are constant and correspond to the peak of values of the parameters distributions found in the calibration exercise. The second family is related to heterogeneous traffic where the parameters values have a normal distribution where the mean corresponds to the peak found in the calibration results; Figure 8-b shows a slightly increasing flow-density data points scattering if compared to the homogeneous scenario simulation. Such slight increase in scattering of the flow-density data points while

keeping the corresponding triangular fundamental diagram characteristics underlines the robustness of the model.

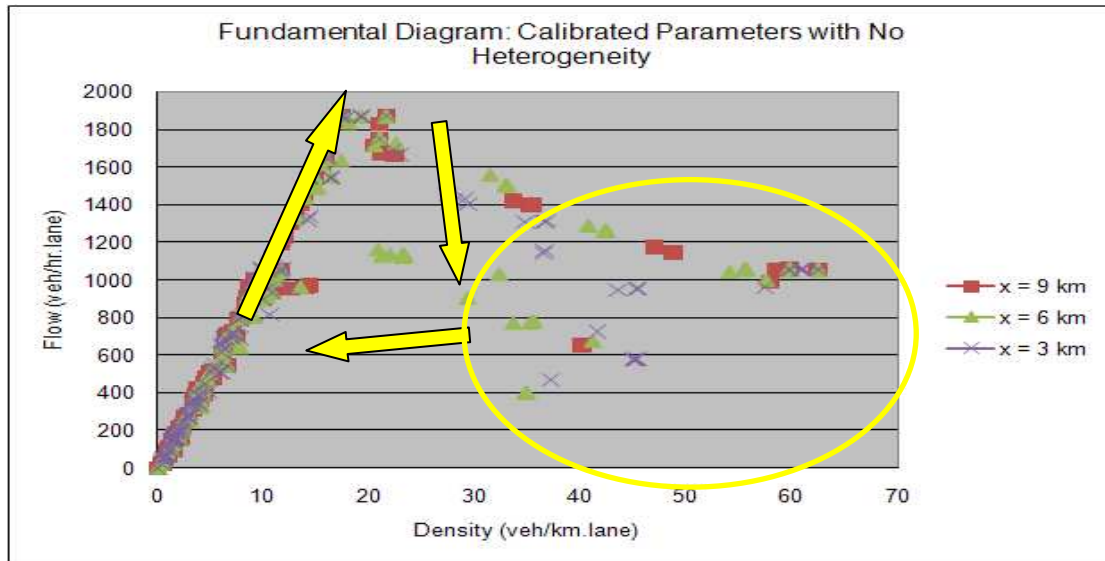


Figure 8-a: Flow-density relation. The results are based on virtual one-minute detectors at the indicated locations.

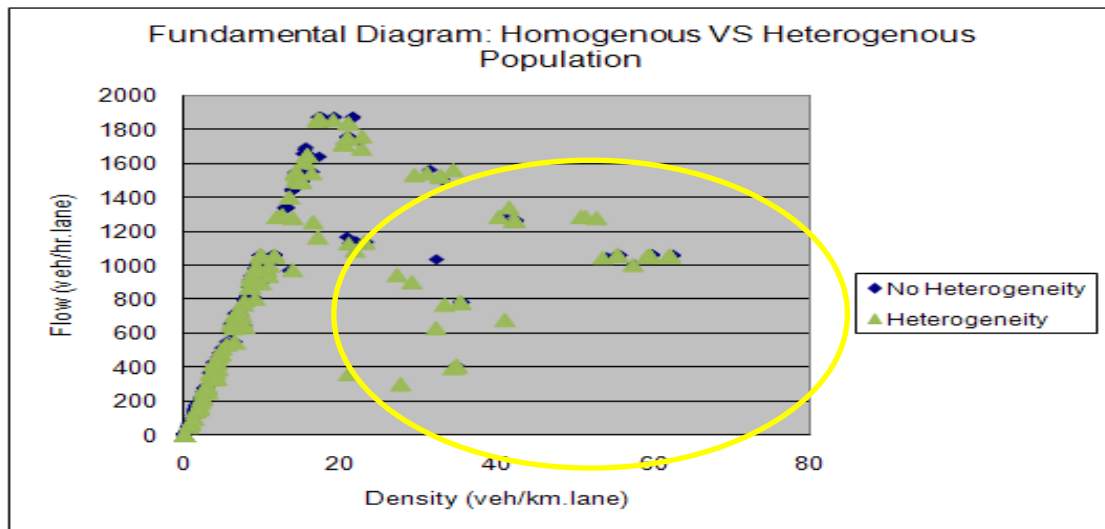


Figure 8-b: Flow-density relation with and without heterogeneity

6.3. Crash Investigation

Further extensive sensitivity analysis is performed to test the effect the parameters values and heterogeneity on the crash creation and distribution. It should be noted that for the simulation runs performed using the calibrated model parameters, no collisions were observed. When changing the parameter values and the corresponding distributions, collisions did occur. The basic results are shown in Table 8:

- There are two types of scenarios: with heterogeneity and without heterogeneity. For the scenarios with heterogeneity, we assume, for each parameter, symmetrically truncated Gaussians defined by the mean, the standard deviation and the range. When changed, the mean, standard deviation and the range are given in the second and the fourth columns of Table 8. As for the “no-heterogeneity” scenarios, only the mean value is used in the simulation as a deterministic parameter for all vehicles. In such cases, for each scenario, one parameter value is changed (as specified in the second and the fourth columns of Table 8).
- The non-changed parameter values are those found in the calibration exercise for data set 1. In the case of allowed “Inter-Driver Heterogeneity” (See Table 8), the corresponding distribution is assumed to Gaussian. The mean is considered to be the peak value observed in Figure 6, the range is taken to be the range that allows for only realistic parameter values to be obtained – i.e. defined by the minimum and the maximum values presented in Table 2, and the standard deviation is assumed to be equal to the standard deviation provided in Table 2. In the case of “Constant Change – All Vehicles” (i.e. homogenous traffic - See Table 8), the rest of the parameter values are set to be equal to peak values observed in Figure 6.
- Once two vehicles collide, simplified post-collision dynamics are assumed with the two impacted vehicles having aligned centers of gravities (no “off-set” allowed): the two vehicles decelerate at the maximum deceleration rate until reaching complete stop. The remaining vehicles continue maneuvering around these collided vehicles in the blocked lane(s). The authors did not force the vehicles to disappear as this is the current practice in other microscopic traffic simulation software and this unrealistic treatment is exactly what the authors are trying to avoid.

Incidents are created while there is an interplay between the weight parameters (W_m , W_+ and W_c) and the time parameters (τ = anticipation time and RT = reaction time). Regarding the weight parameters, the relative decrease of W_m (a lesser value than $0.098 < 1$) with respect to W_+ ($= 1$ initially) starts producing incidents. On the other hand, when increasing W_+ to 2 (until we have a linear value function with $W_m = W_+$), no incidents are created even when tailgating is favored and higher throughput is observed: the incident creation in rear-end collisions is related to a deceleration behavior rather than an acceleration behavior. Finally, as observed in the homogeneous case (2nd row), the relative weight between W_m and W_+ contribute in creating the incident while the crash weight's (W_c) role seems to be producing the flow dynamics and instability; such instabilities (stop and go, tailgating) constitutes an encouraging environment for incident scenarios. On the other hand, reaction time and anticipation time contributes for the creation of incident in a different manner; when comparing the reaction time scenarios in the heterogeneous traffic versus the homogeneous traffic, it is the heterogeneity in the reaction time that produced the highest number of incidents. This may suggest the role of correlation between two successive vehicles in stabilizing traffic conditions. As for the anticipation time, when low anticipation times are used or when a

high discrepancy between the anticipation times of two successive vehicles exists, the probability of incidents creation increases. Accordingly, inter-driver heterogeneity is a major aspect than needs to be understood when studying safety in vehicular traffic.

Table 8: Sensitivity Analysis Results: Heterogeneity VS Crash Distribution.

Bottleneck Scenario	Inter-Driver Heterogeneity			
	Inter-Driver Heterogeneity 1		Inter-Driver Heterogeneity 2	
Parameter	(mean, std, range)	# Accidents	(mean, std, range)	# Accidents
Wm	1, 0.5, 2	30	0.5, 0.5, 1	9
W+	1, 0.5, 2	0	2, 0.5, 2	0
Wc	100000, 10000, 40000	0	100, 50, 200	0
Beta Utility Uncertainty	5, 0.5, 2	0	100, 25, 100	
Tau	4, 0.5, 2 (1.3 sec)	0	4, 1.5, 7 (1.3 sec)	4
Reaction Time	2, 1, 3.8	249	1, 0.5, 0.9	0
Bottleneck Scenario	Constant Change - All Vehicles			
	Change 1		Change 2	
Parameter	Value	# Accidents	Value	# Accidents
Wm	1	0	3	0
W+	2	0	5	0
Wc	50000	0	500	0
Beta Utility Uncertainty	100	0	0.1	0
Tau	2	0	1	0
Reaction Time	5	27	1	0
	Change 3		Change 4	
	Value	# Accidents	Value	# Accidents
Wm	0.1	9	0.5	0
W+	20	0	100	0
Wc	5	0	0.5	0
Beta Utility Uncertainty	100000	0	0	0
Tau	1 (inter-arrival 0.1)	0	1 (1.3 seconds RT)	4
Reaction Time	1.3	0	2	0

7. CONCLUSIONS AND FUTURE RESEARCH NEEDS

In this paper, we propose a stochastic car-following model incorporating (a) common observations of traffic flow, (b) findings of prospect theory on the form of the perceived (subjective, generalized) utility, (c) risk taking, and (d) perception limits and subjective behavioral fluctuations. All model parameters can be attributed to these four “dimensions” as follows:

- a- The model produces an essentially triangular fundamental diagram characterized by the desired free-flow speed v_0 , the effective vehicle length l_{eff} composed of

the actual vehicle length and the minimum gaps s_0 , and a time gap resulting from the risk-related parameters via Equation 18. Furthermore, the model reproduces traffic flow instabilities and the often observed widely scattered congested flow-density data (Fig. 9).

- b- The prospect-theoretic subjective utility is modeled as a function of the objective utility by the parameters w_m (weighting of losses relative to gains), γ (weighting of utility differences far from the reference point relative to the sensitivity at the reference), and a_0 (the acceleration range that is considered “near the reference”). As in the original formulation of Kahneman and Twersky (1974), the formulation incorporates a power law of exponent γ . However, the suggested subjective utility (Equation 3.1) is smooth and has a finite gradient at the reference point. Both of these properties are desirable properties and allow an efficient model implementation. While such functions are often drawn in a qualitative way, an explicit function with these properties has never been proposed for microscopic traffic modeling.
- c- The risk-related aspects are characterized by the weighting w_c of crashes (the higher w_c is, the more conservative a driver is), the time horizon τ (the higher τ is, the more conservative the driver is since no reactions of the drivers are assumed in determining the crash risk over this time horizon), and the relative speed uncertainty α (the higher α is, the more uncertain a driver will be about the future speed of the lead vehicle and the more conservative he/she will be). The steady-state time gap emerges from these risk-related aspects as an increasing function of these parameters (Equation 18).
- d- The perception limits and subjective behavioral fluctuations are characterized by the intra-driver variability parameters β, τ_{corr} , and the reaction time.

There may be models with a smaller number of parameters accurately capturing traffic trajectory data, such as Gipps's model (Gipps, 1981) or the Intelligent Driver Model (IDM) (Treiber et al., 2000). However, the model suggested in this paper shows how general decision making tendencies/components influence the specific driving behavior which, consequently, may be approximated by “simpler” models. The presented formulation allows to relate each proposed parameter to fundamental behavioral components (risk-taking, uncertainty and perception limits) seen in different decision making/psychology theories (i.e. Prospect Theory).

To assess the properties of the proposed model, we have calibrated it on trajectory data and tested it for validity in terms of reproducing scattered flow-density data and realistic congestion dynamics. Based on the extensive numerical analysis, this study showed that the GA approach is suitable to calibrate car-following models with complex structures and capturing inter-driver heterogeneity. The study shows that the parameters related to prospect theory are estimated in accordance with this theory, i.e., losses are weighted higher than gains ($w_m > 1$), and the subjective sensitivity decreases far from the

reference point ($\gamma < 1$), cf. Table 1. Moreover, inter-driver heterogeneity generally outweighs intra-driver heterogeneity.

After testing the model for incidents creation, the presented utility-based structure seems to be more robust (cognitive aspect) than existing models when two vehicles collide (Hamdar and Mahmassani, 2008). Both individual and chain type accidents can be produced using weight parameters and time parameters. However, incident creation is not based on a simple relaxation of safety constraints since our model has no explicit safety criteria. Instead, the probability of crashes results from a complex interplay of the three risk-related parameters and the reaction time.

In future studies, this calibration exercise should be performed on different data sets where the inter-driver dynamics are recorded for longer durations and on a longer stretch of freeway. Calibration should be performed while considering intra-driver heterogeneity and parameters inter-correlation. Finally, the incident-related sensitivity analysis should be generalized to include fixed-object crashes, and the results need to be calibrated with real-life incidents scenarios.

8. ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 0927138. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

- 1- Ahmed, K. I., 1999, *Modeling Driver's Acceleration and Lane Changing Behaviors*, Ph.D. Dissertation, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts.
- 2- Brockfeld, E., R. D. Kuhne and P. Wagner, 2005, *Calibration and Validation of Microscopic Traffic Flow Models*, National Research Council (U.S.). Transportation Research Board Meeting (84th: 2005: Washington, D.C.). Preprint CD-ROM.
- 3- Chandler, R., R. Herman and W. Montroll, 1958, *Traffic Dynamics: Studies in Car-Following*, Operations Research 6, pp. 165-184.
- 4- Cheu, R. L., X. Jin, K. C. Ng, Y. L. Ng and D. Srinivasan, 1998, *Calibration of FRESIM for Singapore expressway using Genetic Algorithm*, Journal of Transportation Engineering, Vol. 124, No. 6, pp. 526-535.
- 5- Chu, L., H. X. Liu, J. S. Oh and W. Recker, 2004, *A Calibration Procedure for Microscopic Traffic Simulation*, National Research Council (U.S.). Transportation Research Board Meeting (83rd: 2004: Washington, D.C.). Preprint CD-ROM.
- 6- Gazis, D., R. Herman and R. B. Potts, 1959, *Car-Following Theory of Steady-State Traffic Flow*, Operations Research 7, pp. 499-505.
- 7- Gipps, P. G., 1981, *A Behavioral Car-Following Model for Computer Simulation*, Transportation Research 15B, pp. 101-115.

- 8- Federal Highway Administration (FHWA), 2004. *NGSIM Task E.1-1: Core Algorithms Assessment, Final Report*, Cambridge Systematic, Inc.
- 9- a) FHWA, 2005, *NGSIM I-80 Data Analysis (4:00 pm to 4:15 pm)*, Summary Report, Cambridge Systematic, Inc, Massachusetts.
- b) FHWA, 2005, *NGSIM I-80 Data Analysis (5:00 pm to 5:15 pm)*, Summary Report, Cambridge Systematic, Inc, Massachusetts.
- c) FHWA, 2005, *NGSIM I-80 Data Analysis (5:15 pm to 5:30 pm)*, Summary Report, Cambridge Systematic, Inc, Massachusetts.
- 10- Hamdar, S. H. and H. S. Mahmassani, 2008, *From Existing Accident-Free Car-Following Models to Colliding Vehicles: Exploration and Assessment*, Transportation Research Record, No. 2088, pp. 45-56, 2008.
- 11- Hamdar, S. H., M. Treiber, H. S. Mahmassani and A. Kesting, 2008, *Modeling Driver Behavior as a Sequential Risk Taking Task*, Transportation Research Record, No. 2088, pp. 208-217, 2008.
- 12- Hamdar, S. H., Treiber, M. and H. S. Mahmassani, 2009, *Calibration of a Stochastic Car Following Model Using Trajectory Data*. National Research Council (U.S.). Transportation Research Board Meeting (88th: 2009: Washington, D.C.). Preprint CD-ROM.
- 13- Herman, R., W. Montroll, R. B. Potts and R. W. Rothery, 1959, *Traffic Dynamics: Analysis of Stability in Car-Following*, Operations Research 7, pp. 86-106.
- 14- Hourdakis, J., P. G. Michalopoulos and J. Kottommannil, 2002, *A Practical Procedure for Calibrating Microscopic Traffic Simulation Models*, National Research Council (U.S.). Transportation Research Board Meeting (82nd: 2002: Washington, D.C.). Preprint CD-ROM.
- 15- Kesting, A., M. Treiber and D. Helbing, 2007, *MOBIL: General Lane Changing Model for Car-Following Models*, Transportation Research Record, No. 1999, pp. 86-94.
- 16- Lu, G., B. Cheng, Y. Wang and Q. Lin, 2013, *A Car Following Model Based on Quantified Homeostatic Risk Perception*, Journal of Mathematic Problems in Engineering, Volume 2013, <http://dx.doi.org/10.1155/2013/408756>.
- 17- Treiber, M. and A. Kesting, 2013, *Traffic Flow Dynamics – Data, Models, and Simulation*, Springer.
- 18- National Highway Traffic Safety Administration (NHTSA), 2007, *2005 Traffic Safety Facts Annual Report*. Washington, DC: DOT HS 810 631.
- 19- Ossen, S. and S. Hoogendoorn, 2007, *Driver Heterogeneity in Car-Following and Its Impact on Modeling Traffic Dynamics*, Transportation Research Record, No.1999, pp. 95-103.
- 20- Punzo, P., B. Ciuffo and M. Montanino, 2012, *Can Results of Car-Following Model Calibration Based on Trajectory Data Be Trusted?*, Transportation Research Record, No. 2315, pp. 11-24.
- 21- Thiemann, C., M. Treiber and A. Kesting, 2008, *Estimating Acceleration and Lane-Changing Dynamics Based on NGSIM Trajectory Data*, National Research Council (U.S.). Transportation Research Board Meeting (87th: 2008: Washington, D.C.). Preprint CD-ROM.

- 1
2
3
4 22- Treiber M, K. Hennecke and D. Helbing, 2000, *Congested Traffic States in*
5 *Empirical Observations and Microscopic Simulations*, Physical Review E,
6 Volume 2, No. 2, pp. 1805-1824.
7
8 23- Treiber, M. and A. Kesting, 2013, *Traffic Flow Dynamics: Data, Models and*
9 *Simulation*, Springer.
10 24- Treiber, M., A. Kesting and D. Helbing, 2006, *Delays, Inaccuracies and*
11 *Anticipation in Microscopic Traffic Models*, Physica A 360, pp. 71-88.
12
13 25- Wiedemann R. and U. Reiter, 1992, *Microscopic Traffic Simulation, the*
14 *Simulation System Mission*, Project ICARUS (V1052) Final Report. Brussels,
15 CEC.
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65