

Summer 9-18-2018

Driver Distraction in Microsimulation of a Mid-Block Pedestrian Crossing

Darryl Joseph Michaud
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds



Part of the [Transportation Engineering Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Michaud, Darryl Joseph, "Driver Distraction in Microsimulation of a Mid-Block Pedestrian Crossing" (2018). *Dissertations and Theses*. Paper 4553.
<https://doi.org/10.15760/etd.6438>

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Driver Distraction in Microsimulation of a Mid-Block Pedestrian Crossing

by

Darryl Joseph Michaud

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science
in
Civil and Environmental Engineering

Thesis Committee:
Miguel A. Figliozi, Chair
Christopher Monsere
Avinash Unnikrishnan
Robert Bertini

Portland State University
2018

©Darryl Joseph Michaud 2018

ABSTRACT

Traffic simulation has become an invaluable part of the traffic engineering toolbox. However, the majority of driver models are designed to recreate traffic performance based on interactions among vehicles. In keeping with this pursuit, most are fundamentally built to avoid collisions. This limits the applicability of using these models for addressing safety concerns, especially those regarding pedestrian safety performance. However, by explicitly including some of the sources of human error, these limitations can, in theory, be overcome. While much work has been done toward including these *human factors* in simulation platforms, one key aspect of human behavior has been largely ignored: driver distraction.

This work presents a novel approach to inclusion of driver distraction in a microsimulation or agent-based model. Distributions of distraction events and inter-distraction periods are derived from eye-glance data collected during naturalistic driving studies. The developed model of distraction is implemented – along with perception errors, visual obstructions, and driver reaction times – in a simulated mid-block pedestrian crossing.

The results of this simulation demonstrate that excluding any of these human factors from the implemented driver model significantly alters conflict rates observed in the simulation. This finding suggests that inclusion of human factors is important in any microsimulation platforms used to analyze pedestrian safety performance.

TABLE OF CONTENTS

Abstract	i
List of Tables	vi
List of Figures	viii
Glossary of Variables.....	x
Mathematical Symbols	x
Macroscopic Traffic Models.....	xi
IDM.....	xii
Yielding.....	xiii
Human Factors	xiv
Pedestrians.....	xv
Introduction	1
Motivation.....	1
Research Questions	2
Outline.....	3
Literature Review	5
Pedestrian Safety	5
Conflicts.....	6
Simulation	9

Distraction	13
Summary	19
Driver Modeling	21
Modeling Humans	23
Macroscopic Models	25
Microscopic Models	28
Intelligent Driver Model	31
Modeling Human Factors	40
Delayed Reaction	43
Imperfect Perception	54
Distraction	59
Driver Yielding	68
Perception of Pedestrians	71
Yielding Kinematics	73
Summary	74
Pedestrian Modeling	77
Generalized Force Model	79
Pedestrian gap acceptance	83
Summary	86

Agent-Based Modeling.....	88
Repast Symphony	90
Model Implementation	92
Road Environment	92
Scheduler	94
Drivers	96
Pedestrians.....	99
Visual Obstructions	102
Conflicts.....	104
Face Validation.....	105
Fundamental Diagram	105
Results.....	109
Significance of Including Human Factors	109
Baseline	111
Full Model	112
Human Factors Analysis	112
Summary	126
Other Agent Rates.....	127
Perception Distance	135

Conclusions	141
References	144
Appendix	172
Appendix A – Variable Values Used in Simulations	172
Appendix B – Macroscopic Equivalent to IDM	174

LIST OF TABLES

Table 1 – Empirical IDM Values	36
Table 2 – IDM Parameters Used in Simulation	38
Table 3 – Linear Regression of Speed Quantiles to Posted Speed Limit (PSL)	39
Table 4 – Pedestrian Parameters.....	81
Table 5 – Linear Regression: Baseline Model	111
Table 6 – Linear Regression: Full Model	112
Table 7 – Linear Regression: Estimation Errors Only	115
Table 8 – ANOVA: Estimation Errors vs. Baseline	115
Table 9 – Linear Regression: Full Model Excluding Estimation Errors.....	116
Table 10 – ANOVA: Estimation Errors vs. Full Model	116
Table 11 – Linear Regression: Visual Obstruction Only	117
Table 12 – ANOVA: Visual Obstruction vs. Baseline	118
Table 13 – Linear Regression: Full Model Excluding Visual Obstruction	118
Table 14 – ANOVA: Visual Obstruction vs. Full Model	119
Table 15 – Linear Regression: Reaction Times Only	120
Table 16 – ANOVA: Reaction Times vs. Baseline	120
Table 17 – Linear Regression: Full Model Excluding Reaction Times	121
Table 18 – ANOVA: Reaction Times vs. Full Model.....	122
Table 19 – Linear Regression: Distraction Only	123
Table 20 – ANOVA: Distraction vs. Baseline	123

Table 21 – Linear Regression: Full Model Excluding Distraction	124
Table 22 – ANOVA: Distraction vs. Full Model.....	124
Table 23 – Predicted Conflicts from Linear Models.....	125
Table 24 – Predicted Serious Conflicts from Linear Models.....	125
Table 25 – Predicted Collisions from Linear Models	125
Table 26 – Linear Regression: Full Model – PPH=400, VPH=800 (Speed \geq 30kph).....	127
Table 27 – Linear Regression: Full Model – PPH=300, VPH=800 (Speed \geq 30kph).....	127
Table 28 – Linear Regression: Full Model – PPH=200, VPH=800 (Speed \geq 30kph).....	128
Table 29 – Linear Regression: Full Model – PPH=400, VPH=600 (Speed \geq 30kph).....	128
Table 30 – Linear Regression: Full Model – PPH=400, VPH=400 (Speed \geq 30kph).....	128
Table 31 – Linear Regression: Full Model – PPH=400, VPH=200 (Speed \geq 30kph).....	128
Table 32 – Predicted Conflicts from Linear Models at Various Agent Rates.....	130
Table 33 – Predicted Serious Conflicts from Linear Models at Various Agent Rates	130
Table 34 – Predicted Collisions from Linear Models at Various Agent Rates.....	131
Table 35 – Linear Fits of Response Curves to Changes in Vehicle Volumes.....	134
Table 36 – Parameters Describing How Linear Fits to VPH Change with PSL.....	134

LIST OF FIGURES

Figure 1 – Safety Pyramid	6
Figure 2 – Greenshields’ Fundamental Diagram	26
Figure 3 – Shifted Weibull Fitting	53
Figure 4 – Preliminary Glance Distribution Fitting.....	62
Figure 5 – Glance Cdf Fitting	65
Figure 6 – Pedestrian Position	82
Figure 7 – Generated Pedestrian Gap Acceptance.....	85
Figure 8 – Modeled Crosswalk.....	93
Figure 9 – Simulation Schedule.....	94
Figure 10 – Multiple Threat	103
Figure 11 – Modeled Vehicle And Pedestrian	103
Figure 12 – Simulated Fundamental Diagram	108
Figure 13 – Empirical Fundamental Diagram.....	108
Figure 14 – Simulation Results: Baseline	111
Figure 15 – Simulation Results: Full Model.....	112
Figure 16 – Simulation Results: Estimation Errors vs. Baseline.....	115
Figure 17 – Simulation Results: Estimation Errors vs. Full Model	116
Figure 18 – Simulation Results: Visual Obstruction vs. Baseline	117
Figure 19 – Simulation Results: Visual Obstruction vs. Full Model	118
Figure 20 – Simulation Results: Reaction Times vs. Baseline	120

Figure 21 – Simulation Results: Reaction Times vs. Full Model.....	121
Figure 22 – Simulation Results: Distraction vs. Baseline	123
Figure 23 – Simulation Results: Distraction vs. Full Model.....	124
Figure 24 – Linear Response of Conflict Rates to Changes in Vehicle Volumes	132
Figure 25 – Linear Response of Serious Conflict Rates to Changes in Vehicle Volumes	132
Figure 26 – Linear Response of Collision Rates to Changes in Vehicle Volumes.....	133
Figure 27 – Model Results, Various Perception Distances	136
Figure 28 – Heatmap: PSL = 50kph, Various Perception Distances	137
Figure 29 – Heatmap: PSL = 70kph, Various Perception Distances	138
Figure 30 – Simulation Results Extended to High PSL, Perception Distance = 60m	139
Figure 31 – Simulation Results Extended to High PSL, Perception Distance = 100m.....	139

GLOSSARY OF VARIABLES

ABM – Agent-Based Modelling (defined on page 88)

ADRT – Acceleration/Deceleration Reaction Time (defined on page 44)

BRT – Brake Reaction Time (defined on page 47)

CAH – Constant Acceleration Heuristic (defined on page 34)

Conflict – Safety-critical interaction with $TTC \leq 1.5s$ (defined on page 6)

EIDM – Enhanced IDM (defined on page 33)

HDM – Human Driver Model (defined on page 43)

IDM – Intelligent Driver Model (defined on page 31)

NDS – Naturalistic Driving Study (defined on page 17)

PSL – Posted Speed Limit

Serious Conflict – Safety-critical interaction with $TTC \leq 1s$

MATHEMATICAL SYMBOLS

CDF – Cumulative Distribution Function (defined on page 48)

γ = Shift parameter

k = Shape parameter of the Weibull distribution

λ = Scale parameter of the Weibull distribution

$\ln(y)$ = Natural logarithm of some value y

M = Mean of some quantity or distribution

μ = Scale parameter of the lognormal distribution

η_i = Instances of a normal distribution with $M = 0$ and $S = 1$

PDF – Probability Density Function (defined on page 51)

S = Standard deviation of some quantity or distribution

σ = Shape parameter of the lognormal distribution

θ_w = Adjusted scale parameter of the Weibull distribution ($\theta_w = \lambda^k$)

U_i = Instances of a uniform distribution between 0 and 1

$\xi(t)$ = White noise

\hat{y} = Maximum-likelihood estimation of some variable y

y_s = Instance of a variable in a three-parameter (“shifted”) distribution

MACROSCOPIC TRAFFIC MODELS

n = Number of vehicles in a section of road

Φ = Net influx of vehicles along a stretch of road (e.g. from on-ramps and off-ramps)

q = Flow rate of vehicles – vehicles passing a stationary point per unit time

ρ = Macroscopic density of traffic flow – vehicles per unit length of road

V = Macroscopic velocity of traffic flow

often defined as the *harmonic mean* or *space mean* speed: $V = \frac{N}{\sum_{n=1}^N \left(\frac{1}{v_n}\right)}$

IDM

$a = \frac{dv}{dt} = \dot{v}$ = Acceleration

a_{ACC} = Adaptive cruise control acceleration – the incorporation of the Enhanced IDM and CAH into a functional car-following model

a_{CAH} = Acceleration given by the constant acceleration heuristic

a_{free} = Free-flow acceleration of EIDM

a_{max} = Maximum comfortable acceleration

b = Maximum comfortable deceleration (in absolute value)

δ = Free acceleration exponent – a description of how acceleration changes as the desired speed is approached

l = Subscript designating a variable applies to the front vehicle of a leader-follower pair

s = *Netto* space headway: the space between the rear bumper of the front vehicle and the front bumper of the rear vehicle of a leader-follower pair

s_0 = Jam headway – the space left in front of a vehicle when traffic comes to a standstill

s_1 = Non-linear headway term

s^* = Preferred headway

T = Preferred *netto* time-headway: $T = s/v$

τ_{react} = Driver reaction time

Δv = Velocity difference of a leader-follower pair: $\Delta v = v_f - v_l$

v = Velocity

v_f = Maximum, or *free-flow*, velocity

$$Z = \frac{s^*}{s}$$

YIELDING

a_{hy} = Acceleration needed to come to a full stop before the crosswalk – a “hard yield”

a_{sy} = Deceleration necessary for all pedestrians to be out of harm’s way by the time a vehicle reaches the crosswalk, yet not requiring the vehicle to come to a full stop – a “soft yield”

d_{stop} = Distance from edge of crosswalk at which a vehicle stops

t_{hy} = Time to come to a full stop with an acceleration of a_{hy}

T_{safe} = Time until a pedestrian will be safely beyond the path of an approaching vehicle

TTC = The time until two moving bodies would collide if each were to continue along its current path with its current speed – “time to collision”

TTC_{min} = The minimum TTC reached during an interaction between two road users

HUMAN FACTORS

D_i = Duration of an individual distraction event

d_{perc} = Distance from a pedestrian at which a driver can perceive that the pedestrian is present and intending to cross

ε = Variation coefficient, representing the relative standard deviation of a perceived value from the true value

ID_i = Duration of an individual inter-distraction period

λ_{ID} = Arrival rate of distraction events: $\lambda_{ID} = \frac{1}{\langle ID_i \rangle}$

ω = Rate of relative angular change: $\omega = \frac{d\theta/dt}{\theta}$

σ_r = Standard deviation of relative approach rate: a scaling parameter for errors in estimation of relative velocity

$\tilde{\tau}$ = Persistence time of perception errors

τ_{ADRT} = Acceleration/deceleration reaction time

τ_{BRT} = Brake perception reaction time

τ_{DR} = Device response time of the braking system

τ_M = Movement time of driver's foot between accelerator and brake pedal

τ_{RT} = Total brake reaction time ($\tau_{RT} = \tau_{BRT} + \tau_M + \tau_{DR}$)

θ = Apparent optical angle subtended by vehicle ahead

V_s = Statistical variation coefficient: the relative standard deviation of s^{est} from the true headway

$w_s(t)$ and $w_v(t)$ = Stochastic portions of Weiner processes (for headway and relative velocity, respectively)

w_{veh} = Vehicle width

y^{est} = Driver or pedestrian estimate of some quantity y

PEDESTRIANS

A and B = Empirically-derived constants for pedestrian interaction

c_1 and c_2 = Constants describing distribution of pedestrian critical gaps

d_{ij} = Scalar distance between pedestrian centers-of-mass: $d_{ij} = \|\mathbf{r}_i - \mathbf{r}_j\|$

d_{iW} = Scalar distance from pedestrian to barrier

\mathbf{e}^0 = Normalized (length = 1) pointing vector towards pedestrian destination

\mathbf{f}_{ij} = Interaction force vector pointing from pedestrian j to pedestrian i

\mathbf{f}_{iW} = Interaction force vector pointing from a barrier toward pedestrian i

m = Pedestrian mass

\mathbf{n}_{ij} = Normalized pointing vector between interacting pedestrians: $\mathbf{n}_{ij} = (\mathbf{r}_i - \mathbf{r}_j)/d_{ij}$

\mathbf{n}_{iW} = Normalized pointing vector away from barrier

p_{gap} = Probability of a pedestrian accepting a gap or lag

π = Probability of a pedestrian rejecting a gap or lag: $\pi = 1 - p_{gap}$

r = Pedestrian radius

\mathbf{r}_i = Vector location of pedestrian i

r_{ij} = Sum of interacting pedestrian radii: $r_{ij} = r_i + r_j$

τ_{ped} = Characteristic time of pedestrian acceleration

v = Speed

v^0 = Natural walking speed

INTRODUCTION

MOTIVATION

There has been a shift in public opinion and city planning toward livable communities, wherein walking and public transit are increasingly seen as attractive transportation modes. The U.S. Census Bureau's American Community Survey began collecting data on transportation modes in 2005. That year, an estimated 3.3 million Americans reported walking as their primary mode of commuting to work. By 2013, that number had risen to nearly 4 million. [GAO 2015] This increase in pedestrian volumes is accompanied by increased exposure of pedestrians to dangerous encounters with motor vehicles. Pedestrians do not have the safety benefits of traveling within a multi-ton protective metal housing; and, therefore, are more vulnerable than drivers to injury or death in the event of a collision.

5,987 pedestrians were killed in traffic crashes in 2016 – a 9% increase in fatalities from 2015 – making 2016 the deadliest year for pedestrians since 1990. [NHTSA 2018] While the overall number of traffic fatalities decreased by 17.8% between 2006 and 2015, the number of pedestrian fatalities increased by more than 12% over the same ten-year period. Pedestrians now make up 15% of all road fatalities, the highest proportion in the history of FARS data. [NHTSA 2017]

Given the rarity of pedestrian collisions and considering the ethical concerns of case-control experiments, simulation provides an invaluable tool in studying pedestrian

safety. The Federal Highway Administration (FHWA) commissioned a report on using simulations to study traffic safety in 2001. [Gettman and Head 2003] However, that report found that “[s]imulation of pedestrians’ movements, awareness of pedestrians by vehicles, and vehicle-pedestrian interactions are not as well developed as vehicle-vehicle model components in available traffic simulation models... Only a percentage of conflict events between pedestrians and vehicles are because of ‘normal’ driving and pedestrian behaviors (jaywalking and **mid-block pedestrian crossings are not typically modeled**). Also, sight-distance restrictions and **driver distractions play a large part in conflict events between pedestrians and vehicles...** These elements are not modeled in current traffic simulations, but should be an important part of future work in traffic simulation modeling.” [emphasis added] This remains an outstanding issue, as even recent attempts at comparing the results of traffic conflict studies to microsimulation often explicitly exclude pedestrian interactions. [e.g. Ambros, Turek, and Paukert 2014] Incorporating these two factors (mid-block crossings and driver distraction) into a simulation is the main focus of this work.

RESEARCH QUESTIONS

This work explores the limitations of modeling pedestrian safety without inclusion of human factors in the applied driver model. Since the focus is on demonstrating the limitations of existing simulation models, a purpose-built agent-based model of a mid-block pedestrian crossing was developed.

The driver model used is an extension of the Human Driver Model (HDM). [Treiber et al. 2006] The HDM is, itself, an extension of the Intelligent Driver Model [Treiber et al. 2000] to include the human limitations of imperfect perception and delayed reaction. In addition to these human factors, the model presented herein includes distraction and obstructed lines of sight. It relies on distributions from previous research to assign values of various parameters to agents, then allows those agents to interact with one-another based on the rules defined. The rate of pedestrian-vehicle conflicts (explicitly defined in later sections) is analyzed to quantify the effects of excluding the various human factors from simulation.

OUTLINE

This work begins with a summary of the results from a review of previous research. Broadly, the topics covered are pedestrian safety and traffic conflict analysis, modeling of vehicle-pedestrian interactions, and modeling of human factors (specifically distraction) in microsimulation.

The next two sections are concerned with driver and pedestrian models. After brief taxonomies of existing models, the details of the models used in this work are presented.

In the following section, agent-based models are discussed. Then, the specifics of the simulation environment created are outlined. This includes the details of agent calculations and cognition, as well as interactions among agents and classes of agents.

The final sections present the results of the developed simulation and conclusions drawn from the analysis, with implications for future research.

LITERATURE REVIEW

PEDESTRIAN SAFETY

On average, in the decade between 2006 and 2015, 4,666 pedestrians died every year in traffic crashes. [NHTSA 2017] However, studying the circumstances that lead to pedestrian collisions can be difficult. As with many traffic safety issues, ethical and practical constraints preclude running controlled experiments in the usual sense. Historical data are often the only sources of information available, and police reports often lack the detailed information needed to reconstruct an accident in a meaningful way. [Schaap 2012] Beyond the fact that not all accidents are reported, the “level of reporting is unevenly distributed with regard to type of road users involved. Vulnerable road users are, for instance, heavily under-represented in the police accident statistics compared to what hospital registrations and other studies show.” [OECD 1998] Crash statistics are, therefore, useful in validation but not in model creation.

Another fundamental limitation to studying pedestrian safety is that, despite the high likelihood of injury in the event of a crash, such crashes are rare events – on the order of a few per century at a given location.¹ It is, therefore, often useful to study pedestrian-vehicle *conflicts*.

¹ In a study of 2,000 uncontrolled crossing locations in the United States [Zegeer et al 2005], the average rate observed was “one pedestrian crash per crosswalk site every 43.7 years.”

CONFLICTS

In traffic conflict analysis, a methodology proposed more than 50 years ago,² [Perkins and Harris 1967] the precise definition of a *conflict* is an area of active debate. [Zheng et al. 2014] A conflict can be qualitatively, yet effectively, defined as “an observable situation in which two or more road users approach each other in space and time to such an extent that there is a risk of collision if their movements remain unchanged.” [Amundsen and Hydén 1977]

Road-user interactions can be conceptualized as belonging to a portion of the pyramid in Figure 1. The volume of a section in this pyramid represents, qualitatively, the frequency

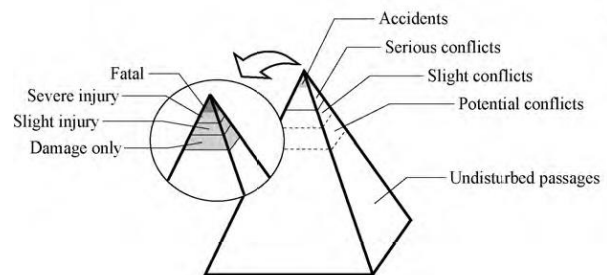


FIGURE 1 – SAFETY PYRAMID
[LAURESHYN 2010, ADAPTED FROM HYDÉN 1987]

with which each level of interaction occurs. The severity of interactions is inversely proportional to how often they occur. It has been demonstrated that “traffic conflicts... produce estimates of average accident rates nearly as accurate, and just as precise, as those produced from historical accident data.” [Migletz, Glauz, and Bauer 1985]

A variety of metrics for describing the severity of a conflict have been proposed. Perkins and Harris [1967] defined conflicts as events requiring evasive action to avoid a

² Though the topic was being explored a decade earlier [Forbes 1957]

collision. A subsequent study [Campbell and King 1970] was, however, unable to demonstrate a correlation between conflicts defined in this manner and collisions (though the sample size was admittedly small). In an attempt to further quantify the severity of a conflict, Hayward [1971, 1972] proposed the measurement of *time-to-collision* (TTC): “the time required for two vehicles to collide if they continue at their present speed and on the same path.” Hayward used a value of 1 second to define a vehicle-vehicle conflict; yet, “threshold values used appear to be mainly intuitively determined instead of based on systematic research.” [van der Horst 1990]

If the involved parties do not change course, but still avoid collision by only a fraction of a second, the TTC value is infinite – despite the fact that a collision very nearly occurred. This limitation led Allen, Shin, and Cooper [1978] to define the *post-encroachment time* (PET) of a conflict: If *encroachment time* is defined as the period in which a vehicle infringes on the travel path of another vehicle, the PET “is identified as the time from the end of encroachment to the time that the through vehicle actually arrives at the potential point of collision.” While this method can perform better than TTC for vehicle-vehicle conflicts, its applicability is limited since the value is undefined if one of the actors comes to a full stop. [Grayson et al. 1984]

Collisions are decisive events – they either occur or not. Conflicts, on the other hand, are more qualitative in nature. In an effort to eliminate any ambiguity around the topic, an international study was convened. Researchers from ten countries gathered in Malmö, Sweden to observe traffic at three urban intersections with mixed traffic.

[Grayson et al. 1984] While recording video of the intersections, the researchers tallied interactions they considered as constituting conflicts and compared results. The various measures of conflict severity were then tabulated for these interactions. [van der Horst 1984] Of the 337 conflicts analyzed in detail, 72 (~21%) were conflicts between cars and pedestrians. The type of road user involved in the interactions was not found to significantly affect the TTC thresholds for researchers deeming them as conflicts – all types had median threshold values near 1.5 seconds. [van der Horst 1990] The overall mean for the minimum-TTC (TTC_{min}) of an interaction deemed a *conflict* was 1.53 seconds. The median was 1.52 seconds.

In a later analysis of the Malmö study, [van der Horst 1990] various extensions of the TTC concept were compared in their accuracy for identifying interactions regarded as *conflicts* by observers:

- *Minimum-TTC* – TTC_{min} , the lowest TTC value observed during the interaction
- *Minimum-acceleration* – the greatest (in absolute value) deceleration observed by either road user
- *Minimum-distance* – the minimum distance between two road users during the course of an interaction
- *Minimum-TTC-distance* – the distance between road users at the moment of *minimum-TTC*
- *Minimum-TTCA* – the minimum TTC value observed, assuming a constant acceleration from the moment of measurement
- *Minimum-TTCA-distance* – the distance between road users at the moment of *minimum-TTCA*

- *TTC-brake* – TTC at the onset of braking (defined as the moment deceleration exceeds -0.5 m/s^2)
- *PET* – post-encroachment time (defined above)

Of these, TTC_{min} was observed to have the strongest correlation to interactions deemed as conflicts: “Based on these results, it can be concluded that, first of all, the TTC_{min} measure is an important variable in discriminating between normal and critical encounters. Furthermore... a minimum value of 1.5 s appears to be of prime importance.” This finding is doubly important, since it has been effectively argued that information on TTC is more directly available to a driver than explicit knowledge of velocity or distance. [Lee 1976; Janssen, Michon, and Harvey 1976]

Efforts have been made to automate some of the process of tabulating conflicts. [e.g. Ismail, et al. 2014] However, data acquisition and reduction remains costly and time-intensive. Further, this approach is, necessarily, reactive in nature: Changes in safety can only be quantified after infrastructure changes are made. Moving beyond this limitation requires accurate simulation of the pedestrian-vehicle interaction. This further necessitates accurate recreation of the human factors that lead to collisions.

SIMULATION

Transportation projects are generally enacted from perspectives on the larger end of the spectrum. [Batty 2001] Municipalities and regions are concerned with allowing people to move between distant locations. This approach is often most concerned with macroscopic measurements of traffic flow such as average speeds,

densities, and throughputs. While these metrics have correlations with safety performance, smaller-scale modeling is required to address the safety of specific sites. So-called *microscopic* models attempt to provide such a detailed view by mathematically representing the behaviors of individual road users. These models are implemented using *agent-based* simulations, in which individual actors with pre-defined characteristics interact with others in their environment based on explicit rules. This approach provides a means of studying the safety of road designs prior to implementation in the real world.

Some existing simulation platforms do take pedestrian conflicts with vehicles into account, but the models implemented focus on flow instead of conflicts. They also are rather difficult to adapt to specific research questions. [Bazzan and Klügl 2013] Of the software systems that model pedestrians at all, only three can accommodate detailed modeling. [FHWA 2004] These are VISSIM, PARAMICS, and DRACULA. The documentation for these is conspicuously lacking in principles or directions on the use of these capabilities. [Kittelson 2016]

VISSIM is the most capable of accepting user-defined behavior algorithms. [Rouphail and Chae 2002] In VISSIM, vehicle-pedestrian interactions are governed by “priority” rules, in which given percentages of each user will decide to yield to other user types when critical gap measures allow. [Chae 2005] The critical gap assignment to pedestrians is constant for each pedestrian population, so the only method of assigning distributed values is through initializing multiple user-defined populations. [Chae 2005]

Further, the “priority” rules do not seem to be at all combined with the car-following models in place.

When the FHWA’s Surrogate Safety Assessment Model [Gettman et al. 2008] was applied to output from VISSIM, analysis found that “the VISSIM model underestimated the pedestrian-vehicle conflicts.” [Wu, Essam and Abou-Senna 2016] The authors found that the simulation underestimated conflicts by approximately 19%.

Many studies have attempted to quantify pedestrian risk using logistic regression to correlate crash data to various factors such as the local built environment [Dumbaugh, Li, and Joh 2013], mutual awareness [Roth, Flohr, and Gavrila 2016], socio-economic status and travel patterns [Elias and Shiftan 2014], or all of the above [Quistberg et al. 2015]. However, few have attempted to recreate these patterns in simulation.

Some pedestrian motion simulations include moving vehicles for pedestrians to avoid [e.g. Liu et al. 2017; Li, Qian, and Luo 2012]. However, the modeled vehicles are essentially unaware of nearby pedestrians. A simulation of pedestrians and vehicles interacting in a shared space has been developed. [Anvari et al. 2015] In it, pedestrians and drivers avoid one another based on mutually-repulsive forces and rule-based conflict resolutions. Safety considerations were not included in this model. A cellular automata-based model for congested pedestrian and vehicle networks has been proposed, [Zhang and Chang 2014] but conflicts are resolved using “competition

factors” calibrated to agree with observations, and human factors are not considered. Cellular automata have also been used to model pedestrian conflicts with vehicles at signalized intersections. [Li et al. 2012] Here, again, conflict rules are stripped to the essentials and human factors are not considered.

Two notable exceptions of simulations that address pedestrian-vehicle interactions were found. The first [Zheng et al. 2015] is a simulation of jaywalking pedestrians and driver reactions to them. This model features detailed modeling of pedestrian gap-acceptance and crossing speed. It also explores driver yielding behavior in terms of yield rates and decision distances based on instrumented vehicle observations.

The second is SAFEPED – a three-dimensional multi-agent simulation of multi-modal urban environments. [Waizman, Shoval, and Benenson 2015] It features realistic limitations in pedestrian and vehicle movements, as well as visual obstructions between agents. It focuses on motion planning and obstacle avoidance based on a robotic motion-planning algorithm developed by Fiorini and Schiller [1998]. The SAFEPED model incorporates steering and pedestrian evasive action in the simulation and adds parked cars to obscure agent vision at crosswalks. It does explicitly include reaction times, but these are assumed to be uniformly distributed within a predefined range. In neither example are further human factors such as perception errors or distraction considered.

DISTRACTION

In a detailed analysis of vehicle-pedestrian interactions from a naturalistic driving study [Habibovic et al. 2013], the authors found three main causation patterns for safety critical pedestrian incidents: obstructed line of sight, distraction, and unexpected pedestrian behavior. Despite its salience in traffic safety, the published literature includes very few methods of simulating distraction.

Saifuzzaman and Zhang [2015] undertook a review of the literature on human factors in car-following models. The authors concluded that “in our extensive literature review we observed very few experiments designed for obtaining human factors critical for car-following modelling.” Further, they found that, in microscopic simulation software packages, “many human factors which are crucial for describing human car-following behavior are, by and large, ignored.”

It should be noted that VISSIM does, in fact, already have much of the programmatic architecture needed to implement distraction for safety-critical interactions. One of the many parameters available for manipulation is the “sleep” parameter (or “temporary lack of attention”). [PTV 2011] This parameter is used to lower the capacity of congested links and is ignored in cases that include “emergency braking.” Further, no indication is given that this “sleep” state can be used to affect interactions with other modes. Quadstone Paramics also features an “awareness” parameter. However, this parameter seems to affect the distance at which drivers become aware of information from their environment. [Hidas 2005]

Some researchers have developed methods of included human factors in driving behavior. Cody et al. [2008] explored the impacts of task-specific driver distraction on lane-change behavior. Most authors have focused on the effects of human factors on car-following models. Bevrani and Chung [2012] extended the Gipps car-following model to include reaction time, errors in detecting speed differences, and imperfections in drivers' speed adjustments. Treiber et al. [2006] extended their earlier work on the Intelligent Driver Model to include delayed reaction and errors in driver perception of distance and relative velocity. Neither of these models, however, include distraction. Xin et al. [2008] implemented a driver-specific "scanning interval" (the update interval for each agent's perception of traffic around her) that interacts with driver perception limitations to create an effective distracted state. The appropriate range of these values was found by calibrating the results to vehicle trajectory data. A model named HUTSIM [Archer and Kosonen 2000] uses a similar approach, adding random delays to the update frequencies of driver awareness.

Saifuzzaman et al. [2015] developed a method of incorporating Fuller's task-capability interface model [Fuller 2005] into car-following equations. The central hypothesis of Fuller's model is that drivers adjust their behavior to maintain "task difficulty homeostasis." In the car-following model developed by Saifuzzaman et al., this is accomplished by varying the drivers' free-flow speeds, acceleration patterns, and following distances based on a representation of the mental effort and perceived risk of the current driving environment. Distraction is implicitly involved in the "risk

compensatory parameter,” and explicitly added as a “reaction time increment.” Values for these parameters were fit using a genetic algorithm to measured trajectory data.

An interesting approach to the problem is the development of so-called *cognitive architectures* to recreate the human driver from the ground up. A foundational example of this is the Adaptive Control of Thought-Rational (ACT-R) model. [Salvucci 2006] ACT-R combines the “underlying psychological theories that incorporate basic properties and limitation of the human system” with a task analysis of vehicle control to simulate the act of driving. Another example is the Queueing Network-Model Human Processor developed by Liu, Feyen, and Tsimhoni [2006] for simulating driving while multi-tasking. These approaches have great promise in analyzing the effects of engaging in specific tasks while driving, but generalization to the full range of human multi-tasking activity is not the intended purpose.

One notable model that explicitly models distraction as an isolated phenomenon is the “errorable” car-following model developed by Yang and Peng [2009]. The authors present a car-following model that includes explicit representation of reaction times, perception errors, and distraction. Without a generative model for distraction, they indirectly characterized its duration and prevalence by attributing to it any significant deviations between measured car-following trajectory data and the predictions of their model. The authors found both distraction events and the periods between them to be approximately lognormal. The duration of individual distracted states, D_i , had a mean duration and standard deviation of 1.85 seconds and 0.89 seconds, respectively; while

the duration of individual inter-distraction periods, ID_i , had a mean and standard deviation of 26.62 and 0.75 seconds, respectively. These data stand out in that they represent a reproducible model of distraction events; however, they fundamentally rely on the accuracy of the car-following model presented instead of empirical data on distraction.

The omission of human factors, including distraction, from simulation platforms is due to a dearth of models for quantifying their occurrence in a form suitable to simulation. Distraction takes innumerable forms when diverting a driver's attention from the road. There are, therefore, a plethora of approaches for quantifying its prevalence. Many studies have been performed that focus exclusively on the distractions involved in cell phone use. For a meta-analysis of cell phone distraction to drivers, see [Caird et al. 2008].

While cell phone use represents only one of the many distracting tasks a driver may engage in, it provides a starting place for modeling distraction. The average increase in reaction time found in the meta-analysis by Caird et al. [2008] was used to represent distracted states by Przybyla et al. [2012]. Specifically, during a distracted state, reaction times (τ_{BRT}) were increased by a random sampling from a normal distribution with a mean of 0.36 seconds and standard deviation of 0.42 seconds. The prevalence of these distracted states was fit by calculating predicted numbers of crashes and fitting these values to observed counts.

One of the more extensive studies of the details of driver distraction is a detailed simulator study of the durations and effects on performance of twenty-two selected distracting activities. [Angell et al. 2006] The report provides a wealth of data on the detrimental effects of the selected activities on driving performance, as well as detailed analysis on eye glance behavior and “object-and-event” detection. However, it did not attempt to quantify the prevalence of these behaviors in every day driving.

Data from laboratory studies and driving simulators are useful in determining the behavior around specific examples of distracting behaviors; however, they cannot provide insight into the overall prevalence of such actions in normal driving circumstances. If normal driving behavior is to be modeled, naturalistic driving studies (NDSs – those that involve passive observation of drivers during normal daily driving) provide the most appropriate data for examining distraction. [Hurts, Angell, and Perez 2011]

Stutts et al. [2003] performed one such NDS. The authors used a collection of inward- and outward-facing cameras to observe 70 drivers for approximately three hours each. Videos were manually analyzed with distractions grouped into nine categories. The duration and frequency of each of the behaviors involved was tabulated. Due to the subjective nature of the behaviors under study, the agreement between coders as to when distractions began and ended “only reached about 65% or 70%.” Excluding conversing with passengers, the authors found drivers engage in these activities 16.1% of the time their vehicles are moving.

While no distributions were derived, and no cumulative patterns were presented by Stutts et al. [2003], their data were recently used to derive the only example similar to this work for providing a data-driven model for driver distraction that is suitable for microsimulation. The platform presented in that work [Lindorfer et al. 2018], begins by developing a “segregation matrix” for each simulated driver that represents, based on observed percentages, whether a particular driver will ever engage in each of the distracting behaviors tabulated. From there, each type of distraction is generated with Poisson arrival patterns derived from the overall average proportion of time it was observed in the NDS. These distraction events are then randomly assigned to drivers who are not currently otherwise distracted, and whose segregation matrices show that they are susceptible to that type of distraction.

Time proportions are quoted as aggregated means and standard deviations in [Stutts et al. 2003], leaving the authors to assume a log-normal distribution based on the observations of Yang and Peng [2009]. The method of moments is used to fit distraction duration (D_i) distributions because the first two central moments of the empirical distribution presented in the data source are insufficient for “more sophisticated types of distribution fitting such as e.g. maximum likelihood estimation.” [Lindorfer et al. 2018]

Beyond the above assumptions, the major limitation of this approach is that it is based on data that required extensive manual reduction in a method that is not likely to be repeated. The data source used in the model developed herein relies on eye-glance

analysis from an NDS. [Dingus et al. 2006] The data used were manually coded, but a variety of methods have been developed for automatically recognizing distraction based on eye movements and driving performance. [e.g. Liang, Lee, and Reyes 2007]

SUMMARY

The limitations of using historic crash data for studying pedestrian safety necessitated the creation of surrogate safety measures. Among these is the concept of a *conflict*, which occurs when two road users' current trajectories put them on a collision course. While this is somewhat of a subjective concept, it can be quantified using a variety of metrics.

Time-to-collision (TTC) is a promising metric that has received significant attention. TTC represents how long in the future a collision would occur between two parties, given their current velocities. More specifically, the minimum TTC that occurs over the course of an interaction has been shown to be a good surrogate to crash statistics for quantifying safety performance. For vehicle-pedestrian interactions, a minimum TTC value below 1.5 seconds is generally considered to represent a *conflict*. While also subjective, a further distinction can be made that a minimum TTC below 1.0 second represents a *serious conflict*. [Kraay, van der Horst, and Oppe 1986]

Conflict data is time- and resource-intensive to collect and analyze. It is also only available after a project is completed. Therefore, simulation is an essential tool in traffic safety analysis. While microscopic and agent-based traffic simulations are numerous for

vehicular or pedestrian behavior in isolation, the interaction between the two user groups has received less attention. Two notable exceptions are [Zheng et al. 2015] and SAFEPEd. [Waizman, Shoval, and Benenson 2015]

Where vehicle-pedestrian interactions are considered, the dearth of explicit models and specific empirical data make simplifying assumptions necessary. Among the topics generally over-simplified in simulation are the limitations of human drivers. These human factors are, however, at the foundation of the breakdowns that make the road a dangerous setting. Distraction stands out as a human factor that is largely absent from models of driver behavior suitable for use in simulation.

Yang and Peng [2009] presented the first exception to this gap in the literature. Their “errorable” driver model derived a distribution of distraction from observed deviations from an applied car-following model. Recently, Lindorfer et al. [2018] presented another, based on detailed decomposition of distracting behaviors from a prior naturalistic driving study. Both works represent valuable contributions to the field. However, both have their limitations. Use of the former has its foundation in the assumption that any significant deviations from expected driving behavior are attributable to distraction. The latter is based on a single, specialized set of data that is unlikely to be repeated. The work presented herein presents another approach, based on eye-glance behavior data from a large naturalistic driving study. Data of this type are commonly used for a variety of purposes and are relatively easy to collect. These features make eye-glance data attractive as a basis for modeling distraction.

DRIVER MODELING

Despite the startling statistics of casualties on the road, collisions are rare events at any given location. Modeling is therefore often a necessary tool for analyzing the potential for danger. However, after decades of active research, a universally accurate and robust model of road user behavior remains an elusive target. This is not surprising, as human behavior results from an amalgam of complex internal and interactive processes, many of which are still not fully understood. This makes the problem a difficult one; but it also means there are many angles from which the task can be approached. Traffic models have been developed based on the work of cognitive scientists, physicists, psychologists, and computer scientists – each discipline adding to our understanding of the patterns we see.

Fundamentally, driving can be broken down into three levels of control: strategic (trip planning), tactical (maneuvering), and operational (basic skills). [Michon 1971] Car-following and lane-keeping would be tactical, while the actual movement of the pedals and steering wheel are operational. These realms generally interact in a top-down manner: Decisions at the strategic level set the stage for tactical decisions, which define the necessary operational tasks. This hierarchy can break down, however, if circumstances in the road environment interrupt decisions made at higher levels. Many subtasks can be handled subconsciously, but when the unexpected happens, the driver must bring the task involved into conscious control. [Reason et al. 1990]

With such a broad and diverse landscape under review, it is perhaps best to begin with a fundamental truism and build from there: Road users move in two dimensions. Within this constraint, pedestrians generally have the freedom to move in any direction they choose, whereas motor vehicles are typically confined to lanes. This makes it possible and convenient to decompose vehicular motion on a straight road into orthogonal components: longitudinal motion along a lane and transverse motion within and between lanes. It is therefore not surprising that driver models predate and far outnumber pedestrian models. As mentioned above, driving is a means of getting from one place to another, so it is also not surprising that the majority of these driver models focus mainly on the longitudinal component.

Within the confines of this simulation there are no traffic signals or changing speed limits, so longitudinal movement is simply a combination of car-following behavior and pedestrian yielding. The lateral movement of driver-car units on a straight road (under normal circumstances) consists of lane-choice and position within the lane. In this model, lane changing is ignored in the interest of simplicity. The added mental workload of changing lanes makes this an important area for further research, but the complexity is outside the scope of this research. Exact position within the lane would have second-order effects on yielding behavior and gap acceptance, but its effect was deemed insignificant in the current exploration. Further, while lane-changes far in advance of a crosswalk would constitute a form of yielding, no research could be found on the prevalence of this behavior. Finally, if a driver deviates from strictly forward

movement to avoid a pedestrian at the last moment, that would constitute a conflict. The presence of this conflict is all that is being investigated; and its aftermath is, again, beyond the scope of the model presented.

MODELING HUMANS

Beyond modeling patterns in the physical behavior of drivers, researchers have developed theoretical cognitive models to posit potential thought processes behind these patterns. In the realm of driver behavior, there are two broad categories of cognitive models (though they are not mutually exclusive) [Shinar 2007a]: those designed to describe the human actor in terms of cognitive limitations (e.g. Wickens' [1984] "bottleneck" of limited attention and information processing resources) and those that attempt to describe the strategies drivers adopt to accomplish their goals. While these theories do not attempt to describe the full cognitive capacities of their subjects, they present testable frameworks into which the results of cognition can be parameterized and, potentially, predicted. "Absent the theories, it is almost impossible to specify what new countermeasures might emerge. Thus, what is a standard operating procedure for many human factors researchers (using models) might require an act of faith from practicing highway engineers who do not normally invoke theories of human behavior." [Kantowitz et al. 2004]

Among the earliest of these reductionist theories of cognition (wherein complex behaviors are broken down into component parts) is the assertion that "the mind is made up of certain sub-forces – the perception, the intellect, the emotions, and the

will.” [Hammond 1876] This formulation of the mind (referred to as PIEV, after changing the word *will* to *volition*) is the basis of a later architecture for modeling human practical reasoning within software: the belief/desire/intention model (BDI). [Bratman 1987] In this model, agents within a software program have a set of knowledge about their surroundings – their “beliefs,” and a set of goals they are trying to fulfill – “intentions.” This terminology demonstrates the contributions of the approach: agents (and the humans they represent) act based on the information they have, even if this differs from the ground truth; and their actions are intended to achieve an ultimate goal – as opposed to traditional programming, which is instead rigidly task-oriented. Based on its *beliefs* and *desires*, a software agent performs actions chosen from among a set of plans. These are its *intentions* – the behaviors chosen.

One fundamental cognitive model is based on the theory that decision makers are subconsciously maximizing an equation that balances the outcomes of available actions – the so-called *expected utility theorem*. [Neumann and Morgenstern 1944] An early adaptation of this approach, *prospect theory*, instead considers a weighted utility equation of the potential risks and rewards of possible courses of action. [Kahneman and Tversky 1979] Another cognitive model asserts that drivers alter their behavior to maintain a comfortable level of anxiety while driving. [Taylor 1964] This was later adapted to argue that drivers adjust their behavior to stabilize their subjective estimates of the probability of a collision – *risk homeostasis*. [Wilde 1982] More recently, it was suggested [Fuller 2005] that what drivers attempt to maintain is a stable level of task

difficulty, adjusting their speed and behavior to adapt to more or less challenging scenarios: this is the *task-capability interface*.

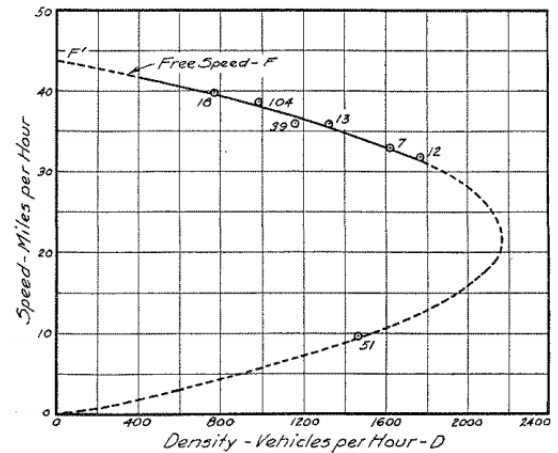
Many attempts have been made to base behavior models on such cognitive theories, bridging the gap between the engineering and psychological approaches. The Atomic Components of Thought (or “Adaptive Control of Thought-Rational” – ACT-R) constitutes a ground-up approach to modeling driver behavior. [Salvucci 2005] It is a cognitive architecture model that utilizes theories of the mind to enable exploration of how specific secondary tasks are accomplished alongside a main activity such as driving. Hamdar, Mahmassani, and Treiber [2014] developed another model that intrinsically incorporates driver uncertainty into a prospect utility maximization framework. Saifuzzaman, et al. [2015] proposed yet another based on Fuller’s task-capability interface. Absil [2008] proposed a behavior model specifically for use in microsimulation; however, the author cited the lack of an “accurate enough model of a human being for distraction to be properly implemented.” Reichardt [2008] even developed a driver model that attempts to recreate driver emotional responses to traffic situations, and how these changes affect risk acceptance behavior.

MACROSCOPIC MODELS

All longitudinal driver models are essentially derived from the fundamental relation of traffic [Wageningen-Kessels et al. 2015], which describes the intrinsic relationship between the density (ρ) of vehicular traffic and the speed (V) of that traffic.

This correlation was first presented by Greenshields [1935], who derived a parabolic speed-flow relationship. Some of his data are shown in Figure 2.

From this fundamental relationship between density (or, equivalently, *headway* – the space between cars – *s*) and speed, two general approaches for describing traffic can be pursued. First, the *macroscopic* approach aggregates the motion of vehicles into average properties



**FIGURE 2 –
GREENSHIELDS' [1935] FUNDAMENTAL DIAGRAM**

that describe traffic flow in terms associated with fluid dynamics, typically using partial differential equations (PDEs). The mathematics of these macroscopic traffic dynamics have been a topic of active research for decades (e.g. the LWR model [Lighthill & Whitham 1955; Richards 1956] or gas-kinetic models [Prigogine & Herman 1971]). These models explore the relationships among the three essential variables of traffic flow: *density* (ρ), *velocity* (V), and *flow rate* (q – vehicles passing a point per unit time):

$$q = V\rho \quad [1]$$

Along with the definition of flow (Equation [1]), macroscopic models typically include a representation of the conservation of vehicles.³ Put simply, the change in the

³ The derivation here loosely follows that in Chapter 2 of [Kachroo et al 2008]

number of vehicles in a road segment between two points, x_1 and x_2 , is given by the number of cars entering the segment, minus the number leaving the segment, plus the net influx of vehicles from on- and off-ramps. We can represent this latter influx over a distance, dx , as $\Phi(x, t)dx$. Since, under normal circumstances, no vehicles are created or destroyed on a road segment, the number of cars within the segment is given by

$$n = \int_{x_1}^{x_2} \rho(x, t) dx \quad [2]$$

The change in n is then given by $\frac{dn}{dt} = q_{in} - q_{out} + \int_{x_1}^{x_2} \Phi(x, t) dx$. If we allow x_1 and x_2 to vary with time, this becomes a partial derivative, and the integral form of the conservation law becomes

$$\frac{\partial}{\partial t} \int_{x_1}^{x_2} \rho(x, t) dx = q_{in} - q_{out} + \int_{x_1}^{x_2} \Phi(x, t) dx \quad [3]$$

Now, by the fundamental theorem of calculus,

$$q_{in} - q_{out} = - \int_{x_1}^{x_2} \frac{\partial}{\partial x} q(x, t) dx \quad [4]$$

Substituting this into Equation [3], we have

$$\int_{x_1}^{x_2} \left[\frac{\partial}{\partial t} \rho(x, t) + \frac{\partial}{\partial x} q(x, t) - \Phi(x, t) \right] dx = 0 \quad [5]$$

Since the only quantity that equals 0 regardless of the endpoints over which it is integrated is the 0-function, we can set the quantity within the square brackets to 0; and we have the *continuity equation*:

$$\frac{\partial \rho}{\partial t} + \frac{\partial q}{\partial x} = \Phi \quad [6]$$

To constrain a system of three variables, we need three equations. For most formulations, the third relationship is an equation for the average speed of traffic. One foundational approach, the Lighthill-Whitham model [Lighthill & Whitham 1955], assumes the equilibrium velocity to be a function of density: $V(x, t) = V_e(\rho(x, t))$. Even with such simplifying assumptions, many of the collective, self-organized phenomena of instabilities and nonlinear interactions in traffic flow can be examined from this perspective; and much progress has been made through this approach toward comprehending the behavior of traffic as a whole.

MICROSCOPIC MODELS

The alternative approach recognizes that traffic patterns result from countless decisions and actions taken by individuals, based on their surroundings as well as their personal attributes and current states. While the large number of actors means that general characteristics of these patterns can be recreated from a macroscopic perspective, the details are obscured. The aim of a *microscopic* approach, on the other hand, is to enable examination of individual interactions in detail. Micro-simulation models of traffic number in the hundreds [Brockfeld et al. 2003], each with its own strengths and weaknesses. However, broadly speaking, there only a few general categories of microscopic modeling.

The earliest microscopic models were so-called *safe-distance* models, such as the Pipes [1953] and Gipps [1981] models, which express a driver's (the *follower's*) behavior in terms of her speed and the distance to the vehicle ahead (the *leader*). The General Motors family of car-following models generalize many of these "follow-the-leader" systems. [Gazis, Herman, and Rothery 1961] Another approach, *action-point* models (first formalized by Wiedemann [1974]), utilize thresholds for psychophysical determination (a driver's ability to perceive changes in the relative motion between herself and the car ahead) to define various regimes, each associated with an appropriate acceleration response.

Two newer categories of microscopic models have been developed. The first of which is referred to as *cellular-automata* (CA) models. In these approaches, (first introduced by Cremer and Ludwig [1986], but most famously implemented by Nagel and Schreckenberg [1992]) as opposed to the forms presented previously, space is discretized into *cells*. Each cell is either occupied by one vehicle or empty, and a list of rules determines when each automaton will move, and into which cell. An even newer category of models was proposed in the 1990s [Kikuchi & Chakroborty 1992], called *fuzzy logic* models. These models attempt to recreate the vague heuristics of driver decisions in a natural language-based format, instead of the deterministic equations of other approaches. To accomplish this, logical operators are applied to variables that take values between 0 and 1 based on the state variables of the local traffic.

The final general category of microscopic conceptualizations of vehicle traffic are termed *stimulus-response* models. In these models, a follower's acceleration, $a = \frac{dv}{dt} = \dot{v}$, is determined by her own velocity, v ; her current headway, s ; and the speed of the car ahead (or, more explicitly, the difference between her speed and that of the leading car), $\Delta v = v_f - v_l = \dot{s}$.⁴ A general form of stimulus-response models, incorporating the driver's reaction time, τ_{react} ,⁵ was published very early [Chandler et al. 1958]:

$$a(t) = f(v(t), s(t - \tau_{react}), \dot{s}(t - \tau_{react})) \quad [7]$$

The broad array of approaches should not be surprising. Driving is a complex behavior, and (so far) no driver model has been able to capture all the behaviors of real drivers. To complicate matters more, it has been shown that different models work better for different drivers. [Ossen and Hoogendoorn 2010] In other words, not only do drivers have responses of different magnitudes to the cars ahead of them, they appear to actually be responding to different stimuli. While this fact is troublesome for any detailed modeling of traffic flow, car-following behavior is not the primary objective of this exploration.

When 13 of the most prominent models were compared in terms of how well they fit the behavior of drivers on a one-lane road, two models tied for the best performance: the Nagel and Schreckenberg CA model [1992] and a stimulus-response

⁴ All variables in the following derivation are applied to the following car, so the subscript is dropped for clarity.

⁵ This is equivalent to the τ_{ADRT} of later sections.

model called the Intelligent Driver Model. [Treiber et al. 2000] The latter was chosen for the current simulation due to its strong performance, ease of implementation, and the fact that each of the model's parameters has a reasonable interpretation and is empirically measurable. [Kesting and Treiber 2008a] Further, it is one of the simpler developed models in terms of number of parameters. This is a good thing, as a review of car-following models found that "complex models likely [do] not produce better results." [Brockfeld et al. 2004] In the words of one of the authors of the IDM, "a model for a real system should be as simple as possible, but not simpler." [Helbing et al. 2002]

INTELLIGENT DRIVER MODEL

The Intelligent Driver Model (IDM) is a time-continuous stimulus-response car following model. [Treiber et al. 2000] A macroscopic equivalent of the model, developed by the authors, [Helbing et al. 2002] is presented in Appendix B. In the original formulation of the IDM, the behavior of a single leading vehicle, l , determines that of the following vehicle. A driver's change in speed, $\frac{dv}{dt} = \dot{v}$, at any time $t + \Delta t$ is determined by the situation at time t : namely, her speed, v ; the speed of the car ahead (or, more explicitly, the difference between her speed and that of the leading car, $\Delta v = v - v_l$); and the current headway, s ,⁶ by the following equations:

⁶ Headways and time gaps in the IDM are *netto* gaps (Italian for *net*), meaning they measure the distance from the front bumper of a following car to the rear bumper of its leader. This is in contrast to *brutto* gaps (Italian for *gross*), which measure front-bumper to front-bumper distances.

$$\begin{aligned}\dot{v} &= a_{max} \left(1 - \left(\frac{v}{v_f} \right)^\delta - \left(\frac{s^*}{s} \right)^2 \right) \\ s^* &= s_0 + \max \left(0, s_1 \sqrt{\frac{v}{v_f}} + vT + \frac{v\Delta v}{2\sqrt{a_{max}b}} \right)\end{aligned}\quad [8]$$

The above equations demonstrate that the IDM is a combination of an open-road strategy: $\dot{v} = a_{max} \left[1 - \left(\frac{v}{v_f} \right)^\delta \right]$, and a deceleration strategy $\dot{v} = -a_{max} \left(\frac{s^*}{s} \right)^2$ that only becomes relevant if the headway is close to the *preferred headway*, s^* .⁷ While these parameters can vary based on a driver's external situation and internal state, they are defined as follows: [Kesting et al. 2009]

- v_f – maximum, or *free-flow*, speed
- T – preferred time-headway ($T = s/v$)
- s_0 – jam headway, the space left in front when traffic comes to a standstill
- s_1 – non-linear headway term. This parameter is set to 0 in most analyses of the model.
- a_{max} – maximum comfortable acceleration
- b – maximum comfortable deceleration (in absolute value)
- δ – a description of how acceleration changes as the desired speed is approached. $\delta = 0$ yields a linear acceleration curve. $\delta \rightarrow \infty$ corresponds to always accelerating at a_{max} . A value of δ close to 4 agrees with empirical data.

⁷ With a large enough Δv , the equations would possibly produce a safe headway shorter than the jam headway. This is avoided with the $\max()$ function in the calculation of s^* .

In its original formulation, the IDM has unrealistically large decelerations when the speed is above the desired free-flow speed. The remedy for this limitation, as devised by the authors of the IDM [Treiber and Kesting 2013a, Ch 11], is a piecewise acceleration equation, referred to as the Enhanced Intelligent Driver Model (EIDM):

$$a_{free}(v) = \begin{cases} a_{max} \left[1 - \left(\frac{v}{v_f} \right)^\delta \right] & \text{if } v < v_f \\ -b \left[1 - \left(\frac{v_f}{v} \right)^{\delta \cdot a_{max}/b} \right] & \text{if } v \geq v_f \end{cases} \quad [9]$$

While there are no changes in speed limit to create such a situation in this model, the a_{free} term defined above is useful in correcting the next deficiency – the fact that gaps in platoons traveling near the desired speed, v_f , become greater than s^* , and the parameter T loses its meaning. This can be avoided by distinguishing between the cases of $s \leq s^*(v, \Delta v)$ and $s > s^*(v, \Delta v)$, while insisting that \dot{v} remains continuously differentiable at the transition point:⁸

$$\begin{aligned} \left. \frac{dv}{dt} \right|_{v < v_f} &= \begin{cases} a_{max}(1 - z^2) & \text{if } z = \frac{s^*(v, \Delta v)}{s} \geq 1 \\ a_{free}(1 - z^{2a/a_{free}}) & \text{otherwise} \end{cases} \\ \left. \frac{dv}{dt} \right|_{v \geq v_f} &= \begin{cases} a_{free} + a_{max}(1 - z^2) & \text{if } z \geq 1 \\ a_{free} & \text{otherwise} \end{cases} \end{aligned} \quad [10]$$

⁸ In the original formulation of the IIDM [Treiber and Kesting 2013a], the cases of Equations [9] and [10] were $v \leq v_f$ and $v > v_f$. This lead to division by 0 in the case of $s = s^*$. Through correspondence with the authors, this issue was resolved by adjusting the ranges to those stated here.

Though less extreme, the IDM also produces overreactions when short headways occur between two cars going nearly the same speed. [Kesting, Treiber, and Helbing 2010] Assume, for example, a driver moves from one lane into another whose vehicles are traveling at nearly the same speed. Using the IDM, the sudden drop in headway that follows this move would lead the following car to decelerate drastically. Real drivers, on the other hand, typically would not. This unrealistic behavior can be avoided by enforcing the following rule (subscript l represents the lead vehicle):

$$a_{CAH} = \begin{cases} \frac{v^2 \tilde{a}_l}{v_l^2 - 2s\tilde{a}_l} & \text{if } v_l(v - v_l) \leq -2s\tilde{a}_l \\ \tilde{a}_l - \frac{(v - v_l)^2 \theta(v - v_l)}{2s} & \text{otherwise} \end{cases} \quad [11]$$

Here, $\tilde{a}_l = \min(a_l, a)$, represents the *effective acceleration*, or the lower of the pair of vehicles; and $\theta(x)$ is the Heaviside step function, which is 0 for negative arguments and equal to 1 for positive arguments.⁹ While Equation [11] eliminates excessive decelerations in specific circumstances, it is not a full car-following model. It

⁹ The Heaviside step function, proposed by Oliver Heaviside [Heaviside 1892] is a useful tool in operational calculus; yet it does not necessarily have a meaning when the argument is equal to zero. Heaviside considered himself an explorer in the realms of physics, and considered strict mathematical rigor as stifling to the pioneering spirit. He preferred "the substitution of simpler and more direct processes for the indirect and complicated processes of the highly cultivated mathematician with too rigorous proclivities," going on to assert that "complaints of the want of perfection of the ways and manners of work of explorers on the part of men who are accustomed to more rigorous methods have a considerable element of the ludicrous in them." [Heaviside 1899] Since the numerator disappears when the argument of the function is zero anyway, this limitation is not disqualifying.

should therefore only be applied when the IDM produces unrealistically high values of deceleration. To enforce this, the resulting acceleration (a_{ACC} , since it was derived by the authors as a model for an *adaptive cruise control* system) that is used in the simulation is defined by:

$$a_{ACC} = \begin{cases} a_{IDM} & \text{if } a_{IDM} \geq a_{CAH} \\ (1 - c)a_{IDM} + c \left[a_{CAH} + b \tanh\left(\frac{a_{IDM} - a_{CAH}}{b}\right) \right] & \text{otherwise} \end{cases} \quad [12]$$

Here, c is a so-called “coolness” factor to define the quantity of the correction. $c = 0$ would revert to the IDM acceleration. The authors assign a value of $c = 0.99$, which will be used in the simulation.

PARAMETER VALUES

For those drivers who are best modeled by the IDM (a la Ossen and Hoogendoorn 2010), the appropriate values of these parameters vary widely between countries and regions, among drivers within a region, and even within individual drivers over time (e.g. “frustration effects”). [Treiber and Helbing 2003] There have been many attempts at calibrating them for small groups of drivers, using a variety of methods. Some of these results are presented in Table 1. Methods that rely on integrated measures such as travel time can average out crucial details such as oscillations; while those based on repeated microscopic measurements like headways are heavily influenced by inter- and intra-driver variability. [Treiber and Kesting 2010] However, the behavior of the IDM is rather robust to changes in the distributions of driver parameters. Deviations between observed and simulated behavior have been shown to

remain low even when using parameters averaged over several drivers. [Kesting 2008]

Further, sensitivity analysis of the parameters of the IDM [Ciuffo et al. 2014] shows that varying drivers' preferred time headway, T , can cover the full variability in vehicle speeds or positions in a car-following situation. In fact, the authors assert that T "is the only parameter that in any case needs calibration."

parameter	Treiber and Kesting 2013b	Ciuffo et al. 2014	Schakel et al. 2012	Punzo and Simonelli 2005		Kovács et al. 2016 ¹⁰	Kesting 2008	Kim and Mahmassani 2011	
				M	S			M	S
v_f (m/s)	16.1 ¹¹	20.58	34.4 ¹²	28.362	203.987	15.3 (fixed)	57	85.72	26.55
$T(s)$	1.2	0.54	1.2	0.69	0.046	0.86	1.29	1.266	0.507
$s_0(m)$	1.53	2.0	3 (fixed)	0.743	0.13	2.0 (arbitrary)	2.22	2.172	1.152
a_{max} (m/s ²)	1.39	2.83	1.25	2.568	0.619	1.62 (det'd by s_0 choice)	1.25	1.406	1.012
$b(m/s^2)$	0.65	2.92	2.09	1.694	0.493	2 (fixed)	1.57	2.225	1.849

TABLE 1 – EMPIRICAL IDM VALUES

For the simulation, the values reported by Kim and Mahmassani [2011] will be used to define driver behavior. This study was performed using data collected on a 500-

¹⁰ The authors measured these parameters for drivers at urban intersections.

¹¹ These data come from observations on an urban road, so *free-flow* speed is constrained by legal limits. The posted speed limit is not given in the publication, but German roads have a default speed limit of 50 km/s (≈ 13.9 m/s) in built-up areas.

¹² Data for this study were collected on the A20 freeway near Rotterdam, for which the speed limit is 120 kph (≈ 33.3 m/s).

meter stretch of California's Interstate-80 as part of the Next Generation SIMulation project. [NGSIM 2006] It included 465 follower-leader pairs traveling along a six-lane interstate. While the parameters quoted in Table 1 describe the distribution of the raw population, Kim and Mahmassani [2011] observed a distribution in values closer to lognormal for several of the variables. They do not quote the scale parameter and shape parameter for the distributions derived, but we can approximate the values they would have found by inverting the equations derived by Finney [1941] for estimating the mean and standard deviation of a lognormal distribution. For a transformed set of data, $x = \ln(y)$, that has a mean, μ , and standard deviation, σ , the parent population, y , would have mean, M , and standard deviation, S , given by:

$$M = e^{\mu + \frac{\sigma^2}{2}} ; \quad S^2 = e^{2(\mu + \sigma^2)} - e^{2\mu + \sigma^2} \quad [13]$$

We can invert these¹³ to give estimates for the shape and scale parameters of the transformed data set:

$$\mu = 2 \ln(M) - \frac{1}{2} \ln(S^2 + M^2)$$

$$\sigma^2 = \ln(S^2 + M^2) - 2 \ln(M)$$

Estimates derived using these equations are presented in Table 2. Also presented in Table 2 are the extremes of the values allowed in the simulation. Since values are

¹³ Following the methods of [Ginos 2009]

derived using distributions that are asymptotic to the x-axis, the occasional unrealistic value is generated. To avoid this, maxima and minima were assigned (based on reasonable real-world values); and any generated values outside these ranges are explicitly set to the relevant extreme value.

parameter	M	S	distribution	μ	σ	max	min
$T(s)$	1.266	0.507	normal			3.294 ($M + 4S$)	0.252 ($M - 2S$)
$s_0(m)$	2.172	1.152	lognormal	0.6517	0.4979	4.476 ($M + 2S$)	0.444 ($M - 1.5S$)
$a_{max} (m/s^2)$	1.406	1.012	lognormal	0.132	0.6461	5.454 ($M + 4S$)	0.394 ($M - S$)
$b(m/s^2)$	2.225	1.849	lognormal	0.5372	0.7246	7.4 ¹⁴	0.376 ($M - S$)

TABLE 2 – IDM PARAMETERS USED IN SIMULATION [ADAPTED FROM KIM AND MAHMASSANI 2011]

Returning to Table 1, the unrealistically-large desired speeds (v_f) of 57 m/s (≈ 211 kph) and 85.7 m/s (≈ 308 kph) from Kesting [2008] and Kim and Mahmassani [2011], respectively, can be interpreted as resulting from data being fit from observations of bound traffic. Since most drivers did not experience free-flow conditions, the fitting results cannot produce realistic values for this parameter. [Kesting 2008] A more realistic value is needed. However, the appropriate distribution requires some consideration.

¹⁴ This value is equal to the maximum deceleration physically possible by the average car on a dry road. [Greibe 2007] It is also quite close to the value AASHTO gives for emergency decelerations (4.5 m/s²)

Urban free-flow speeds (v_f) vary greatly,¹⁵ and are influenced by a litany of factors. For example, the Highway Capacity Manual's regression model for predicting free-flow speeds has nine inputs. [Moses and Mtoi 2013] Of these variables, the only one included in this simulation is the posted speed limit (PSL). While this is a significant simplification, it has been demonstrated [Fitzpatrick et al. 2001] that the PSL is the most significant factor affecting free-flow speeds.¹⁶ In a study of 104 sites around the United States, Fitzpatrick et al. [2003] found that a linear relationship between quantiles of driver free-flow speeds and the PSL described the data reasonably well. The equations found, as well as the resulting speeds for the median, 85th percentile, and 95th percentile are presented in Table 3

	Q ₅₀	Q ₈₅	Q ₉₅	Quantile
	$3.336 + 0.966 * \text{PSL}$	$7.675 + 0.98 * \text{PSL}$	$10.196 + 0.993 * \text{PSL}$	Linear fit
speed limit	0.911	0.901	0.879	Adjusted R ²
25	27.5	32.2	35.0	
30	32.3	37.1	40.0	
35	37.1	42.0	45.0	
40	42.0	46.9	49.9	
45	46.8	51.8	54.9	
50	51.6	56.7	59.8	
55	56.5	61.6	64.8	

**TABLE 3 – LINEAR REGRESSION OF SPEED QUANTILES TO POSTED SPEED LIMIT (PSL)
[FROM FITZPATRICK ET AL. 2003]**

¹⁵ This is further complicated by the fact that, for researchers attempting to characterize free-flow behavior, the definition of a vehicle in free-flow conditions is not a settled matter. [Vogel 2002]

¹⁶ The correlation found in [Fitzpatrick et al 2001] is between 85th percentile speeds and posted speed limit. Since the generally accepted method for determining speed limits is heavily influenced by 85th percentile speeds, the causal relationship here is questionable; however, this distinction does not reduce the applicability of the finding.

If we assume a Gaussian distribution for v_f , one standard deviation above the mean would represent approximately the 84th percentile, and two standard deviations above the mean would be approximately the 98th percentile. Based on the data in Table 3 a Gaussian distribution with a mean value 2mph ($\approx 3.22\text{kph}$) greater than the PSL¹⁷ and a standard deviation of 4.5mph ($\approx 7.24\text{kph}$) would fit the data rather well. These are then the parameters used to define the distribution of v_f in the simulation. As with the parameters derived from Kim and Mahmassani [2011], the use of a Gaussian distribution allows for unrealistic values of v_f from the tails of the distribution. To avoid this, any generated values more than 4 standard deviations greater or less than the mean will be fixed to $M + 4S$ or $M - 4S$, respectively.

MODELING HUMAN FACTORS

The task of driving involves three main categories of actors: drivers, their vehicles, and the surrounding environment (including e.g. the road, weather, and foreign objects). An influential model in the field of error management [Reason 1990] argues that errors occur when weaknesses in system-level safeguards coincide with natural human errors. This model (referred to as the *Swiss-cheese model*) posits that each of the defenses that a system employs against the natural variation in human behavior has weaknesses or holes in it. These defensive barriers stand between human

¹⁷ The data collected in [Fitzpatrick et al 2003] present the median, not the mean. However, for speed distributions, the difference between these two values tends to be very small. [Berry and Belmont 1951]

actors and any undesired consequences of those actions. In the metaphor from which this model derives its name, each of these safeguards is a slice of Swiss cheese with holes that are constantly in flux with changing circumstances. Sometimes, all the holes line up; and if a person makes a mistake under these circumstances, their error can lead to an undesired outcome – here, a crash.

Due to decades of research and improvements in road and vehicle design, many of the latent conditions that make traffic crashes more likely have been removed. Analysis of crash data suggests [Treat et al. 1979] that, in the vast majority of collisions, human error is to blame. By some estimates, driver error is responsible for as much as 94% of traffic crashes. [Singh 2015] This study makes no effort to examine the system-level factors that allow crashes to occur. It instead attempts to isolate some of the driver behaviors that can lead to such crashes.

The aberrant driver behaviors that can lead to a crash can be categorized as *violations* or *errors*. [Reason et al. 1990] Violations, according to Reason, are “deliberate (though not necessarily reprehensible) deviations from those practices believed necessary to maintain the safe operation of a potentially hazardous system;” while errors are “the failure of planned actions to achieve their intended consequences.”

Significant efforts have been made toward defining and quantifying the categories and effects of driver error; however, “universally accepted taxonomies of driver error and error causing conditions are yet to emerge, and most are beset by a

lack of validation evidence.” [AustRoads 2011] Most of the literature on errors is based on one of two frameworks: accident reports and analysis (along the lines of Hendricks et al. [2001]) or self-reported driver questionnaires (such as the Driver Behavior Questionnaire of Reason et al. [1990]). These exhibit a strong selection bias in that they only include errors that caused crashes (and only those crashes that were reported¹⁸).

The authors of the AustRoads report [2011] divide driver errors into three general categories: recognition/perception (e.g. distraction), decision (e.g. misjudged velocity), and performance (e.g. loss of control). They also include data on the relative frequencies of these categories. However, many of the paper's included errors are not necessarily dangerous and/or are outside the realm of this simulation. While this categorization is useful in modeling the effects of driver failure, the authors lament the dearth of research into “exactly what error causing failures exist across road transport systems, how they influence driver behaviour and which failures lead to what driving errors.” [AustRoads 2011]

The main drawback of using the Intelligent Driver Model [Treiber et al. 2000] (or any traditional car-following model) for the purposes of this analysis is that it is deterministic and, by design, accident-free. This limitation of car-following models can be overcome for some circumstances by inclusion of random noise. [e.g. Jost and Kai

¹⁸ In Oregon, reporting a crash is only mandated if damages exceed \$1500, or the crash causes injury or death.

2003] However, this is at best an imprecise approach to modeling the limitations of a human actor.

Human drivers are not perfect – they can be distracted, have finite reaction times and less-than-perfect perception of distance and relative velocity. One of the goals of this research is a comparison of the relative effects that these underlying limitations have on the driver failures that affect pedestrian safety. Inclusion of these human factors builds on the work of Treiber et al. [2006] and their Human Driver Model (HDM). Along with the delayed reactions and perception errors of the HDM, visual obstruction and a novel approach to modeling distraction is included.

DELAYED REACTION

Humans do not react instantaneously to stimuli. The brain must first recognize the input and decide on an appropriate response – this can be termed the *mental processing time* (τ_{BRT}). Second, during the *movement time* (τ_M), the body executes the selected course of action. Finally, the physical device being used takes time to perform the action – the *device response time* (τ_{DR}). [Green 2000]

While the effects of finite driver reaction times on the flow of traffic have been considered for decades [e.g. Chandler et al. 1958, Newell 1961], the quantification of the phenomenon is not a settled matter. Driver reactions to stimuli can vary based on many factors, both endogenous (e.g. driver age, gender, urgency, expectancy, and cognitive load [Green 2000]) and exogenous (e.g. headway [Kim and Zhang 2011], traffic

mix [Siuhi and Kaseko 2016], or even the color of the object eliciting the reaction [Lee, Lee, and Boyle 2009]). From among the diverse domain of driver reactions, two specific realms of reaction time can be easily delineated: delay in the acceleration and deceleration behavior of car-following, and delayed braking responses. [Mehmood and Easa 2009]

ACCELERATION/DECELERATION REACTION TIME

Acceleration/deceleration reaction times – *ADRT*, defined as delays in changes to vehicle speed effected through movement of the gas pedal – are shorter than braking reaction times (*BRTs*). This makes sense, as the behavior is a constant feedback loop of reaction to expected changes and requires very small movements. Finite *ADRT* values ($\tau_{ADRT} > 0$) are a realistic way of introducing instability into simulated traffic flows. [Kesting and Treiber 2008b] Interestingly, while non-zero *ADRT* adds instability to traffic flow on small scales, distributed values can improve large-scale platoon stability (when compared to simulations with the reaction times of all drivers being equal). [Treiber, Kesting, and Helbing 2007]

Since *ADRT* manifests in an essentially continuous manner, studies seeking to quantify it rely on genetic algorithms [e.g. Hamdar, Mahmassani, and Treiber 2013] or comparison of simulation to empirical behavior through statistical software. [e.g. Siuhi and Kaseko 2016] Since the actual acceleration is all that is measured in these studies, τ_{ADRT} implicitly includes any movement or device response delays.

Unfortunately, of the few studies that report values for τ_{ADRT} , none were found that sought to describe a distribution. Siuhi and Kaseko [2016] found acceleration reaction times to be 0.1 seconds slower, on average, than deceleration reaction times (0.8s and 0.7s, respectively). This makes sense, since deceleration is often a safety-critical behavior; however, this complexity was not included in this simulation, since the flow of traffic is only indirectly related to driver yielding behavior.

The values chosen instead come from Hamdar, Mahmassani, and Treiber [2013], who found a mean (M_{AD}) of 0.658857 seconds, with a standard deviation (S_{AD}) of 0.726583 seconds. The lognormal distribution was chosen for modeling τ_{ADRT} in this simulation since it is commonly observed in random processes that feature low average values, large variances, and are exclusively positive. [Limpert et al. 2001] However, the lognormal distribution allows for values approaching 0. Since humans have non-zero reaction times, a shifted (or *three-parameter*) lognormal distribution was used. The probability density of the three-parameter lognormal function can be expressed as

$$f(x|\mu, \sigma, \gamma) = \frac{1}{(x - \gamma)\sigma\sqrt{2\pi}} \exp \left[-\frac{(\ln(x - \gamma) - \mu)^2}{2\sigma^2} \right]$$

Here, γ is the *shift parameter*, which gives moves the minimum value away from 0. In a study seeking to decompose driver reaction times into their constituent parts [Lister 1950], an irreducible minimum *perception time* of 0.25 seconds was found. Therefore, γ_{AD} in this simulation equals 0.25 seconds. The mean, M , and standard

deviation, S , of a set of values that fit the three-parameter lognormal distribution are given by:

$$M = \gamma + e^{\mu + \frac{\sigma^2}{2}}$$

$$S^2 = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)$$

Rearranging and plugging in the values from Hamdar, et al. [2013]:

$$\mu_{AD} = \frac{1}{2} \ln \left[\frac{(M_{AD} - \gamma_{AD})^2}{1 + \left(\frac{S_{AD}}{M_{AD} - \gamma_{AD}} \right)^2} \right] \approx -1.60692$$

$$\sigma_{AD}^2 = \ln \left[1 + \left(\frac{S_{AD}}{M_{AD} - \gamma_{AD}} \right)^2 \right] \approx 1.42506$$

To implement this in the simulation, a number, η_i , is generated from a normal distribution for each driver using Java's built-in random number generator. Since this gives values from a Gaussian with mean of 0 and standard deviation of 1, the produced number is multiplied by σ_{AD} and added to μ_{AD} . Euler's number, e , is raised to the result; and this is added to γ_{AD} .¹⁹

$$\tau_{ADRT} = \gamma_{AD} + e^{\mu_{AD} + \eta_i \sigma_{AD}} \quad [14]$$

¹⁹ The lognormal distribution is asymptotic to 0 as the independent variable approaches infinity. Therefore, applying it here can lead to unrealistically large reaction times in rare cases. In both the datasets collected in [Hamdar, Mahmassani, and Treiber 2013], the histograms of reaction times derived using the authors' genetic algorithm have empty bins for values above 2.5 seconds (with a few outliers at values greater than this, up to 3.2 seconds). Since 2.5 seconds is already a value of ADRT that produces unrealistic behavior (speeds in excess of 2x the desired maximum), any values above this are set to 2.5 seconds.

This delay in driver reaction is implemented by means of a *delay-differential equation*.²⁰ [Bellman and Cooke 1963] At each time step of the model, the acceleration given by the IDM is calculated and stored. This is then recalled by each agent after her native reaction time, τ_{ADRT} , has passed. If $t - \tau_{ADRT}$ falls between two stored accelerations, linear interpolation is used to split the difference.

BRAKE REACTION TIME

The literature is replete with studies quantifying driver brake reaction times (τ_{BRT}) in various situations. In a review of 30 such studies, Green [2000] delineated some of the results. In his review, he found mean total brake reaction times (τ_{RT} – the sum of perception-reaction time – τ_{BRT} , movement time – τ_M , and device response time – τ_{DR}) of 0.7 to 0.75 seconds, 1.25 seconds, and 1.5 seconds for expected events, unexpected but common events, and surprise events, respectively. In an exploration of the effects of cognitive load on driver reaction times to pedestrian incursions, Lee et al. [2009] found a mean τ_{BRT} of 1.35 seconds in the baseline of driving without a secondary task. Another group [Laberge et al. 2004] found a mean τ_{BRT} of 1.31 seconds while exploring the effect of distracting conversations on driver reactions to pedestrian encroachment.²¹

²⁰ Originally referred to as “differential-difference equations.”

²¹ This is the mean of the values the authors found for their “easy” and “difficult” driving situations.

All the studies listed above, while useful in quantifying mean or median values, do not attempt to describe the observed distribution of reaction times. To simulate distributions in delayed reaction, a function must be chosen to fit the data. A few assumptions are required to choose an appropriate function: the function should be strictly positive, the survivor function²² should be monotonically decreasing, and the function should allow for asymmetry about the mean, as stated by Green [2000] “[reaction time] data are almost always skewed toward longer values.” This last criterion means that, the median observed reaction time is often lower than the mean (one of the few studies to report both values [Chang, Messer, and Santiago 1985] found mean values averaged approximately 0.2 seconds higher than median values). The simplest distribution that fits these requirements is the Weibull distribution.²³ Further, it has been observed that the standard deviation of reaction time distributions increases linearly with the mean. [Wagenmakers and Brown 2007] The Weibull distribution fits this criterion as well.²⁴ The cumulative distribution function (CDF – the probability that a random value chosen from the distribution will be less than or equal to a threshold, t) for the Weibull function is defined for all values of $t \geq 0$ as:

²² If the cumulative distribution function, $CDF(t) \equiv F(t)$, is the probability that a variable will take a value less than or equal to t , then the survivor function, $S(t) = 1 - F(t)$, is the probability that the value will be greater than t .

²³ The empirical mean and standard deviation values used to derive the ADRT distribution in the previous section could not be simultaneously fit by a shifted Weibull, given the chosen shift parameter – therefore, the three-parameter lognormal distribution was used instead.

²⁴ This relationship is in terms of changes to the scale parameter, λ .

$$F(t) = 1 - e^{-(t/\lambda)^k}, \quad f(t) = F'(t) = \frac{k}{\lambda^k} t^{k-1} e^{-(t/\lambda)^k} \quad [15]$$

where λ is the *scale parameter*, and k is the *shape parameter*. Haque and Washington [2014] observed driver perception reaction times (τ_{BRT}) to pedestrians approaching marked crosswalks (from the time the pedestrian was in view and started moving towards the crosswalk to the moment the driver released pressure on the accelerator) and fit the observed data to a Weibull distribution, finding average baseline values for λ and k of 1.47 and 3.043, respectively.²⁵

It should be noted that the authors include distraction in their measurements of reaction time. The numbers discussed here are from the baseline case (no distracting activities). Since the experiment was performed in a driving simulator, it is assumed these values represent the performance of a driver free from distraction. There is no way to ensure a subject is free from internal distractions, but it is not unreasonable to assume participants in a driving simulator experiment are focused on their task.

The Weibull distribution produces values approaching 0 seconds; which, due to the physical and cognitive processes involved, are impossible. Early research attempted to quantify the magnitudes of these constraints, and to establish an *irreducible minimum* brake reaction time. Two quality examples of such projects [Johansson and

²⁵ NB: a different parameterization of the Weibull distribution is used in this source, and the equation given for the survival function within it has a typo – in the source’s parameterization, the Weibull survivor function is $e^{-(\lambda t)^P}$

Rumar 1971; Lister 1950] found a value of 0.4 seconds. A third [Schweitzer et al. 1995] found minimum values ranging from 0.38 to 0.42 seconds, depending on drivers' level of preparedness for an emergency braking procedure. This also matches the minimum BRT found by Mehmood and Easa [2009]. Lister decomposed this value into a mental processing time of 0.25 seconds, and a movement time (τ_M) of 0.15 seconds. The values measured by Haque and Washington [2014] should include this mental processing time (τ_{BRT}), but not the movement time. The shifted, or *three-parameter*, Weibull function can be applied to produce a distribution that does not violate this constraint. Defining the Weibull shift parameter, γ_w , and (for convenience) an adjusted scale parameter, $\theta_w = \lambda^k$, this can be represented – similar to Equation [15] – by

$$F(t) = 1 - e^{-(t-\gamma_w)^{k_s}/\theta_{w,s}}, \quad f(t) = F'(t) = \frac{k_s}{\theta_{w,s}} (t - \gamma_w)^{k_s-1} e^{-(t-\gamma_w)^{k_s}/\theta_{w,s}} \quad [16]$$

The Weibull distribution is quite flexible. With a given shift parameter, appropriate shape and scale parameters can be found to produce a set of points that would be fit by the un-shifted distribution found by Haque and Washington [2014]. This is accomplished iteratively:

Step 1) Ranges of possible values for the shifted shape and scale parameters, k_s and λ_s , are chosen.

Step 2) For each, a shifted Weibull distribution is generated.

Step 3) The un-shifted parameters, k and λ , with the maximum likelihood of producing this data set from an un-shifted Weibull distribution are calculated.

The values of k_s and λ_s that produce values of k and λ closest to the values reported in Haque and Washington [2014] are chosen and used in the simulation. The methods of the maximum likelihood calculation are presented here.

For each point, the probability density function (PDF, the derivative of the CDF: $f(x) = F'(x)$) calculates the probability of measuring the given value. Thus, the product of all such probabilities is the *likelihood* of the data series arising from the distribution defined by the given parameters. We can write the density function of the un-shifted Weibull distribution as $f(x) = \frac{k}{\theta} x^{k-1} e^{-x^k/\theta}$, then the Likelihood function is [this derivation follows Cohen 1965]

$$L(X|k, \theta) = \prod_{i=1}^n \frac{k}{\theta} x_i^{k-1} e^{-x_i^k/\theta} \quad [17]$$

To find the maximum of an equation in two variables, we take the gradient and find where it is equal to 0. The variables can be isolated into partial derivatives, and each equation is set to zero separately: $\frac{\partial L}{\partial \theta} = 0$, and $\frac{\partial L}{\partial k} = 0$. In practice, it is often easier to take the natural logarithm of each side before taking these derivatives since the products can be decomposed into summations. The logarithmic function is defined on

the domain of the data and monotonically increasing, so this *log-likelihood* function will have maxima²⁶ in the same locations as Equation [17]. Thus, the log-likelihood is

$$\mathcal{L}(X|k, \theta) = \ln(L(X|k, \theta)) = n \ln\left(\frac{k}{\theta}\right) + (k-1) \sum_{i=1}^n \ln x_i - \frac{1}{\theta} \sum_{i=1}^n x_i^k$$

Taking the partial derivatives and setting each equal to 0:

$$\frac{\partial}{\partial \theta} \mathcal{L}(X|k, \theta) = -n + \frac{1}{\theta} \sum_{i=1}^n x_i^k = 0 \Rightarrow \theta = \frac{1}{n} \sum_{i=1}^n x_i^k \quad [18]$$

$$\frac{\partial}{\partial k} \mathcal{L}(X|k, \theta) = \frac{n}{k} + \sum_{i=1}^n \ln x_i - \frac{1}{\theta} \sum_{i=1}^n x_i^k \ln x_i = 0 \quad [19]$$

Substituting Equation [18] into Equation [19], we have

$$\frac{\sum_{i=1}^n x_i^k \ln x_i}{\sum_{i=1}^n x_i^k} - \frac{1}{k} = \frac{1}{n} \sum_{i=1}^n \ln x_i \quad [20]$$

For a set of points generated from each combination of shifted parameters (k_s and λ_s), Equation [20] is solved iteratively for k . Plugging these results into Equation [18] for $\theta = \lambda^k$, and comparing to the empirical values, we find $\lambda_s = 1.20$ and $k_s = 2.435$.

²⁶ Explicitly, this only finds local extrema. Proving that the locations derived are maxima requires finding the Hessian (the second derivative matrix) and proving that it is negative-definite. For a proof that this is the case here, see [Scholz 1996].

With this and $\gamma_w = 0.25s$, we have a shifted Weibull distribution that produces data that would be fit by the un-shifted Weibull parameters found by Haque and Washington. In Figure 3, data generated by an unshifted-Weibull having the values found by Haque and Washington

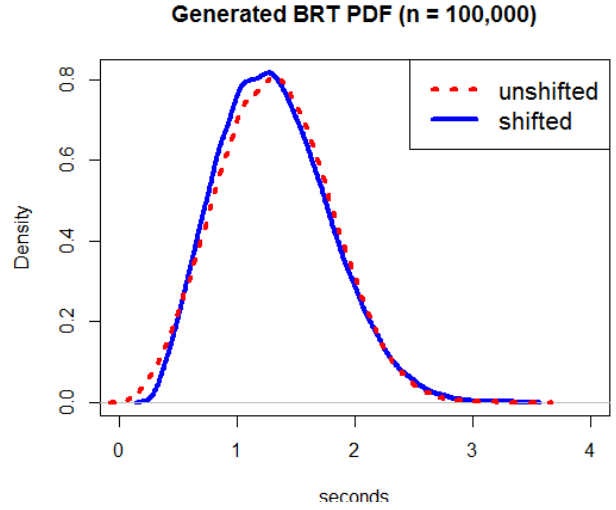


FIGURE 3 – SHIFTED WEIBULL FITTING

[2014] are shown in comparison to those generated using the shifted parameters derived above. The mean and standard deviations of the two distributions are within a few hundredths of a second, yet the shifted distribution has a minimum value of 0.266s.

To generate braking reaction time (BRT) values that fit the shifted Weibull distribution, we simply invert the distribution function, $F(t)$, in Equation [16], and insert instances of a uniform distribution, U_i , bounded on [0,1):

$$\tau_{BRT} = \gamma_w + \lambda_s [-\ln(1 - U_i)]^{1/k_s} \quad [21]$$

Since τ_{BRT} is only the perception reaction time, the movement time (τ_M – kept at a constant 0.15s [Lister 1950] for simplicity) and the device response time (τ_{DR}) must be added. For τ_{DR} , vehicles in the simulation are considered to have four-wheel disk brakes actuated by hydraulic cylinders. In a study of the performance of such braking systems [Grover et al. 2008], the authors found that vehicles in a hard-braking event achieve nearly the maximum deceleration approximately 0.2 seconds after the driver's

initial application of the brake pedal.²⁷ Thus, a combined movement and device response time of 0.35s is added to the τ_{BRT} derived from the three-parameter Weibull distribution to give the total reaction time τ_{RT} .

The delay-differential equations used in implementing ADRT cannot be applied directly to BRT, as the accelerations affected by the latter are calculated at each timestep to bring the vehicle to a stop at a specific location based on current speed and distance from that location. Using values calculated at some prior time will cause every driver to overshoot her stopping point. Therefore, an approach based on an adaptation to an *iterated coupled map* was enacted.²⁸ The calculated value for deceleration is stored at each time step. When a time, t , is reached for which the value stored at time $t - \tau_{RT}$ is lower than the current car-following acceleration, the driver begins to yield; however, the yielding deceleration is calculated based on current values of location and velocity.

IMPERFECT PERCEPTION

Perceived headway and velocity relative to a leading vehicle are subject to driver estimation errors.²⁹ To model these errors, the acceleration equation takes the same form; but estimated values (s^{est} and v_l^{est}) replace the true values of the gap to the

²⁷ The study found vehicles reach maximum deceleration at approximately 0.26s after initial pedal application, but have exceeded 90% of that value within 0.2s

²⁸ For an exploration of how this differs from a delay-differential equation, and the resulting effects on traffic stability, see [Kesting and Treiber 2008b]

²⁹ This section adapted from [Treiber and Kesting 2013a: pp 210-213]

vehicle ahead and of the leading vehicle's speed. It is assumed the speedometer provides a consistent (even if not entirely accurate) value for the driver's own speed. "In most driving situations, the *relative* estimation error for the gap, or, equivalently, the error of the logarithm of the gap, turns out to be essentially constant," allowing us to write

$$\ln(s^{est}) - \ln(s) = V_s w_s(t) \quad [22]$$

V_s , the statistical *variation coefficient*, represents the relative standard deviation of s^{est} from the true headway. Following Treiber and Kesting [2013a], its value is set to 0.1, and the error is assumed to have no bias. The evolution of driver error in time is modeled here as a Wiener process, represented by the stochastic variable w_s . The explicit derivation and evolution of this term will be discussed below.

Drivers estimate the speed of the vehicle ahead relative to their own based on the change in apparent optical angle subtended by the vehicle. [Lee 1976] By the small angle approximation, a vehicle at distance s , of width w_{veh} is seen as $\theta \approx w_{veh}/s$. "Based on experiments we assume the error of the *rate of relative angular change* to be constant" [Treiber and Kesting 2013a]

$$\omega = \frac{d\theta/dt}{\theta} = \frac{\frac{w_{veh}}{s^2} \Delta v}{w_{veh}/s} = \frac{\Delta v}{s} = \frac{1}{TTC} \quad [23]$$

Uncertainty in estimation of Δv is proportional to distance (therefore time to collision, $TTC = \frac{s}{|\Delta v|}$, has constant uncertainty).

“Assuming a constant standard deviation, σ_r , of the relative approach rate (on the order of $\sigma_r = 0.01s^{-1}$) we obtain”

$$v_l^{est} - v_l = -(\Delta v^{est} - \Delta v) = -s \left(\frac{1}{TTC^{est}} - \frac{1}{TTC} \right) = -s\sigma_r w_l(t) \quad [24]$$

w_l describes a stochastic variable analogous to the w_s used in calculating the evolution of estimated errors for headway. Every driver has her own independent set of $\{w_s(t), w_l(t)\}$, which are all independent from each other. The same approach is used to implement the same class of errors for drivers yielding to pedestrians and pedestrians yielding to drivers. Further, each pedestrian has two sets of these stochastic variables, $\{w_{s,l}(t), w_{l,l}(t)\}$ and $\{w_{s,r}(t), w_{l,r}(t)\}$, for vehicles approaching from the left and right, respectively.

To generate stochastic variables within the model, a *stochastic process* must be defined. In the following, it is assumed that $w_s(t)$ and $w_l(t)$ are instances of a stationary (having no general drift velocity) process, $w(t)$. This process can be defined by (i) its time-dependent distribution function, and (ii) its *autocorrelation* function – which describes the correlation between the process at two times as a function of the time elapsed between measurements. Driver perception errors are assumed, following [Treiber and Kesting 2013a], to follow a standard Gaussian distribution:

$$w(t) \sim N(0,1) \rightarrow \langle w(t) \rangle = 0, \quad \langle w^2(t) \rangle = 1 \quad [25]$$

Ignoring distraction for the moment, errors in a driver’s perception are assumed to be “memoryless” – meaning future probability densities can be determined based

only upon knowledge of the current state of the system. This is a loose definition of a *Markov process*. They do, however, exhibit a certain *persistence*: If, for example, a driver underestimates a distance at one moment, she is likely to continue to do so in the next moment. Mathematically, this means that, unlike white noise, the errors at two times are positively correlated for a specified *persistence time*, $\tilde{\tau}$, of “a few seconds up to one minute.” [Treiber and Kesting 2013a] (A persistence time of 20 seconds was chosen for the simulation.) This can be described by the following autocorrelation function:

$$\langle w(t)w(t') \rangle = \exp\left(-\frac{|t - t'|}{\tilde{\tau}}\right) \quad [26]$$

WIENER PROCESS

An ordinary differential equation with a noise term is the basic, heuristic definition of a Langevin equation: $\frac{dx}{dt} = a(x, t) + b(x, t)\xi(t)$.³⁰ A more mathematical definition of “noise” is a “rapidly varying, highly irregular function,” such that for $t \neq t'$, $\xi(t)$ and $\xi(t')$ are statistically independent. Also required is that $\langle \xi(t) \rangle = 0$ so any non-zero mean is contained within $a(x, t)$; therefore,

$$\langle \xi(t)\xi(t') \rangle = \delta(t - t').$$

³⁰ This derivation based on Ch 4 of [Gardiner 1983]

This is Dirac's $\delta(x)$, which is equal to 0 for all $x \neq 0$.³¹ $\xi(t)$ then represents standard, idealized *white noise*. With this definition, it is now possible to specify a stochastic differential equation that satisfies the above conditions:

$$\frac{dw}{dt} = -\frac{w}{\tilde{\tau}} + \sqrt{\frac{2}{\tilde{\tau}}} \xi(t) \quad [27]$$

This equation, describing what is known as a *Wiener process*, represents the solution to the Fokker-Planck equation in one variable, $w(t)$, with a drift coefficient of zero (essentially an Ornstein-Uhlenbeck process that has decayed to a stationary state), and a diffusion coefficient of 1.

The formal solution is then

$$w(t) = \sqrt{\frac{2}{\tilde{\tau}}} \int_{-\infty}^t e^{-(t-t')/\tilde{\tau}} \xi(t') dt' \quad [28]$$

To implement this in a model, an *update scheme* must be derived. The exact update scheme for an Ornstein-Uhlenbeck process of this form is:³²

$$w_i = e^{-\Delta t/\tilde{\tau}} w_{i-1} + \sqrt{1 - e^{-2\Delta t/\tilde{\tau}}} \eta_i \quad [29]$$

Treiber and Kesting use an approximation (which is implemented in this model):

³¹ Formally, the δ -function only makes sense within an integral, mapping a function to a number:

$\int_{-\infty}^{\infty} \delta(x) f(x) dx = f(0)$.

³² For a derivation, see [Gillespie 1996]

$$w_i = e^{-\Delta t/\tilde{\tau}} w_{i-1} + \sqrt{\frac{2\Delta t}{\tilde{\tau}}} \eta_i \quad [30]$$

Here, η_i are instances of a computer-generated independent, identically-distributed random number from a normal distribution with an expectation value of 0 and unit variance. The approximation is accurate when $\Delta t \ll \tilde{\tau}$.

DISTRACTION

Implicitly assumed in all the above models is that the actors involved not only receive the relevant information about the world around them, but that the information is processed (either consciously or subconsciously) and acted upon. However, as discussed above, the driver's attention resources are limited; so, this is not always the case. Distraction can come in many forms and defining it explicitly – even within the confines of the driving activity – is a debated topic in the literature. For our purposes, the following definition will suffice: “Driver distraction is a diversion of attention away from activities critical for safe driving towards a competing activity.” [Lee et al. 2008] Distraction while driving is dangerous: In Treat's study of accident causations [Treat et al. 1979], “improper lookout” and “inattention” were the two leading causes of accidents for which a definite cause could be determined. In 2014, of the 32,675 [FARS 2014] traffic fatalities in the United States, 3,179 (nearly 10%) were explicitly recorded by police as caused by some form of distraction. [Distraction.gov]

In the overwhelming majority of instances, however, becoming distracted has no serious consequences for a driver. This “latent distraction” does not manifest as

deterioration in observable measures of driving performance, but in readiness to react to new (possibly safety-critical) developments. [Schaap et al. 2013] Quantifying the occurrence of distraction is, therefore, a difficult concept. One promising approach is to relate attention to visual allocation (where the driver's eyes are looking). A relatively early attempt at doing so [Wierwille & Tijerina 1998] combined mentions of driver pre-crash distractions in crash narratives from police reports with independent calculations of the frequency and duration of visual allocation to in-car distractions. Despite the vague data, the authors found significant correlation ($R = 0.982$)³³ between accidents citing a specific distracting activity and the exposure level of the activity (exposure = average-single-glance-duration × average-number-of-glances × frequency-of-use). The authors of an analysis of the 100-Car Naturalistic Driving Study [Dingus et al. 2006] (which is discussed below) found that, in their dataset, “almost 80 percent of all crashes and 65 percent of all near-crashes involved the driver looking away from the forward roadway just prior to the onset of the conflict.”

In another analysis of the 100-Car NDS [Liang et al. 2012], the authors compared a variety of algorithms based on different combinations of recent glance histories, adjacent short glances, and the locations of off-road glances. They concluded that none of these complications added predictive power to simple measures of instantaneous

³³ This value does not include crashes that referenced looking at the speedometer, which was a significant outlier.

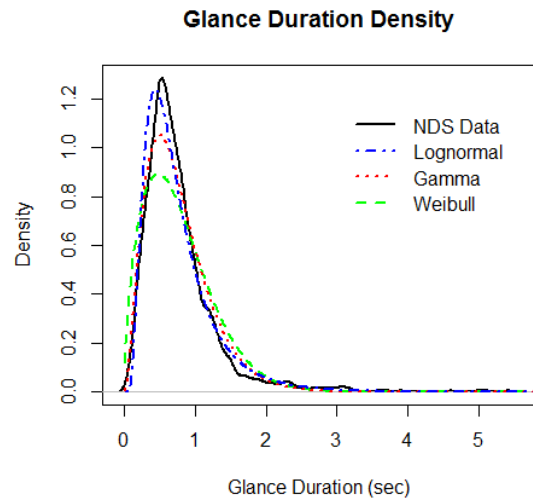
glance duration. Despite the limitations, the time drivers spend with their eyes not on the road is a reasonable proxy for distraction, and at least gives a lower bound. Many studies have been done on visual attention under controlled conditions; however, recent developments in sensor technology, as well as image and data analysis, have made large-scale studies of driving possible outside the artificial limits of explicit experiments. These naturalistic driving studies (NDSs) provide a wealth of insights into typical driver behavior.

Two large-scale NDSs have been performed in recent years. The 100-car NDS recorded extensive data on the behaviors of 241 drivers over nearly 43,000 hours of driving. [Dingus et al. 2006] This served as a pilot project for the NDS portion of the second Strategic Highway Research Program (SHRP2), which included more than 3,000 drivers over more than 1 million hours of driving. [Victor et al. 2015] Included in both data collections were video records of drivers' faces, from which the direction in which their eyes are pointing (glance data) can be extracted.

As part of a study on how visual attention influences crash risk, [Bärgman et al. 2015] glance behavior was extracted from 223 20-second video segments taken from the SHRP2 dataset. In this dataset, the authors found that drivers were not looking at

the forward roadway³⁴ 21.17% of the time.³⁵

In another study, 4,083 6-second observation windows from the 100-car NDS were analyzed to find how the duration of off-road glances is distributed. [Rootzén & Zholud 2016] The authors focused on glances longer than 2 seconds, but they graciously published their extracted glance data with the analysis.



**FIGURE 4 –
PRELIMINARY GLANCE DISTRIBUTION FITTING**

From these data, after removing periods with no glances away from the road or glances truncated by the observation window, 1740 observations remain. The fit-performance of a few distributions to these measurements was checked visually using built-in R methods. Figure 4 shows that, based on this preliminary analysis, the lognormal distribution appears to be the best fit. This is not surprising since, as observed in the section on reaction times, lognormal distributions are common in random processes that

³⁴ The authors defined “looking at the forward roadway” based on the definition given in [Victor 2005]: fixations falling within a circular area 16 degrees in diameter, centered on the *road center point*. The *road center point* was defined as the most frequent gaze angle from the driver’s baseline driving data.

³⁵ It should be noted that each of these time periods was selected as a baseline measurement to be compared to the behavior of the same driver (on the same day) in a time period that ended with a crash or near-crash event. Since each of the drivers selected for the baseline data was involved in a crash or near-crash later the same day, there is a selection bias toward drivers potentially more prone to distraction. In fact, matched baselines (chosen using a similar set of criteria) exhibited slightly higher percentages of time with eyes off the road than random baselines in another study of the SHRP2 data [Victor et al 2015, figure 6.7]; however, a literature review did not yield any summary data of random baseline glance behavior from the SHRP2 data.

feature low average values, large variances, and are exclusively positive. [Limpert et al. 2001] The PDF of the lognormal distribution is:

$$f(X_i|\mu, \sigma^2) = \frac{1}{X_i\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\ln(X_i) - \mu)^2}{2\sigma^2}\right]$$

Here, μ is the *scale parameter*, and σ is the *shape parameter*. To fit this distribution to the data, as with the Weibull distribution used for reaction time, the *maximum likelihood* method can be used.³⁶

The likelihood function of the lognormal distribution, given a set of $X = \{X_1, X_2, \dots, X_n\}$, is equal to the product of the probability densities of each X -value:

$$\begin{aligned} L(\mu, \sigma^2|X) &= \prod_{i=1}^n [f(X_i|\mu, \sigma^2)] \\ &= (2\pi\sigma^2)^{-n/2} \left[\prod_{i=1}^n X_i^{-1} \right] \exp\left[\sum_{i=1}^n \frac{-(\ln(X_i) - \mu)^2}{2\sigma^2} \right] \end{aligned} \quad [31]$$

As in the analysis of delayed reaction, we now take the logarithm of this likelihood equation. The gradient of the resulting log-likelihood equation is found, and each equation is separately set to zero: $\frac{\partial L}{\partial \mu} = 0$, and $\frac{\partial L}{\partial \sigma^2} = 0$.³⁷ Solving the resulting equations for μ and σ^2 gives the maximum likelihood estimators, $\hat{\mu}$ and $\hat{\sigma}^2$:

³⁶ This derivation follows that in [Ginos 2009]

³⁷ For proof that the resulting extremum is, indeed, the maximum – i.e. that the Hessian (the second derivative matrix) is negative-definite – see [Ginos 2009].

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n \ln(X_i) \approx -0.455 \quad [32]$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (\ln(X_i) - \hat{\mu})^2 \approx 0.373 \quad [33]$$

From these formulas, the relationship between the lognormal distribution and the normal distribution is obvious: If X fits a lognormal distribution, the natural logarithm of X , $\ln(X)$, is normally distributed; and the *scale parameter*, μ , can be interpreted as the mean of $\ln(X)$, while the *shape parameter*, σ , is related to the standard deviation of $\ln(X)$. Conversely, if Y fits a normal distribution, then $X = e^Y$ will be lognormally distributed. Therefore, if we generate a normally-distributed random variable, η_i , with mean $\hat{\mu}$ and standard deviation $\hat{\sigma}$, then raise Euler's number (e) to the power of the numbers generated, the resulting values, D_i , will have a distribution that approximates that of the observed durations of distraction:

$$D_i = e^{\eta_i \hat{\sigma} + \hat{\mu}} \quad [34]$$

As a check, the cumulative distribution function (CDF) of 1000 numbers created by this method is compared to the CDF of the SHRP2 data in Figure 5.³⁸ In practice, the minimum duration of a distraction event in the model will be one tick of the model clock.

³⁸ The stepping behavior in the CDF of the SHRP2 data [Rootzén & Zholud 2016] is an artifact of the 0.1-second precision in with which it was made available.

Each period of distraction, D_i , and the subsequent inter-distraction period, ID_i , are modeled here as independent samples from identical distributions (iid).³⁹ Assume a system can be in state A or B . If the time between a transition into state A and the subsequent transition out of it, Y_A , is independent of

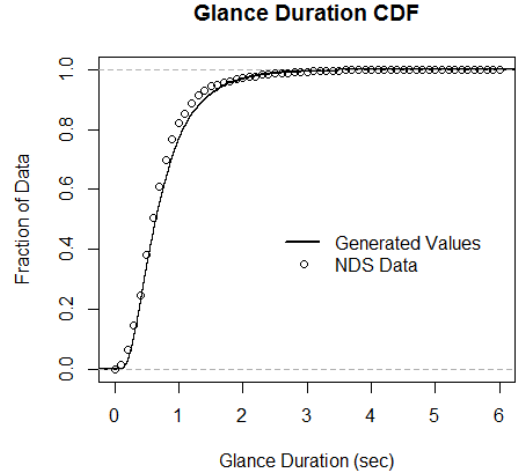


FIGURE 5 – GLANCE CDF FITTING

previous transition times – and likewise for each Y_B – and the distribution of these dwell times is known, then the process can be modeled as an alternating renewal process.⁴⁰ If the expectation value of the distribution of Y_A is M_A and the expectation of Y_B is M_B , then the long-run proportion of time spent in state A is given by

$$P_A = \frac{M_A}{M_A + M_B} \quad [35]$$

Rearranging Equation [35], and applying it to the distraction intervals from the NDS discussed above, we have:

$$\frac{1}{\lambda_{ID}} = \frac{M_X(1 - P_D)}{P_D} \quad [36]$$

³⁹ This assumption is not quite realistic. In fact, we have every reason to assume that a person looking away from the road at a secondary task will look back at it soon thereafter if the task remains incomplete. While this poses a limitation for modeling a given individual's long-term distraction behavior, safety-critical situations are immediate results of individual errors. Since the topic of interest is the crash risk of the population, the overall statistical behavior should not be affected; and, as discussed in [Liang et al 2012], instantaneous, memoryless glance behavior is the best predictor for crash risk.

⁴⁰ For more information on renewal processes, and a derivation of Equation [35], see e.g. [Cox 1967]

Here, P_D represents the proportion of time drivers spend with their eyes off the road; M_X is the arithmetic mean of the X_i observations;⁴¹ and $\lambda_{ID} = \frac{1}{M_{ID}}$ is defined (in anticipation of the distribution that will be used to model it) as the arrival rate of distraction events, given an average inter-distraction interval of M_{ID} . Combining our analysis of the glance data from Rootzén & Zholud [2016] with the total proportion of time drivers look away from the road from Bärghman et al. [2015], the resulting arrival rate is $\lambda_{ID} \approx 0.352 \text{ sec}^{-1}$ – or, a little less than three seconds between each glance away from the road.

The arrival of distraction events will be modeled as a Poisson process. Therefore, the probability of an inter-arrival time equal to Z , $\pi_{ID}(Z)$, will follow the exponential distribution:

$$\pi_{ID}(Z) = \lambda_{ID} e^{-\lambda_{ID} Z}$$

To generate intervals with this distribution within the model, we invert the equation:

$$Z = -\frac{\ln(\pi_{ID}/\lambda_{ID})}{\lambda_{ID}} \quad [37]$$

Generating π_{ID} as instances of a uniformly-distributed random variable, U_i , will then give us values of Z that fit the exponential distribution – with one caveat: the result must be positive, so the quantity within the natural logarithm must be less than one.⁴²

⁴¹ This is not equal to $e^{\hat{\mu}}$, as one might expect. $e^{\hat{\mu}}$ is, in fact, the geometric mean.

⁴² $\ln(1) = 0$. While this is physically possible (There is nothing ensuring the driver checks the road between attending to distractions.), this behavior would have been recorded in the NDS as a single distraction event. To avoid this, we insist that the quantity within the logarithm is strictly less than 1.

We therefore insist $\pi_{ID} < \lambda_{ID}$. Since the random input, $\pi_{ID} < 1$, this is accomplished by linearly scaling it as $\pi_{ID} * \lambda_{ID}$.⁴³

$$ID_i = -\frac{\ln(U_i)}{\lambda_{ID}} \quad [38]$$

To scale the measured distraction intervals (that fit the distribution defined by Equations [32] and [33]) into units of the model, we have to pass from measured quantities, X_i , into model units, X'_i . To do so, the quantities must pass through the logarithmic function; therefore, so must the scaling parameters. If we have a value of c seconds, and our model uses a granularity of d seconds per tick, then $\ln\left(\frac{c}{d}\right) = \ln(c) - \ln(d)$. Therefore, $\hat{M}' = \hat{M} - \ln(d)$. Substituting this value into the equation for variance (Equation [33]), and applying similar logic, gives an interesting result: Looking at the quantity within the parentheses, $\ln(X_i) - \ln(d) - \hat{M}' = \ln(X_i) - \ln(d) - \hat{M} + \ln(d) = \ln(X_i) - \hat{M}$. This is the same value as the original. Thus, while \hat{M} is scaled by the transition into model units, \hat{S} remains unchanged. Similar logic applies to the generation of stop distances from the crosswalk and to IDM parameters observed to follow a lognormal distribution.

When vehicles within the simulation are “distracted,” they do not update external information (e.g. headway, speed differential with other cars, or the presence

⁴³ A second caveat arises from the implementation of the model in Java. Java’s `Random.nextDouble()` method returns uniformly-distributed random values on the interval $[0,1)$. Since $\ln(0) \rightarrow -\infty$, a small value must be added to the result to avoid errors. The value added is 10^{-15} . This creates a maximum period without distraction of more than 10 minutes – an arbitrary, but not-unrealistic restraint.

of crossing pedestrians). They do however remain aware of their own speed relative to their desired maximum velocity. This is accomplished by fixing the values of z in Equation [10] and a_{CAH} in Equation [11] at the pre-distraction levels and ignoring pedestrians.

DRIVER YIELDING

When not distracted, if drivers see a pedestrian intending to cross they can, and often do, choose to yield. In fact, in many States drivers are legally required to stop for a pedestrian who has indicated they intend to cross. While attempts have been made to model this behavior [e.g. Schroeder and Roupail 2011], many factors come into play that are beyond the scope of this exploration. Signage, crosswalk design and treatments, and social norms are just a few of the factors that have significant influence on driver yielding behavior. For a review of some of the research on how treatments from around the world affect pedestrian safety, see [Campbell et al. 2004].

When a driver yields to a crossing pedestrian, her vehicle can block drivers behind her in adjacent lanes from being able to see any pedestrian in the crosswalk. Therefore, the distance from the edge of the crosswalk at which the vehicle comes to rest (d_{stop} – the x in Equation [39]) has a direct effect on pedestrian safety. (This scenario, referred to as a *multiple threat* [Snyder 1972], will be discussed in the *Model Implementation* section on *Visual Obstructions*.)

In Oregon, it is recommended that drivers come to a stop 30 feet (9m) from the crosswalk, but this is not legally required. The MUTCD recommends 20 – 50 ft (6.1 – 15 m). Many complex factors have been shown to correlate with d_{stop} including behavior of nearby vehicles, immediate history of the driving trip, pedestrian behavior, and vehicle type, among others. [Figliozi and Tipagornwong 2016] Advance warning signs and pavement markings have been shown to increase yielding rates and result in a qualitative increase in yielding distances [Pulugurtha et al. 2012, Berger 1975], yet a literature review produced few studies that quantify the distribution of baseline or post-treatment behavior.

Due to the wide array of infrastructure treatments available, and the complexity of the behaviors involved, the model developed herein will be concerned with unsignalized mid-block crossings, marked only with lines on the road surface. The effect of these treatments is to give drivers advanced warning that pedestrians may be entering the roadway ahead. This enhanced situational awareness is intended to increase the probability that a driver's attention will be on the task at hand.

d_{stop} has been measured for crossings that do not include a stop-bar. [Houten, Malenfant, and McCusker 2001] However, implementing these distances led to unrealistically high numbers of conflicts from visual obstruction in the simulation. Therefore, stop-bars were added 9 meters from the side of the crosswalk nearest approaching drivers. This is the only crosswalk treatment included in the simulation.

Even in the absence of safety treatments, driver yielding is influenced by a variety of factors. For example, drivers are more likely to yield if they perceive a pedestrian as disabled [Pancer et al. 1979] or in need of assistance [Malamuth, Shayne, and Pogue 1978]. The probability of drivers yielding is also significantly increased when pedestrians make their intention to cross known: Pedestrians standing just off the curb are yielded to significantly more often than those standing at the curb edge. [Harrell 1993] When drivers are traveling at high speeds, on the other hand, they are less likely to yield to pedestrians. [Bertulis and Dulaski 2014]

Motorist yielding behavior has seen significant research, and a variety of models are available based on driver characteristics. [e.g. Sun et al. 2002] In the model implemented for this study, however, a simplified approach is taken. Drivers do not yield unless a pedestrian has decided the gap available is sufficient to cross safely. The details of pedestrians' decision-making processes are presented in the *Pedestrian Modeling* section. In short, each pedestrian has a statistically-assigned minimum time-gap between approaching cars within each lane, and only crosses when traffic allows. If a pedestrian has entered the roadway, drivers (after their calculated reaction time, assuming they are not distracted) yield accordingly. When the distance to the nearest vehicle is small and a pedestrian's criteria to begin crossing are met, however, pedestrians signal their intention to cross and wait for the approaching cars to yield. This recognition of mutual awareness is assumed to be absolute, and the form it takes is

not addressed. Specifics of this are given in the *Model Implementation* section. For more detailed analysis of this interaction, see [e.g. Matúš 2014 or Zhuang and Wu 2014].

Whether a driver is yielding to a waiting (or already crossing) pedestrian, the attendant deceleration is predicated on the driver seeing the pedestrian. When and where this perception occurs is crucial to safe crossings. This is the basis of the American Association of State Highway and Transportation Officials (AASHTO) minimum *stopping sight distance* requirements: Drivers need room to stop. Ignoring infrastructure limitations such as hills or turns, the question arises of how far away a driver can see a pedestrian. Assuming a visual acuity capable of discerning objects subtending 1 minute of arc-angle (normal, “20/20 vision”), a 5’ tall pedestrian can be seen from more than 5km away. [Shinar 2007b] Visual capacity is, therefore, not the limiting factor.

PERCEPTION OF PEDESTRIANS

The distance at which drivers perceive pedestrians, d_{perc} , is a topic that has seen significant amounts of research [Kwan and Mapstone 2004]; however, the vast majority of such studies focus on night-time sight distances. [Langham and Moberly 2003] Those performed in daylight conditions seem to exclusively focus on the relative benefits of different colors and forms of fluorescent safety garments. Most of these studies limit their measurements to rural roads; however, one [Sayer and Buonarosa 2008] explored the difference in detection distance between “medium” complexity (urban) and “low” complexity (rural) scenes. The authors found average detection distances of 195m and 266m, respectively. Compounding the fact that pedestrians in the study wore high

visibility garments, participants in the study were explicitly instructed to look for pedestrians, making driver expectancy unrealistically high. This is a significant caveat. In a study aimed at differentiating *attention conspicuity* (“the capacity of an object to attract attention”) from *search conspicuity* (“the property of an object that enables it to be quickly and reliably located by search”), researchers found drivers three times more likely to notice something if told to look for it. [Cole and Hughes 1984]

In an attempt to remediate this increased expectancy, Helman and Palmer [2010] measured pedestrian perception distances for English drivers who were asked to simply narrate anything related to the driving task that caught their eye in the outside environment. The authors recorded detection distances of drivers navigating a closed track for mannequins wearing full-body, two-tone fluorescent safety attire with retro-reflective trim in two scenarios. For the first scenario, in which mannequins were placed next to a brightly-colored Highways Agency vehicle with a flashing amber light bar on top, perception distances ranged from 110m and 150m. For the second, which featured mannequins placed adjacent to a large road-work vehicle with flashing lights, detection distances ranged from 106m to 141m. Since these values apply to pedestrians wearing full-body fluorescent clothing, this can be considered an upper bound for normal pedestrian detection. A value for d_{perc} of 100m is used in the simulation for the distance at which an attentive driver can become aware of a pedestrian intending to cross.

YIELDING KINEMATICS

As a driver in the real world approaches a crosswalk, she subconsciously calculates the time until the crosswalk would be reached traveling at her current speed. Based on this calculation, in relation to the observed behaviors of the pedestrians present, it may be necessary for her to decelerate to avoid a collision. If she does decide to yield, the required braking pressure is also calculated subconsciously, based on her experience as a driver. Implementing this in simulation, however, requires calculation of the necessary accelerations. If we assume a constant acceleration, a , a body currently located at the point x_0 and traveling with speed v_0 will travel in time t to a point, x , given by the kinematic equation:

$$x = \frac{1}{2}at^2 + v_0t + x_0 \quad [39]$$

If the body is to come to rest at the end of time t (a *hard yield*), then

$$a_{hy} = -v_0/t \quad [40]$$

Solving for t and substituting into Equation [39], we can solve for the acceleration needed to come to rest at a given location, x :

$$a_{hy} = \frac{-v_0^2}{2(x - x_0)} \quad [41]$$

By plugging this back into Equation [40], we have an equation for the time to execute such a yield:

$$t_{hy} = \frac{2(x - x_0)}{v_0} \quad [42]$$

If this acceleration is less than that given by the car following equation, the driver uses the yielding acceleration instead and allows the pedestrian to cross.

If the time until a given pedestrian will be out of harm's way, T_{safe} , is greater than t_{hy} , then a hard yield is necessary. Otherwise, the driver need not come to a complete stop to allow the pedestrian to cross out of danger,⁴⁴ and a *soft yield* is sufficient. If we solve Equation [39] for t , and set this equal to the time it will take the pedestrian to be clear, T_{safe} , we can find the acceleration needed:

$$\frac{-v_0^2 + \sqrt{v_0^2 - 2a(x - x_0)}}{a} = T_{safe} \Rightarrow a_{sy} = \frac{-2[v_0 T_{safe} - (x - x_0)]}{T_{safe}^2} \quad [43]$$

Here, and at every step in the simulation, decelerations are limited to values that are physically achievable by average drivers of passenger vehicles on dry roads: -7.4 m/s². [Greibe 2007] As mentioned in the section on *Human Factors*, while the decision to yield to a pedestrian is delayed based on the driver reaction time, the acceleration values are calculated based on current location and velocity.

SUMMARY

Driver behavior represents a broad collection of intricate, interconnected subsystems. This diversity means that a broad variety of approaches can be taken in modeling that behavior. The driver model implemented in the simulation presented in

⁴⁴ In Oregon, this means a driver may not enter the crosswalk until a crossing pedestrian has exited the driver's lane and is at least half way across the adjacent lane.

this work is a microscopic model, limited to car-following behavior and yielding to pedestrians. This is a significant simplification. However, the interplay of driver's mental workload during lane-changes and the act of yielding to pedestrians is a topic that would benefit from additional research. Also lacking in the literature is quantification of the prevalence of lane-changes as a form of yielding to crossing pedestrians.

The car-following model implemented in this work is based on a collection of extensions to the Intelligent Driver Model (IDM). [Treiber et al. 2000] Namely, it combines the stabilizing edge-cases of the Enhanced IDM [Kesting, Treiber, and Helbing 2010] with the limitations of human drivers described by the Human Driver Model (HDM). [Treiber, Kesting, and Helbing 2006] The HDM incorporates finite reaction times and imperfect perceptions of distance and velocity into the IDM. The parameter values used for these models are taken from empirical data presented in previous research.

The main contribution of this work is development and implementation of a novel approach to modeling driver distraction. Driver distraction is a significant causal factor in many traffic crashes, and its general exclusion from modeling efforts is a glaring limitation of current simulations that aim to model traffic safety. The distributions calculated combine measurements of eye-glance behavior from two studies: The overall time drivers spend looking at the road ahead comes from an analysis of data collected during the SHRP2 project. [Bärgman et al. 2015] Data from the 100-Car Naturalistic Driving Study [Rootzén & Zholud 2016] were then used to derive the distribution of durations for those times drivers look away from the road to attend to another task.

Admittedly, although the sample of glance durations is numerically large (1,740 periods of observation), each sample is only 6 seconds long. This represents slightly less than 3 hours of data in total, randomly sampled from the 241 participants (including primary and secondary drivers) in the study. However, the use of eye-glance data has become widespread in driving research in recent years. Similar analysis to that presented here, performed on a larger sample of data, is recommended to enhance the generality of the resulting model of distraction.

To test how significant the decision of whether to include distraction is when using a simulation to model safety, the developed model was applied to drivers approaching a simulated pedestrian crosswalk. To minimize externalities, no crossing treatments were included. Driver yielding behavior is also significantly simplified. Drivers do not yield until a pedestrian has indicated his intention to cross the street. There is a significant gap in the published literature regarding how far away drivers are able to recognize the presence of pedestrians at the roadside during daylight hours.

As with driver modeling, diverse approaches have been taken in modeling pedestrian behavior. These various perspectives have much in common with their counterparts in driver modeling. However, since pedestrians are not typically confined to lanes, modelling their behavior can be more complex than doing the same for drivers. As it turns out, this freedom of movement only holds in low pedestrian densities. In dense crowds, pedestrian behavior is “surprisingly predictable,” [Helbing, et al. 2001] even fitting its own fundamental diagram.

One approach to modeling pedestrian movements uses *regression models* (e.g. [Weidmann 1993]) to empirically derive relations among flow variables, enabling prediction of behavior in situations comparable to those from which the relations were derived. Macroscopic models that utilize theories from the study of fluid dynamics and granular flow (e.g. [Piccoli & Tosin 2011]) are more widely applicable, since they are explanatory instead of descriptive. Macroscopic models can also help predict demand, informing decisions regarding the capacity needs of new infrastructure. [Ronald et al. 2007]

Queuing models (e.g. [Løvås 1994]) define travel times along links and add random wait times at bottlenecks to recreate delays. Cellular automata approaches (e.g. [Burstedde et al. 2001]) parallel those in driver modeling by discretizing space into *cells* and defining rules for the occupation of those cells by virtual pedestrians. Discrete

choice models (proposed in [Gipps 1987]) define a finite number of pathways through an area, among which pedestrians decide based on subjective rational choices.

As with driver modeling, treating each actor as an individual can provide greater insight into the dynamics of a crowd. Legion is a commercially-available software system (based on [Still 2000]) that simulates pedestrians with senses and decision-making capacities using a proprietary *least effort* algorithm. STREETS [Schelhorn et al. 1999] is another agent-based model that utilizes demand models and the SWARM simulation platform [Swarm 1999] to provide detailed modeling of pedestrian behavior, including socio-economic and psychophysical phenomena.

The diversity of personal behavior patterns in pedestrians, like that of drivers, is a complex topic. These patterns have strong influences on crash statistics. For example, in 2016 more than 2.2 times as many male pedestrians were killed by vehicles than female. [FARS 2016] By fitting behavior models to data collected from diverse samples of the population, these variations should be inherently captured by the model without explicit designation of the population characteristics that correlate with the observed behaviors. While this is generally a reasonable assumption, models based on this approach will not include populations that exhibit extreme values such as the elderly, handicapped, or intoxicated – since these were not present in the samples used to derive the distributions applied.

In general, the act of walking to a destination can be broken down into the three functional levels of control that Michon defined for driver behavior [Michon 1979]. This hierarchical model is discussed further in the section on *Driver Modeling*, but it describes behavior as a generally top-down system of decisions and actions in three realms: *strategic* (e.g. choice of departure time and activity pattern), *tactical* (e.g. activity scheduling and route choice), and *operational* (walking behavior). [Hoogendoorn and Bovy 2004] Once pedestrians have a destination in mind, their behavior can be modeled in a straightforward manner. Within this simulation, it is assumed that pedestrians have already made the decision to walk, as well as where to walk to, and that these decisions have led them to cross the simulated road – leaving only the operational level to be explicitly recreated.

One approach for modeling the operational level of pedestrian movement is through minimizing a cost function through calculus of variations. [Hoogendoorn 2001] In such a model, deviations from the planned route, large accelerations or decelerations, and proximity to other pedestrians or obstacles incur costs that are to be avoided. Another approach to modeling this facet of pedestrian behavior is a category that can be termed *generalized force models*.

GENERALIZED FORCE MODEL

This style of behavior modeling adds together the various socio-psychological and physical forces that determine the behavior of individual actors. The model chosen [Helbing et al. 2000] includes a goal-oriented *motive* force, and *interactive* forces among

pedestrians and with barriers. At each time step in the simulation, the forces acting on a pedestrian are totaled to define the instantaneous acceleration.

Each pedestrian, i , has mass m_i and a natural walking speed, v_i^0 . A constant mass of 80kg is used in the simulation. Average walking speed of pedestrians can vary, but a value of 1.4 m/s (~5 km/hr) was chosen based on the data available. [Zębala 2012; Akçelik 2001] with a standard deviation of 0.26 m/s. [Still 2000]

The direction toward a pedestrian's destination is represented by the normalized vector \mathbf{e}_i^0 ; and his speed, v_i , changes over a characteristic time, $\tau_{ped,i}$. The characteristic time is set to 0.5s. Interactions with other pedestrians, j , or walls (here, taking the form of the curb when yielding and the edges of the pedestrian box or the edges of vehicles stopped within the crosswalk), W , are *interaction forces* \mathbf{f}_{ij} and \mathbf{f}_{iW} , respectively.

Pedestrian acceleration can then be calculated with the following expression:

$$m_i \frac{d\mathbf{v}_i}{dt} = m_i \frac{v_i^0(t)\mathbf{e}_i^0(t) - \mathbf{v}_i(t)}{\tau_{ped,i}} + \sum_{j \neq i} \mathbf{f}_{ij} + \sum_W \mathbf{f}_{iW} \quad [44]$$

The pedestrian interaction terms, are given by $\mathbf{f}_{ij} = A_i \mathbf{n}_{ij} e^{(r_{ij}-d_{ij})/B_i}$ where A_i and B_i are constants chosen to fit observations; $d_{ij} = \|\mathbf{r}_i - \mathbf{r}_j\|$ represents the distance between pedestrians' centers of mass; $\mathbf{n}_{ij} = (\mathbf{r}_i - \mathbf{r}_j)/d_{ij}$ is the normalized pointing vector from pedestrian j to i ; and $r_{ij} = r_i + r_j$ is the sum of the pedestrians' radii. All constants used in the model are given in Table 4. Observational data [Rouphail et al. 1998] suggest pedestrians tend to maintain a buffer zone of approximately 0.75m² while

walking. Since detailed modeling of pedestrian flow is not the focus of this exploration, it is assumed that this area is circular, yielding a radius, r , of approximately 0.87m; though this is known to be much smaller in Asian cultures than North American or European. [Tanaboriboon et al. 1986] It is important to note that pedestrians only feel forces from other pedestrians in front of them – since real walkers cannot see behind them.

m (kg)	v^0 (m/s)		τ_{ped} (s)	A (kg · m/s ²)	B (m)	r (m)
constant	mean	st. dev.	constant	constant	constant	constant
80	1.4	0.26	0.5	2000	0.08	0.87

TABLE 4 – PEDESTRIAN PARAMETERS

Barrier interactions are similarly represented by $f_{iW} = A_i e^{(r_i - d_{iW})/B_i} \mathbf{n}_{iW}$. A pedestrian is over the curb if the distance from his center of mass to the curb, d_{iW} , is smaller than his radius, r_i . The same formulation is used to keep pedestrians within the crosswalk on the road. While this does not allow the model to capture aberrant pedestrian behavior such as jay-walking, that behavior is outside the purview of this exploration.

Barring interactions with barriers or other pedestrians, Equation [44] is an ordinary differential equation for pedestrian velocity. For simplicity, we will consider the scalar form of the equation, which (assuming $v(0) = 0$, and $y(0) = 0$) differentiates to:

$$v(t) = v^0(1 - e^{-t/\tau_{ped}}) \quad \text{and} \quad y(t) = v^0(t - \tau_{ped}(1 - e^{-t/\tau_{ped}})) \quad [45]$$

From this, the exponential quality of pedestrian acceleration is obvious. When a pedestrian is deciding when to cross a road, his crossing time factors into the decision. (Details of this decision are deferred to the next section.) Calculating this, therefore, requires calculating Equation [45]. For simplicity, however, the location $y(t)$ can be approximated for sufficiently large values of t by:

$$y(t) \approx v^0(t - \tau_{ped}) \quad [46]$$

Using values of $v^0 = 1.4 \frac{m}{s}$ and $\tau_{ped} = 0.5s$, Figure 6 shows that by the time the pedestrian has moved 2 meters, the approximation is indistinguishable from the exact solution. Therefore, the approximation of Equation [46] is valid for any lane beyond that closest to the curb. This approximation is used by pedestrians when calculating if a gap is sufficient to allow safe crossing, and by drivers when deciding on an appropriate yielding deceleration.

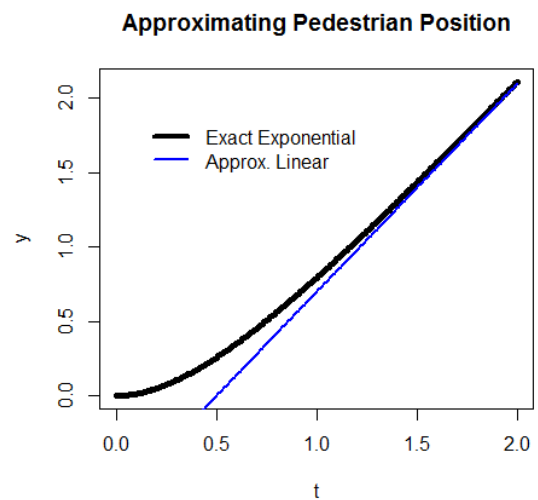


FIGURE 6 – PEDESTRIAN POSITION

For a detailed literature review and analysis of walking behavior, including variations with age, gender, physical ability, culture, temperature,⁴⁵ travel purpose, and many others, see [Daamen 2004].

⁴⁵ Weidmann [1993] found that average walking speed can vary by nearly 20% between 0°C and 25°C (quoted in [Daamen 2004])

PEDESTRIAN GAP ACCEPTANCE

As discussed above, pedestrian movement is amenable to decomposition into the three levels of control Michon used to describe driving behavior. [Michon 1979] Having arrived at the side of a road (based on strategic-level route-choice), the tactical level comes into play; and a pedestrian must decide when to step out and begin crossing. For this discussion, we will define a *gap* as the time-headway between two successive cars. Since the quantity of interest for a pedestrian just arriving at the side of a road is the time until the nearest car reaches the crosswalk (and not the time-headway to the next car, which has already passed), we will define this period as a *lag*.

If the nearest approaching vehicle is multiple lanes away, it will take a finite amount of time before the pedestrian reaches the driver's lane. Thus, the pedestrian may anticipate this fact and move his attention to the gap behind this nearest vehicle. If this latter gap is deemed sufficient for the pedestrian to cross at a comfortable speed, he may not wait for the vehicle to clear the crosswalk before beginning to cross. [Brewer et al. 2006] This phenomenon is referred to as a pedestrian accepting a *rolling gap*. In this model it is assumed that drivers will only yield if the pedestrian's current velocity will carry him into the driver's lane, or an adjacent lane, by the time the vehicle reaches the crosswalk.

There are a variety of approaches to modeling pedestrian gap acceptance. For example, Hachohen, Shvalb, and Shoval [2018] developed a pedestrian crossing decision algorithm in which pedestrians generate a mental risk map of their crossing path with

reference to approaching cars. Another group [Yannis, Papadimitriou, and Theofilatos 2013] calculated a binary logit model that incorporates characteristics of both the approaching vehicles and the pedestrian based on observational data. In the interest of simplicity, the model chosen is one that provides an explicit distribution for the behavior of the study population as a whole.

The study referred to is that by Brewer et al. [2006], in which the authors derived a distribution for acceptable gaps and lags from video observation of more than 600 crossings at 42 sites. The authors' data suggest that the probability of a pedestrian accepting a lag of y seconds is given by:

$$\pi_G = \left(1 - \frac{e^{\beta'(y)}}{1 + e^{\beta'(y)}}\right) * 100\% \quad [47]$$

Logistic regression was used to specify $\beta'(y) = 6.2064 - 0.9420y$ for the combination of all sites.

To generate pedestrians that fit this distribution, we need only invert Equation [47]. Setting $c_1 = 6.2064$ and $c_2 = 0.9420$, we have

$$y = \frac{1}{c_2} \left[c_1 - \ln \left(\frac{1}{\pi_G} - 1 \right) \right] \quad [48]$$

If we then generate π_G as instances of a uniformly distributed random number in the range (0,1), the values given to the pedestrians will fit the stated distribution. There is one caveat: to avoid negative values for a pedestrian's critical gap, the numbers created must be greater than a minimum, $\pi_{G,min} = (1 + e^{6.2064})^{-1}$. The value is

therefore linearly scaled as $\pi_G(1 - \pi_{G,min}) + \pi_{G,min}$. To check this formula, Figure 7 shows the cumulative distribution function of acceptable gaps for 150 calculations of the above formula, compared to the explicit equation for π_G .

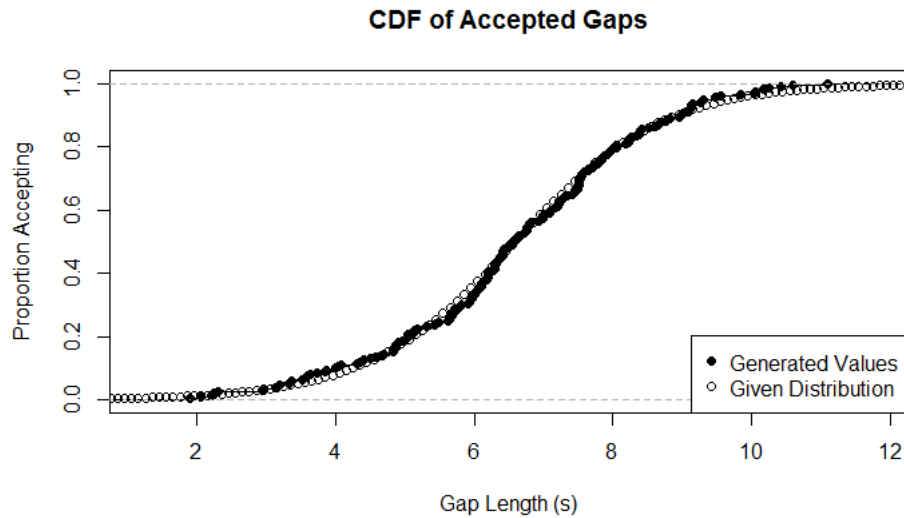


FIGURE 7 – GENERATED PEDESTRIAN GAP ACCEPTANCE

It has been demonstrated that these values are not constant in time for individual pedestrians: As wait times increase, the minimum gaps pedestrians are willing to accept tend to shrink, leading to increased conflicts. [Cheng, Wang, and Li 2013] The data acquired by Brewer et al. [2006], however, relate to the size of gaps that were accepted by pedestrians, not to individual pedestrians' behavior in time. Applying values derived from the resulting distribution does not reproduce the time-dependency of pedestrian gap preferences for individual agents in the model, but it innately recreates realistic crossing behavior. While it has been demonstrated [e.g. Crompton 1979] that

pedestrians are willing to increase their crossing speed after excessive wait times, this behavior was not included in the model.

It has been shown that drivers are more likely to yield to a group of pedestrians than to a single pedestrian. [Schroeder and Roupail 2011] While this behavior is not explicitly included in this model, as more pedestrians accumulate at the road's edge, the odds increase that one of the waiting pedestrians will have a relatively small value for his *acceptable gap*. If the gap is short enough, a pedestrian beginning to cross will cause drivers to yield, enabling other pedestrians to cross simultaneously.

SUMMARY

Due to the greater freedom of movement, pedestrian modeling is arguably more complex than driver modeling. At the level of behavior modeled in this work, however, the mechanics are reasonably well understood. Modeled pedestrian motion is based on the Social Forces Model. [Helbing et al. 2000] Pedestrians are given the same errors in their perception of distance and velocity as in that of drivers. They do not, however, exhibit delayed reactions or distraction. Their decisions to cross are based on applying the gap-acceptance distribution found in [Brewer et al. 2006] to each lane of the road.

Pedestrian behavior is significantly simplified in a few ways. First, while such behavior has been repeatedly demonstrated in observational studies, pedestrians in the model do not alter either their walking speed or the length of gaps in traffic that they deem acceptable based on their waiting time. Second, once a simulated pedestrian

begins crossing, he is effectively no longer aware of approaching vehicles. This will significantly increase the number of serious conflicts and collisions produced by the model, since real pedestrians are (when not distracted) fully capable of altering their trajectory to avoid being run over. This could take the form of aborting a decision to cross or increasing their speed to make sure they are out of harm's way when vehicles arrive.

AGENT-BASED MODELING

The prospect of using computers to simulate complex processes dates to the earliest days of computing. [Metropolis and Ulam 1949] These early methods allowed for huge numbers of iterations through processes that combined deterministic and stochastic steps (the latter modeled by comparing the values of pseudo-random numbers to thresholds) to derive cumulative statistics of the end-state of the overall process. These focused on simple, numerical phenomena that were nevertheless intractable through either analytic solutions or by approximation to a continuum (e.g. the three-body problem of physics [Kalos 1962]). However, the applicability of these tools to transportation was quickly recognized. [Mathewson et al. 1955]

New sources of data and exponential increases in computational power have made feasible the explicit consideration of every action taken by road users. This new capacity is revolutionizing the control, modeling, and design of transportation systems worldwide. This approach can be applied to study the effects of individual drivers' decisions, to explore changes to infrastructure or control strategies in various scenarios, examine atmospheric emissions, test new communication forms, or extrapolate current conditions to aid in managing abnormal events or avoiding gridlocks. One promising new platform for this kind of research is the use of *agents*.

While there is little consensus on terminology, *multi-agent systems* can be regarded as a subset of *agent-based computing*. [Niazi & Hussain 2011] This is a multi-disciplinary field of research endeavors that defines rules for the internal behavior and

interaction of individual entities within a controlled, simulated environment. The use of such a CompuTerrarium [Epstein & Axtell 1996] allows observation of the effects that arise from changes to those rules. For a detailed inquiry in pursuit of defining and delineating the various facets of agent-based computing, see [Jennings 1999].

Arising from object-oriented programming [Dahl & Nygaard 1968], agent-based computing has deep roots in the field of artificial intelligence [Shoham 1993] but has been applied extensively in fields as diverse as ecology [Grimm et al. 2005], economics [Tesfatsion (ed.) 2006], and sociology [Axelrod 2005]. Using agent-based architectures is ideal for scenarios that feature distributed, localized decision making – as opposed to dynamics dominated by centralized control or strict physical laws, which are more amenable to strictly equation-based approaches. [Parunak et al. 1998]

The inherently localized perspectives of agents and the intrinsic ability to explicitly incorporate heterogeneity within such models makes transportation a natural application for the paradigm. It can be applied to all levels and scales of transportation modeling: from the decade-long scopes of travel demand models, to the choices of which route a driver will take on her daily commute, to the split-second decision of when to apply the brakes to avoid a collision. The breadth and scope of such models is also only limited by the researchers' imaginations (assuming the requisite time and computing power is available). Balmer, Nagel, and Raney [2004] demonstrated a platform capable of multi-day simulations of regions with tens of millions of agents that adapt throughout the course of the test run. At the other end of the spectrum, there is

almost no limit to the level of detail in psychological, perceptual, social, and physical behavior that can be included in a model.

Heterogeneity of behavior is not limited to that among drivers, either. Intra-driver variability can also be incorporated. Kesting, Treiber, and Helbing [2009] demonstrated an agent-based incorporation of the frustration drivers feel after spending time in congestion. Reichardt [2008] developed a “cognitive appraisal model of emotion” in which driver emotions vary based on environmental events, and these determine risk acceptance in driving behavior. For an overview of the various applications of agent-based computing in transportation, from multi-modal control and management to simulation, see [Chen and Cheng 2010]. In the context of this exploration, the platform allows for populations within a simulation to actualize behavior patterns based on internal variables that are derived from empirical distributions. By fitting these values to observed distributions, agent-based modeling enables the recreation of behaviors otherwise described by more subjective or categorical, *ad hoc* characteristics such as pedestrian “aggressiveness” or vehicle platooning. [e.g. Schroeder and Roupail 2011]

REPAST SIMPHONY

Agent-based microsimulation has been broadly applied in traffic modeling since its inception. There are a multitude of applications and a commensurate number of open-source and commercially-available platforms. The number of distinct approaches makes an exhaustive review impossible, so only a few examples are presented here.

Some simulations focus on network-scale models: TRansportation ANalysis and SIMulation System (TRANSIMS) is a development of the US Department of Transportation that uses travel demand models and link travel times. [Hope 2010] The Multi-Agent Transportation Simulation (MATSim) is an open-source model that uses queuing models instead of explicit car-following. [Horni et al. (eds.) 2016] VISSIM (a German acronym, the expansion of which translates to “traffic in cities – simulation model”) [PTVGroup.com], along with the host of transportation simulation platforms in the PTV family, is an industry standard. While many of the details of its underlying mechanics are trade secrets, its driver model is based on the Wiedemann psycho-physical model; [1974] and its pedestrian modeling is based on the same social force model [Helbing et al. 2000] used in this simulation.

While many commercially- or publicly-available platforms have modular components that can be adapted for specific applications, the interactions among these modules are constrained. For specific applications that are not easily explored through such mature systems, there are a variety of agent-based computing platforms available for building one’s own simulation. The selection from among these is generally one of choice in terms of the user interface, platform features, and the underlying programming languages supported. [Railsback et al. 2006] For an extensive list of simulation platforms available, see [Tsfatsion & Judd 2016]

One of the more popular platforms is NetLogo [Wilenski & Rand 2015], originally developed by Uri Wilenski at the Center for Connected Learning and Computer-Based

Modeling of Northwestern University. It was designed to be accessible to those without programming backgrounds and has a simple yet powerful interface. Swarm [Swarm 1999] is another powerful platform for simulating multi-agent systems that has been used effectively in modeling pedestrian movements. [Schelhorn et al. 1999]

For this study, the REcursive Porous Agent Simulation Toolkit (Repast) [Collier 2001] Symphony 2.5 framework (an Eclipse-based Integrated Development Environment, “IDE”) was chosen. Repast originated as a Java implementation of Swarm, developed at the Social Science Research Computing department of the University of Chicago, but has diverged into a stand-alone platform. [Railsback et al. 2006] This platform was chosen for its extensive documentation, speed, currency of implementation (2.5.0 was released in October 2017), and the quality of its IDE. [Getchell 2008] In a comparison of Java-sourced agent-based simulation platforms [Tobias & Hofmann 2004], Repast scored highest in a broad assessment of a variety of features.

MODEL IMPLEMENTATION

ROAD ENVIRONMENT

The model developed in the above sections is applied to passenger vehicles approaching an uncontrolled mid-block crossing with no specific treatments beyond a

demarcated crosswalk. The modeled road has four lanes, each 3.3 meters (11 feet)⁴⁶ wide and is 1.5km long. A diagram of the crossing is shown in Figure 8 – Modeled Crosswalk (drawn to scale). This length allows ample room before the pedestrian crossing for traffic to stabilize from any edge effects arising from the stochastic generation of vehicles.⁴⁷ The simulation of this road is spatially quantized to square cells of 0.5m. The space quantization is arbitrary, and its only effect is in the smallest values that the model can define. Java's *double* primitive class can hold positive values down to 2^{-1074} . Since this is many orders of magnitude smaller than the Planck length, choice of this parameter within reason will have no appreciable effects on any physical system being modeled. Distributions for all variables used in the simulations are given in Appendix A.

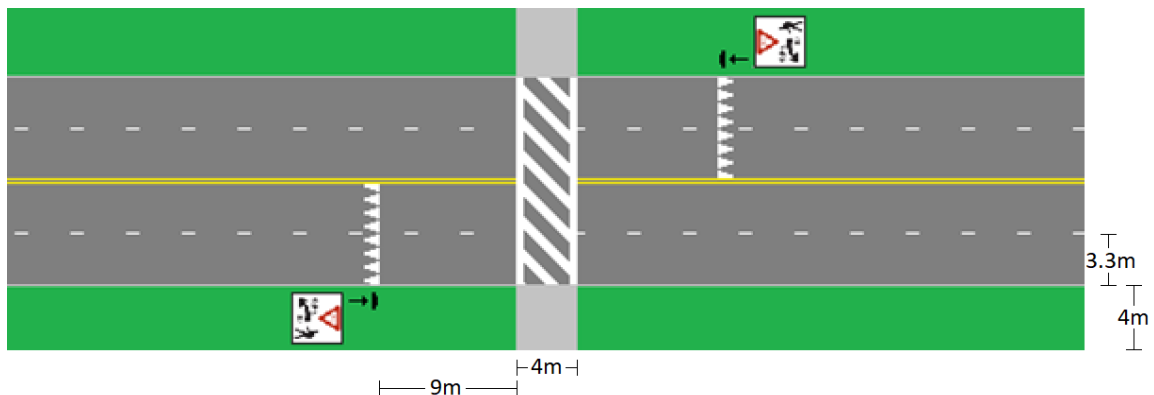


FIGURE 8 – MODELED CROSSWALK

⁴⁶ The standard recommended lane width for arterials is 10 to 12 feet. [AASHTO 2001]

⁴⁷ Vehicles generated by the Poisson process used may be placed nearly on top of one another. To facilitate stabilization, the physical limitations on emergency deceleration are disregarded within 150m of either end of the road.

SCHEDULER

The simulation is structured into a schedule of events. At every *tick* of the simulation clock, an explicit pattern of actions is taken. This is shown diagrammatically in Figure 9. Time within the simulation is quantized to 0.1 seconds per *tick*. This time scale can have significant effects if chosen to be too large, but it has been demonstrated that simulations of the IDM with a time step of 0.1 seconds accurately recreate time-continuous behavior, and do not differ significantly from simulations using time steps as low as 0.01 seconds. [Kesting and Treiber 2008b] Each simulation iteration was run for 50 hours of simulated time, or 1,800,000 ticks. Increasing the duration of each iteration of the simulation allows more time for the stochasticity in the model to average out. However, the real-world time required to run each iteration increases proportionately. A duration of 50 hours was chosen somewhat arbitrarily as a compromise between these two trends. Using this duration, averaging over five runs was sufficient to produce consistent means for any set of input parameters.

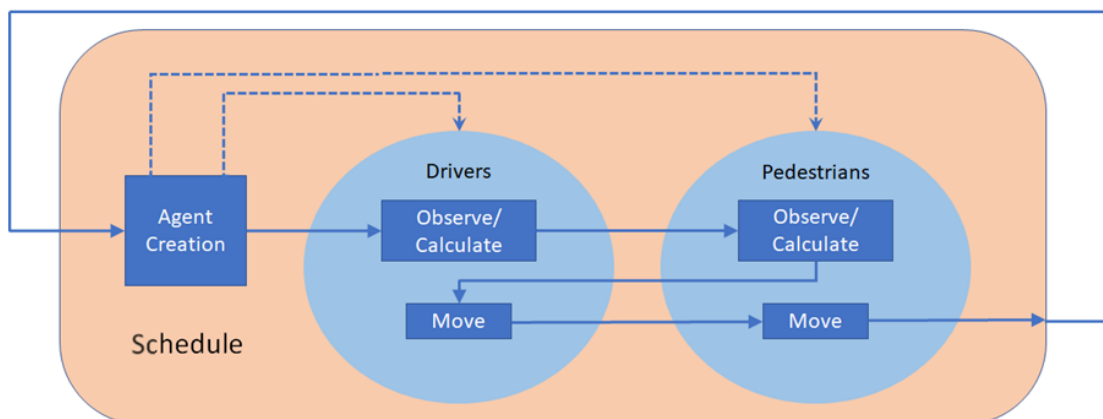


FIGURE 9 – SIMULATION SCHEDULE

The first action taken with every tick is the creation of new agents. Vehicle and pedestrian arrivals are modeled using independent Poisson probability mass functions. With an average number of events per interval, λ , the probability of observing k events in any interval is given by

$$P(k) = \frac{\lambda^k e^{-\lambda}}{k!} \quad [49]$$

$$P(1) = \lambda e^{-\lambda}, \quad P(2) = \frac{\lambda^2 e^{-\lambda}}{2}, \quad \dots$$

The model has four of these processes running simultaneously: one for each direction of both pedestrians and vehicles.⁴⁸ If the result of a uniform random number generator (bounded by 0 and 1) for a given direction of vehicles is below $P(2)$, a car is created in both lanes. If it is above this, but below $P(1)$, one car is created and placed in a random lane. No higher terms are calculated. While this is an unrealistic simplification, the effects are negligible. According to the NCHRP, [Zegeer et al. 2008] the maximum expected flow rate for any functional area classification is 1900 vehicles/hour/lane. With this rate, the combined contribution of all higher terms of the Poisson probability function over the course of a 50-hour simulation is fewer than 85 vehicles – a deviation of approximately 0.02%. With a maximum typical flow rate of 400 pedestrians/hour from the same source, the number of pedestrians not added is less than one per 50-hour simulation.

⁴⁸ Therefore, in practice, the value of λ passed to each generation process is half the overall value of the system.

After the creation of new agents, those already within the simulation observe their environment and calculate the acceleration they will take. Once all agents (drivers and pedestrians) have completed these calculations, they move based on the newly calculated velocities. Any agent that moves beyond the edges of the simulated world is then removed from the environment. Agent cognition is a series of functional modules.

DRIVERS

Drivers are created at the edge of the simulated road with parameters determined by appropriate distributions. Their accelerations are not limited to realistic values while they are within 10% of the road length of their starting edge. This is to avoid undesirable behavior when the Poisson arrival pattern creates cars that physically overlap. They also do not become distracted if their distance from either edge of the simulated road is within 7% of the road length. This is to avoid piling up at either end of the road. Inside these boundaries, the calculation proceeds for each individual as described in the following sections.

CALCULATION

The first step in driver calculation is determination of the distracted state. Drivers are created giving their full attention to the driving task. They then calculate the ultimate duration of this inter-distraction period based on the global Poisson arrival rate of distraction events. The time since the beginning of this period is tracked, and when the duration is exceeded the driver enters a distracted state. The duration of this state is calculated from the global lognormal distribution of distraction events, and the time

since the distraction began is tracked. When this period ends, the driver again becomes aware of her surroundings, and the cycle begins anew. Details of the acceleration patterns in distracted and normal states, as well as derivations of the distributions for these states, are given in the *Driver Modeling* section.

When not distracted, the EIDM acceleration is calculated for the current position and speed relative to the nearest leading vehicle (using values that include errors in speed and distance, as described in the section on *Modeling Human Factors*). This acceleration is stored for later actualization, as described in the *Delayed Reaction* section on ADRT. The acceleration calculated at the appropriate previous point in time is passed to the next steps.

Next, yielding behavior is calculated. If a crossing pedestrian is within the perception limit of a driver, and the driver's view of the pedestrian is not obstructed by any vehicles ahead of it or traveling in the opposite direction (details of this visual obstruction are discussed below), the behavior of the pedestrian is observed. These steps are repeated for all pedestrians that have signaled their intent to cross or have already begun crossing.

If a pedestrian has indicated an intent to cross and will be within a conflict distance (defined here as the pedestrian being within the driver's lane or one of the adjacent lanes) when the driver would cross the crosswalk at her current speed, the driver will decide to yield if doing so is possible at a comfortable deceleration. This value

of “comfortable deceleration,” b_{yield} , is set at a constant 3m/s^2 based on data from Bella and Silvestri [2016]. If the pedestrian has already begun crossing, and their paths bring the two agents within a conflict distance then, also, the driver will yield the right of way. However, if the pedestrian’s current speed will take him outside the vehicle’s adjacent lane by the time the vehicle enters the crosswalk, no change in velocity is required. The details of this deceleration behavior and where the car will attempt to come to a stop are given in the section on *Yielding Kinematics*.

The deceleration calculated for yielding is delayed, as described in the *Delayed Reaction* section on BRT. It bears repeating that, while the decision to yield is delayed, the deceleration passed from this module is based on the current position and velocity. If the delayed acceleration were used, drivers would consistently collide with pedestrians to whom they were attempting to yield.

If the driver is not currently distracted, the deceleration value calculated in the yielding module is compared to the output of the EIDM acceleration calculation. The lower of these two is added to the current velocity to determine how far the vehicle will move in the impending tick. It should be noted that the yielding state decided upon is stored as an enumerated value along with the calculated acceleration: $-1 = \text{no yield}$, $0 = \text{soft yield}$, $1 = \text{hard yield}$. This value cannot be reduced while the pedestrian in question is still crossing. This is to ensure that, after coming to a stop, the driver will not decide that a collision is no longer imminent and begin to accelerate again based on the car-following module.

If the vehicle ahead of a driver has come to a stop without enough room for the driver to fit behind it without blocking the crosswalk, or said lead vehicle is decelerating in a manner that will result in such a situation, the driver will attempt to stop prior to the crosswalk.

The final step for each driver agent is observation of any potential conflicts. To record these, time to collision (TTC) is calculated for all crossing pedestrians. Any TTC below the threshold for a conflict (1.5 seconds) is added to a running log. This log is the output of all simulation runs.

PEDESTRIANS

Pedestrians are created using Poisson processes that are equivalent to, but independent from, those used to generate automobiles. They are generated at the top and bottom of the world, 750m from either end of the road. The point of creation is 4m from the road's edge. The calculation patterns discussed below are enacted by each pedestrian agent at each tick of the schedule clock.

CALCULATION

The models and distributions discussed in this section are explained in detail in the section on *Pedestrian Modeling*. Pedestrian distraction, though a significant issue, [Mwakalonge, Siuhi, and White 2015] is not included in the model. The fundamental mechanics of pedestrian motion are based on the Social Forces Model. [Helbing et al. 2000] Each pedestrian feels a motive force toward a point on the opposite side of the

road. This attractor point is shifted a meter off-center to add a small transverse force, avoiding the potential artifact of exactly balanced forces impeding pedestrian movement. The shift of this point is enacted such that pedestrians pass those going the opposite direction to their own right.

Until a pedestrian has made the decision to cross the street, a barrier force from the curb is acting upon him. This force stops the pedestrian at the road's edge, where all pedestrians gather within a 4x4m box – also enforced via a barrier force. In addition, each pedestrian is acted upon by a repulsive force from any other pedestrians ahead of him. If multiple pedestrians are waiting to cross, this produces a small crowd of people instead of a two-dimensional line waiting at the curb.

From the time of their creation, pedestrians begin observing any approaching vehicles. While vehicles can block the line of sight between a pedestrian and other approaching cars (discussed in greater detail in the next section), pedestrians remain aware of any approaching vehicles they have seen at any time in the past. (The same is not true of drivers, who react only to pedestrians they can currently see.)

At each step, for each lane, the arrival time of the nearest approaching vehicle that the pedestrian is aware of is calculated (including perception errors). Using this, and knowledge of his own acceleration and gap-acceptance behavior, the pedestrian decides if he has enough time to cross in front of each vehicle. If the answer is yes, the pedestrian is said to be accepting a *lag*. If not, and the vehicle will have passed by the

time the pedestrian is within one lane of the vehicle's lane, the pedestrian calculates the arrival time of the vehicle behind the nearest one. If there is sufficient time for the pedestrian to cross between the two, the pedestrian still decides he can cross this lane safely – this is referred to as accepting a *gap*. If the situation in all four lanes is deemed as safe, then the pedestrian will signify intention to cross.

If a vehicle is close enough when this decision is made (within 18m – a somewhat arbitrary distance equal to twice the recommended stopping distance for yielding vehicles), the pedestrian will wait to be acknowledged by the nearby vehicle before crossing in front of it. Further details on this decision-making process are presented in the section on *Pedestrian gap acceptance*.

When a pedestrian has decided to begin crossing, the barrier force of the curb is removed, and he begins walking toward his attractor point on the opposite side of the road. The repulsive forces from pedestrians in front of him are still felt, so pedestrians navigate around each other while crossing in opposite directions. This interaction is not enforced for pedestrians still waiting to cross from the other direction. Otherwise, the combined force of a waiting group would prevent any lone pedestrians from achieving the safety of the opposite curb.

If any vehicles have come to a stop in a pedestrian's path along the crosswalk (that aren't currently yielding to him), a second attractor point is temporarily added to the forces acting on the pedestrian when he comes within a lane's width from the

vehicle's side. This point is located a half-meter behind the rear corner of the vehicle. If the driver of the vehicle stopped on the crosswalk has indicated to the pedestrian that she intends to yield to him, this point is assigned to a half-meter ahead of the front corner of the vehicle. Once past the vehicle, this second attractor point is removed. This, in combination with a barrier force at the edges of vehicles keeps pedestrians from walking through (or, in effect, climbing over) vehicles in their paths.

VISUAL OBSTRUCTIONS

In addition to the human factors described in the *Driver Modeling* section, when a vehicle is close to the crosswalk, another limitation to driver perception of crossing pedestrians becomes relevant: automobiles are mostly opaque. If a vehicle is slowing or stopped to allow a pedestrian to cross, its presence obstructs the line of sight of any vehicles in adjacent lanes behind it. Likewise, if a vehicle has recently traversed the crosswalk but is still in its vicinity, drivers traveling in the opposite direction may not see pedestrians crossing behind it. The drivers of these other vehicles will not be aware that they should be yielding, resulting in the potential for conflicts. The former scenario is referred to as a *multiple threat*, [Snyder 1972] and a diagrammatic example is shown in Figure 10 (pedestrian not drawn to scale). In a nationwide study of pedestrian crashes, [Zegeer et al. 2005] 17.6% of the crashes in marked crosswalks were classified as resulting from multiple-threat events. The stochastic nature of these events makes simulation the ideal approach for modeling them.

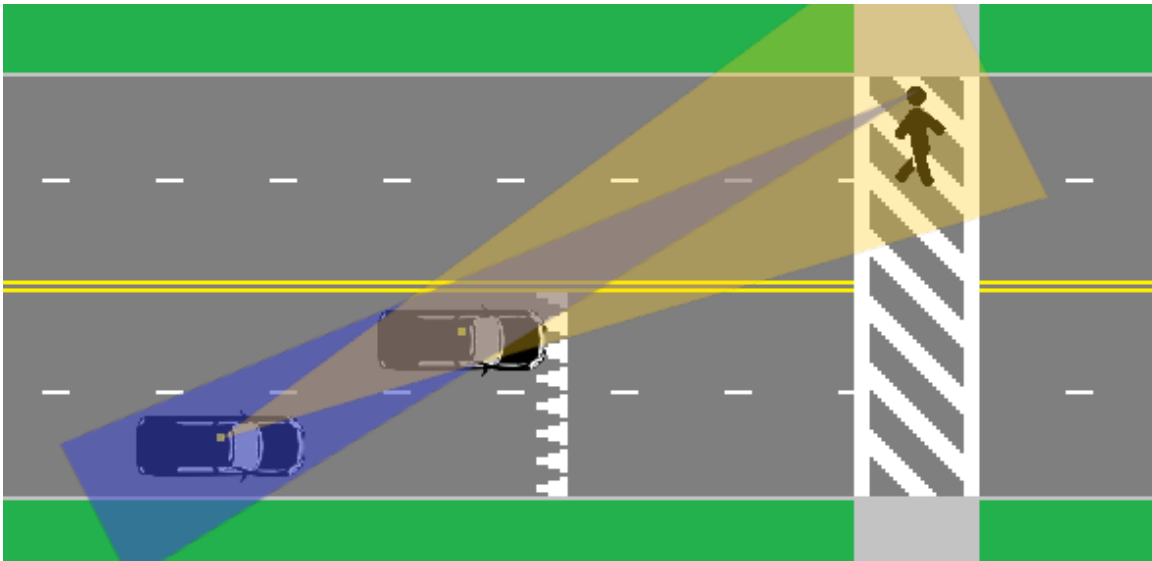


FIGURE 10 – MULTIPLE THREAT

The dimensions for vehicles in the simulation are taken from Edmunds data on new car dimensions. [Woodyard 2007] For simplicity, each vehicle is identically modeled using the average values for “large sedans” between the years 1990 and 2007: 5.28m long, 1.89m wide, and 1.46m tall. This is shown in Figure 11 (drawn to scale).

Pedestrians are assumed to be the average of American male and female heights: 1.4m, with 0.11m from the top of the head to their eyes.

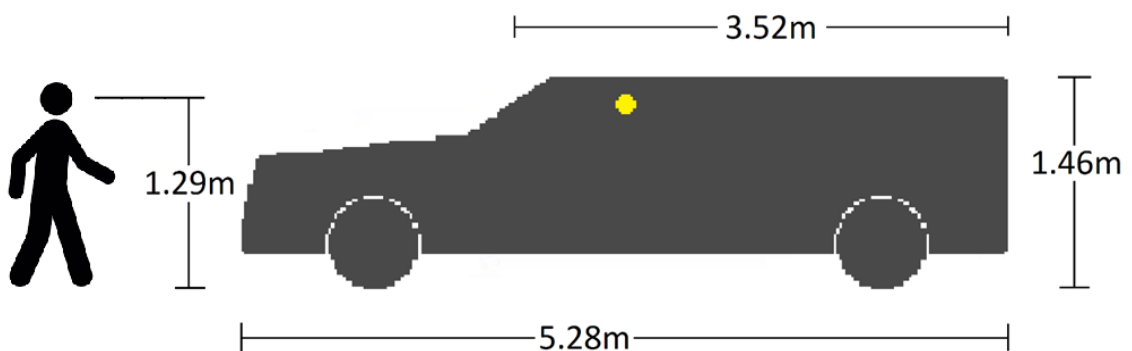


FIGURE 11 – MODELED VEHICLE AND PEDESTRIAN

This puts their eyes at 1.29m above the ground. For simplicity, drivers are assumed to be seated with their eyes at the same height. Both drivers and pedestrians can see over the hood of the vehicles, so only the rear 2/3 of each vehicle is assumed to obstruct the view of drivers and pedestrians around it. The vantage point from which a driver's vision is calculated lies halfway between the front and rear bumpers of her vehicle, and 1/3 of its width (0.63m) from the left side.

CONFLICTS

Conflicts are recorded by each driver agent after all other calculations are made. The time travel time to the longitudinal location of each crossing pedestrian is calculated based on the driver's newly-calculated velocity. If, at that point in the future, any crossing pedestrian's current velocity will have brought him within his radius of the transverse location of the vehicle, that is considered a time to collision (TTC). While it could be argued that a near-miss would still be considered a "conflict" for a pedestrian, this original conception of TTC is used to enable comparison of the output to other studies.

Any such conflict with a TTC of 1.5 seconds or less is recorded in a log, which forms the output of a simulation run. Included in this log are state variables of the pedestrian and driver involved (time since the driver was distracted, the pedestrian's critical gap, the driver's BRT, etc.) as well as a description of their relative motion (current driver speed, acceleration, and longitudinal distance from the pedestrian). Also included, for differentiation between collisions and near misses, is the lateral distance

between the pedestrian and the side of the vehicle. If a collision occurs, this last variable is 0m.

In the analysis of these data, a TTC of 1.5 seconds or less is considered a *conflict*, and a TTC of 1 second or less is considered a *serious conflict*. This is in keeping with the Dutch Objective Conflict Technique for Operation and Research. [Kraay, van der Horst, and Oppe 1986]

FACE VALIDATION

Validation describes checking whether a model, as implemented, gives a “reasonably accurate representation of the real world.” This is in contrast to *verification* which refers to the “process that determines whether the programming implementation of the abstract or conceptual model is correct.” [Xiang, Kennedy, and Madey 2005] By adding breakpoints to relevant points in the code, any portion of the model can be observed to qualitatively verify that it behaves as expected. Also, the modeling platform used, Repast Symphony, features an integrated visualization of the model being run. This makes it straightforward to continually check the face validity of the model’s performance. However, face validity is only a subjective measure of performance.

FUNDAMENTAL DIAGRAM

Data on the macroscopic behavior of traffic in the simulation can be used to make a more quantitative assessment of the model (though comparison of this to empirical data is still done qualitatively). The performance of the driver model can be

checked by generating a fundamental diagram of the simulated traffic flow.⁴⁹ While the introduction of driver errors, delays, and heterogeneity in behavioral parameters creates density waves in the simulation's flow of traffic, the congested states needed to fill out the fundamental diagram were absent. Without bottlenecks or lane changes (which would further complicate pedestrian-vehicle interactions in ways that have not received significant research), no disturbances were able to grow into large-scale breakdown of flow. It has been demonstrated, however, that inhomogeneities in the road can effectively reproduce the disturbances of lane-changes on the collective dynamics of congested traffic. [Treiber, Hennecke and Helbing 2000]

If the location chosen to measure these quantities is upstream of the crosswalk and beyond the limit of drivers' ability to perceive pedestrians, density waves created from drivers yielding to pedestrians will be added to those produced by the introduction of driver errors, delays, and heterogeneity. This enables a demonstration that the driver model, as implemented, recreates realistic traffic flow.

To calculate the relevant values, two imaginary lines are placed along the road at x_1 and x_2 (a distance $dx = 5m$) apart. For each vehicle, i , the time, $t_{i,1}$, at which the vehicle crosses x_1 is recorded. Then, the time, $t_{i,2}$, when the vehicle crosses x_2 is recorded, and the difference, $dt_i = t_{i,2} - t_{i,1}$, is calculated. The number of vehicles, n ,

⁴⁹ For detailed definitions of the quantities discussed, see the *Driver Modeling* section on *Macroscopic Models*.

passing x_2 in a time period, Δt (set to 120 seconds), is counted; and the time, dt_i , that each took to cross the distance dx is recorded. The *flux* is then given by:

$$q = n/\Delta t \quad [50]$$

The *concentration* or *density* of the road, the average number of vehicles per unit length, is calculated as:

$$\rho = \frac{\sum dt_i}{\Delta t * dx}; \quad [51]$$

and the velocity of the flow is calculated as the *harmonic mean speed*:⁵⁰

$$V = \left(\frac{1}{n} \sum \frac{1}{dx/dt_i} \right)^{-1} \quad [52]$$

By measuring these quantities, the fundamental diagram can be created. The results are shown in Figure 12. These measurements are not taken until after a warm-up period equal to twice the length of the simulated road, divided by the speed limit of the simulation instance. For example, with a posted speed limit of 20 kph, this is nine minutes. For a speed limit of 70 kph, this is approximately 2.6 minutes.

Comparing the resulting shape to the shape of empirically observed traffic flow gives some validation of the implemented model. For this purpose, data collected on a freeway in Georgia by Wang et al. [2009] are presented in Figure 13. Congested portions of the fundamental diagram are only filled in when the model is run at low speeds, so

⁵⁰ The harmonic average is chosen because the arithmetic mean is biased towards faster vehicles: Since the measurement points are stationary, more fast cars would pass in a given amount of time, skewing the results. The harmonic mean, on the other hand, inherently weights each measurement by $\frac{1}{v}$. [Knoop, Hoogendoorn, and Zuylen 2009]

the data in Figure 12 are taken from simulations run with a speed limit of 20kph and an average flow of 300 vehicles/lane/hr. The granularity in the data (the visible lines of data points) is due to the process of counting passing cars. Only whole cars can be counted, so the possible values of flux are quantized accordingly.

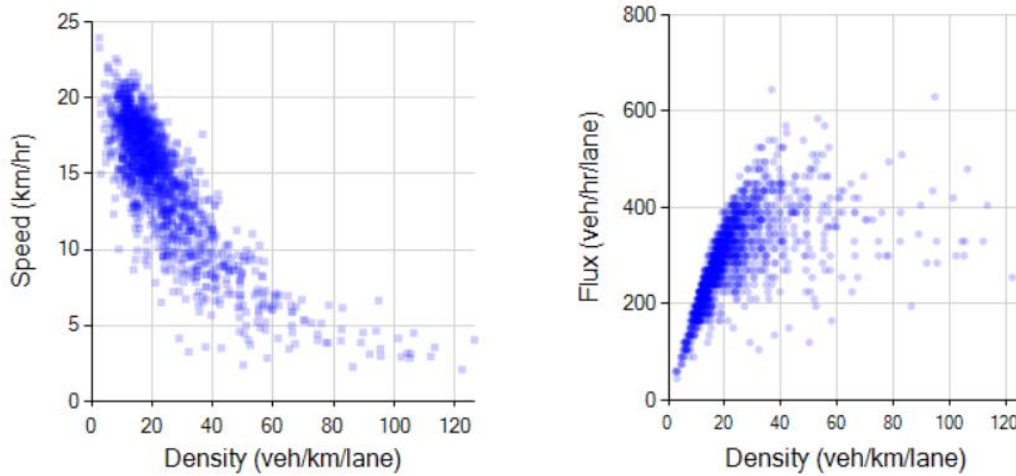


FIGURE 12 – SIMULATED FUNDAMENTAL DIAGRAM

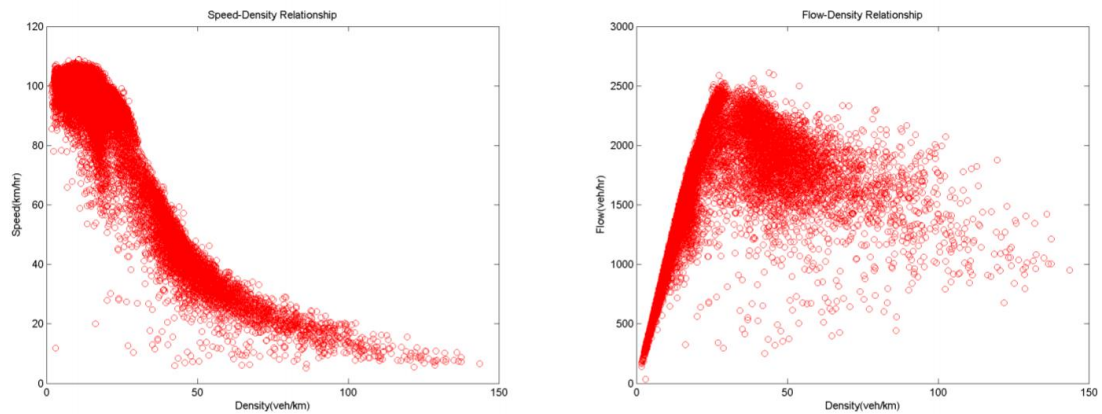


FIGURE 13 – EMPIRICAL FUNDAMENTAL DIAGRAM [WANG ET AL. 2009]

RESULTS

SIGNIFICANCE OF INCLUDING HUMAN FACTORS

The above model was built to address a primary question: To what extent does inclusion of human factors affect the results when using simulations to answer questions related to pedestrian safety? To simplify this stage of the analysis, a single set of agent volumes was chosen. The developed model was run for a pedestrian volume of 400 PPH (*pedestrians per hour* – 200 per hour from each side of the road) and 800 VPH (*vehicles per hour* – 400 per hour in each direction). These illustrative rates were chosen as a compromise: Higher rates create higher conflict counts, and therefore more robust results. However, higher rates require more computational time. Also, if the rates are too high, agent generation exceeds the capacity of the road and the generation rates stop being directly related to how many vehicles move through the crosswalk.

The speed limit was set in increments of 10 kilometers per hour (kph) from 20 to 70kph. Below this speed range, this combination of vehicle and pedestrian rates was higher than the capacity of the intersection. No higher speeds were analyzed since, as recommended in one of the benchmark studies on the topic, [Zegeer et al. 2005], unsignalized crosswalks should not be used on roads with speed limits above 40 mi/hr (~64.4kph). Many of the included combinations of speed and volume (pedestrian and vehicle) are well beyond where the Manual on Uniform Traffic Control Devices [FHWA 2009] warrants a pedestrian signal. However, since these warrants are in place to

minimize vehicle-pedestrian conflicts, observation of conflicts will be most informative beyond their bounds.

In the tables and figures that follow: minimum $TTC \leq 1.5s$ is a *conflict*; minimum $TTC \leq 1.0s$ is a *serious conflict*; and minimum $TTC = 0$ is a *collision*. These will be collectively referred to as *safety-critical interactions*. (Minimum TTC is defined in detail in the *Literature Review* section on *Conflicts*.) Each data point in these figures represents the average hourly rate of these interactions over 250 hours of simulated time (five 50-hour simulations for each combination of speed and human factors). 50-hour runs were chosen somewhat subjectively based on their producing similar outputs from run to run. In shorter simulations, the stochastic nature of the model produces larger discrepancies between runs based on the same input parameters. 50 hours was chosen as a balance between avoiding this behavior and computational time.

At these rates of agent creation, the average rates of safety-critical interactions are approximately linear with respect to posted speed limit, PSL, so linear regression was used to fit the observed interactions to a linear model:

$$y = \beta_0 + \beta_1 \times PSL \quad [53]$$

In testing this linear regression, the null hypothesis is that the coefficients, β_0 and β_1 , are zero. The alternative hypothesis is that the relationship in Equation [53] exists. Thus, if the derived *p-value* of the fit is less than a given level of significance, the null hypothesis is rejected, and the regression is valid. The coefficient of determination, R^2 , is the ratio of the amount of variability in the data explained by the regression

model to the actual amount of variability in the data. For each case, the derived values are included in a table. For reference, the distributions of all model parameters used in the simulations are given in Appendix A.

BASELINE

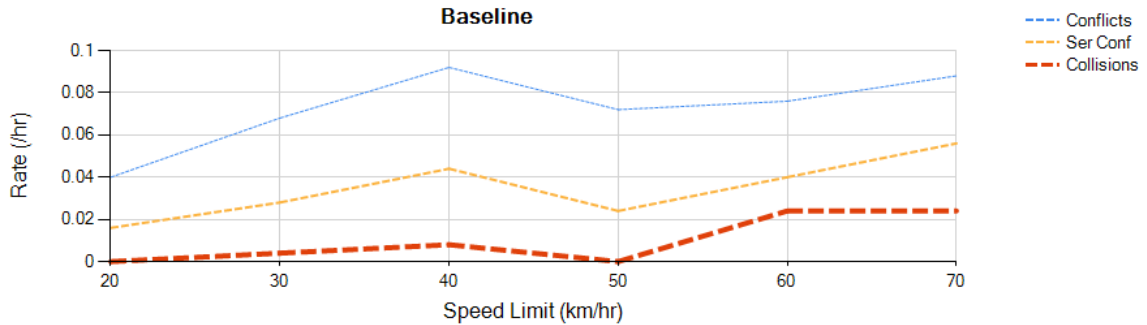


FIGURE 14 – SIMULATION RESULTS: BASELINE

	Value	p-value	Adjusted R ²
β_0 Conflicts	0.0412952	0.0318	0.07555
β_1 Conflicts	0.0006971	0.0770	
β_0 Serious Conflicts	0.0068952	0.48871	0.2191
β_1 Serious Conflicts	0.0006171	0.00531	
β_0 Collisions	-0.0121143	0.0186	0.4401
β_1 Collisions	0.0004914	3.87e-05	

TABLE 5 – LINEAR REGRESSION: BASELINE MODEL

Without human factors, no trends are apparent, aside from serious conflicts and collisions being slightly more likely at high speeds. The only source of conflicts without driver errors are pedestrians accepting unsafe gaps. These result from the continuous distribution from which each pedestrian's critical gap is derived. Despite pedestrian distraction not being explicitly modeled in the simulation, this tail of the gap acceptance distribution may represent pedestrians that are unaware of their surroundings.

FULL MODEL

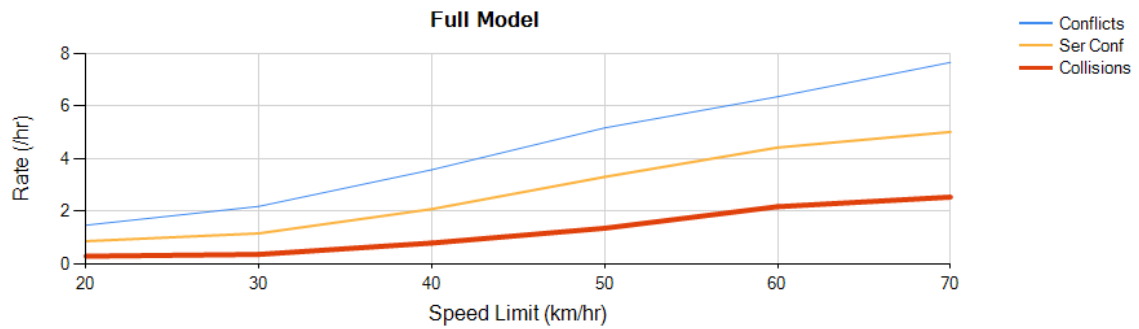


FIGURE 15 – SIMULATION RESULTS: FULL MODEL

	Value	p-value	Adjusted R ²
β_0 Conflicts	-1.3777	4.21e-09	0.9799
β_1 Conflicts	0.1287	< 2e-16	
β_0 Serious Conflicts	-1.26815	2.05e-09	0.9685
β_1 Serious Conflicts	0.09079	< 2e-16	
β_0 Collisions	-0.95745	6.68e-09	0.9339
β_1 Collisions	0.04929	< 2e-16	

TABLE 6 – LINEAR REGRESSION: FULL MODEL

All linear fits in the full model are significant, and the adjusted-R² values suggest a linear fit is sufficient to describe the variance in the response curve.

HUMAN FACTORS ANALYSIS

Each human factor in the driver model (estimation errors, distraction, visual obstruction, and finite reaction times) was analyzed in isolation from two directions. First, the output of the model with only that factor included was compared to the output with no human factors included, the *baseline*. Next, comparison was made between the output with all human factors included, the *full model*, and the output with all except the one factor under exploration.

To demonstrate that the difference between including or excluding a single human factor is statistically significant, two-way Analysis of Variance (ANOVA) techniques were employed. For the results of these tests to be valid, three criteria must be met. First, all observations must be independent of one another: Each run of the simulation was initiated with a randomly-selected random seed, so each data point is entirely isolated from the next. Second, the residuals about the fitted means must be normally distributed. To test this, the residuals of the ANOVA test were analyzed using the Shapiro-Wilk test.

In the Shapiro-Wilk test, the null hypothesis is that the values examined are normally distributed. Therefore, if the derived p-value is greater than a chosen level of significance, the null hypothesis is not rejected, and normality can be assumed. A significance of 0.05 was chosen for this test. The p-value of this test is given in the table for each factor's analysis. A corrolarial deduction can be made from this observation: if the residuals about the mean of each parameter set are normal, then the conditional distributions are also normal.

At low speeds, however, the number of collisions is close to zero. Since this puts a hard floor on the output values, the residuals for these situations are skewed away from zero, violating the normality assumption. Where this is the case, the ANOVA is run twice: once with the full range, and once using only speeds for which the normality criterion is satisfied.

The final criterion for the results of an ANOVA test to be trustworthy is that variances of the test, relative to the calculated means of each variable, are equal. This is referred to as *homoscedasticity*. Since the variance in the conflicts is related to the mean (and this has been shown above to vary linearly with the speed), this criterion is violated. However, as demonstrated by Quensel [1947], violation of this criterion will only have an effect if the sample sizes of each group are significantly different. To avoid this, five runs of each combination were used in the following analysis.

The null hypothesis for ANOVA is that the two sample sets being compared come from the same distribution. Thus, if the p-value calculated is below the chosen level of significance, the null hypothesis can be rejected. The analysis of each human factor discussed below includes results of both the linear regression performed and ANOVA. Since the relationship to speed is shown from the linear regression analysis, only the results relative to inclusion of the human factor (1 degree of freedom) are reported.

ESTIMATION ERRORS

Pedestrians and drivers are prone to imperfect perception of the distance and relative velocity of others on the road. When this human factor is included, the values of Δv and s in the Enhanced IDM (Equations [8] - [12]) are replaced with the driver estimates of these values: Δv^{est} and s^{est} , respectively. These values are calculated in Equations [22] and [24] and vary with each time-step based on an approximated Weiner process (Equation [30]). This is essentially a “random walk,” and can be thought of as noise with inertia.

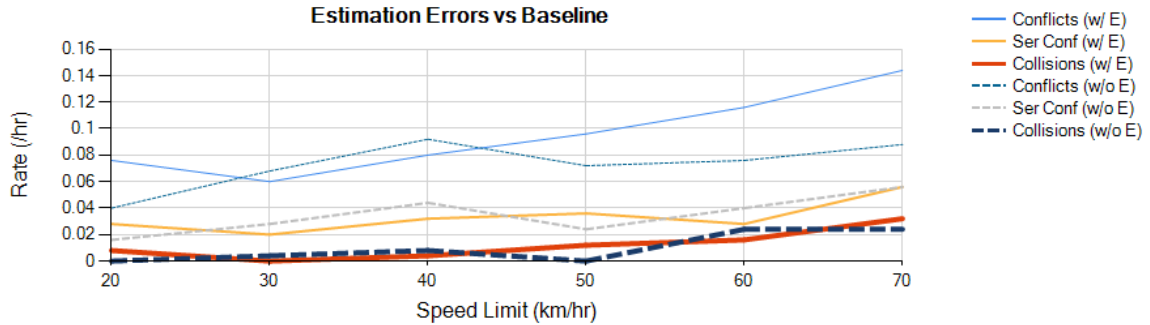


FIGURE 16 – SIMULATION RESULTS: ESTIMATION ERRORS VS. BASELINE

	Value	p-value	Adjusted R ²
β_0 Conflicts	0.027962	0.133004	0.3394
β_1 Conflicts	0.001497	0.000434	
β_0 Serious Conflicts	0.01173	0.307	0.09929
β_1 Serious Conflicts	0.00048	0.050	
β_0 Collisions	-0.0106286	0.15051	0.2627
β_1 Collisions	0.0005029	0.00223	

TABLE 7 – LINEAR REGRESSION: ESTIMATION ERRORS ONLY

The output of regression analysis on the effects of including only driver estimation errors shows that speed significantly affects the number of safety-critical interactions. However, the linear model does not account for much of the variability in the data (low R²). In fact, the effect is barely distinguishable from the stochasticity of the simulation. As would be expected, the intercepts are not statistically different from 0.

	S-W p-value	F value	p-value
Conflicts	0.3094	5.974	0.0182
Serious Conflicts	0.4355	0.065	0.80058
Collisions	5.757e-05	0.486	0.489
Collisions (SL ≥ 30kph)	0.06313	0.063	0.803

TABLE 8 – ANOVA: ESTIMATION ERRORS VS. BASELINE

Including estimation errors in isolation makes no significant difference to the rate of serious conflicts or collisions, and doing so has only a marginally-significant effect on conflicts.

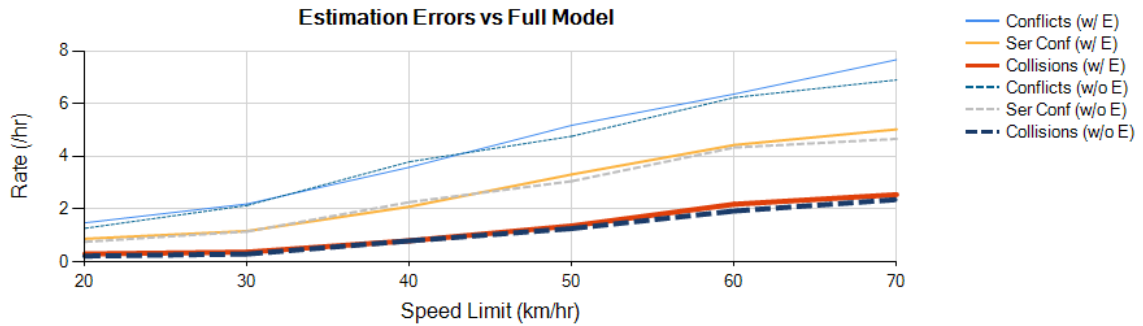


FIGURE 17 – SIMULATION RESULTS: ESTIMATION ERRORS VS. FULL MODEL

	Value	p-value	Adjusted R ²
β_0 Conflicts	-1.1368	5.14e-06	0.9646
β_1 Conflicts	0.1182	< 2e-16	
β_0 Serious Conflicts	-1.13638	1.24e-06	0.9444
β_1 Serious Conflicts	0.08537	< 2e-16	
β_0 Collisions	-0.9214	1.11e-09	0.9404
β_1 Collisions	0.0459	< 2e-16	

TABLE 9 – LINEAR REGRESSION: FULL MODEL EXCLUDING ESTIMATION ERRORS

	S-W p-value	F value	p-value
Conflicts	0.3264	7.707	0.00782
Serious Conflicts	0.4261	3.147	0.0824
Collisions	0.7566	9.825	0.00293

TABLE 10 – ANOVA: ESTIMATION ERRORS VS. FULL MODEL

When the other human factors are included, however, this approximation of the limitations of human visual ability makes a small but significant difference in conflicts and collisions.

VISUAL OBSTRUCTION

Drivers and pedestrians cannot see through cars. This was modeled by excluding any road users from an agent's decision-making processes if the former are within a visual angle obstructed by intervening vehicles. For the baseline, all nearby road users are included in behavioral decisions, regardless of line of sight restrictions.

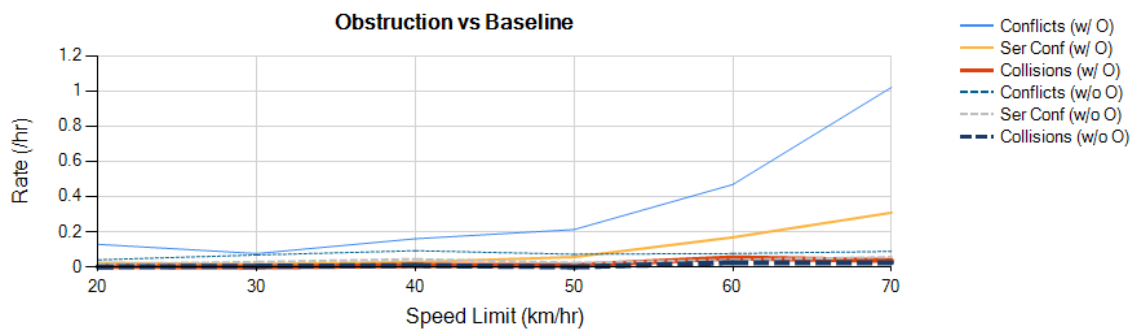


FIGURE 18 – SIMULATION RESULTS: VISUAL OBSTRUCTION VS. BASELINE

	Value	p-value	Adjusted R ²
β_0 Conflicts	-0.38731	0.000435	0.6876
β_1 Conflicts	0.01625	9.1e-09	
β_0 Serious Conflicts	-0.14480	3.22e-05	0.7314
β_1 Serious Conflicts	0.00544	1.07e-09	
β_0 Collisions	-0.0226286	0.00538	0.5291
β_1 Collisions	0.0009029	3.18e-06	

TABLE 11 – LINEAR REGRESSION: VISUAL OBSTRUCTION ONLY

While the linear regression for all safety-critical interactions is significant, the linear regressions do not describe a great deal of the variation. Visual obstructions delay the information that a pedestrian is crossing into the path of an oncoming vehicle. This translates to a decrease in the distance available for the driver to come to a stop. Since the necessary acceleration for stopping within a given distance is related to the square

of the velocity (Equation [41]), and this acceleration is limited to realistic values in the model, a quadratic relationship may be a better fit.

	S-W p-value	F value	p-value
Conflicts	0.5498	409.0	<2e-16
Serious Conflicts	0.1182	165.59	<2e-16
Collisions	0.0006965	13.091	0.000712
Collisions (SL \geq 50 kph)	0.3578	24.0	5.37e-05

TABLE 12 – ANOVA: VISUAL OBSTRUCTION VS. BASELINE

Including visual obstructions as the only human factor in the model makes a significant difference in all safety-critical interactions, though only at high speeds.

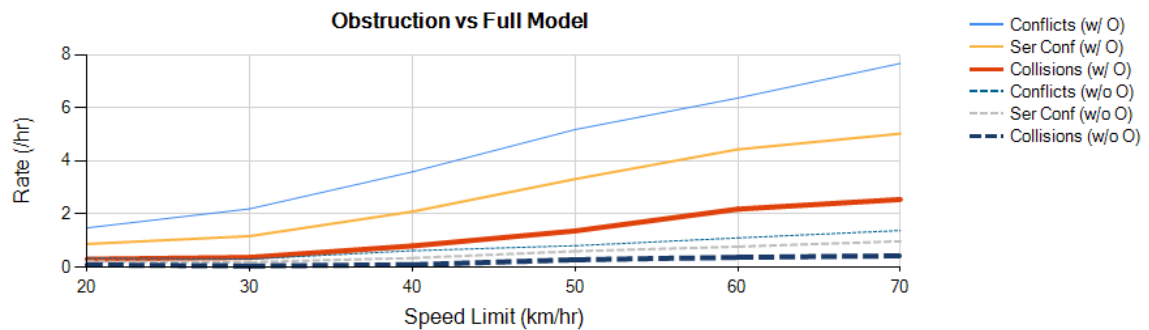


FIGURE 19 – SIMULATION RESULTS: VISUAL OBSTRUCTION VS. FULL MODEL

	Value	p-value	Adjusted R ²
β_0 Conflicts	-0.24110	0.00487	0.8631
β_1 Conflicts	0.02222	7.92e-14	
β_0 Serious Conflicts	-0.23177	0.000861	0.8503
β_1 Serious Conflicts	0.01662	2.79e-13	
β_0 Collisions	-0.151924	0.000769	0.7666
β_1 Collisions	0.008206	1.47e-10	

TABLE 13 – LINEAR REGRESSION: FULL MODEL EXCLUDING VISUAL OBSTRUCTION

	S-W p-value	F value	p-value
Conflicts	0.9607	4277.9	<2e-16
Serious Conflicts	0.8427	4353.8	<2e-16
Collisions	0.04618	1467.0	<2e-16
Collisions (SL \geq 30 kph)	0.1706	1421.7	<2e-16

TABLE 14 – ANOVA: VISUAL OBSTRUCTION VS. FULL MODEL

When excluded from the full driver model, however, the influence of obstruction is unquestionably significant. The model predicts a rate of safety-critical interactions an order of magnitude lower when obstruction is excluded.

REACTION TIME

Human actors do not react instantaneously to information from their environment. This limitation was modeled by delaying drivers' decisions to yield to crossing pedestrians based on empirical distributions of brake reaction times. For the baseline, the output of calculations for yielding decelerations (Equations [41] and [43]) is applied immediately for yielding vehicles. When brake reaction times (BRTs) are incorporated in the model, however, these outputs are delayed (by means of an iterated coupled map, as described in the *Driver Modeling* section on *Modeling Human Factors*) until after a time, τ_{RT} . This total reaction time is equal to the driver-specific *perception reaction time* (τ_{BRT} – given by Equation [21]) plus the *movement time* (τ_M) and *device response time* (τ_{DR}). (The latter two, combined, add a constant 0.35 seconds.) In both cases, acceleration/deceleration reaction times (ADRT – given by Equation [14]) are still applied by means of a delay-differential equation. The effects of ADRT on pedestrian safety are indirect. This was done to avoid confounding these effects with those of BRT.

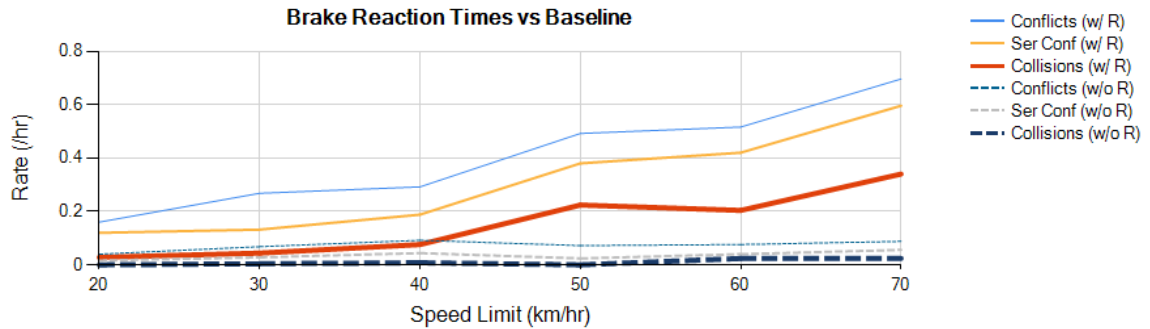


FIGURE 20 – SIMULATION RESULTS: REACTION TIMES VS. BASELINE

	Value	p-value	Adjusted R ²
β_0 Conflicts	-0.06194	0.0884	0.8738
β_1 Conflicts	0.01035	2.54e-14	
β_0 Serious Conflicts	-0.135771	0.000334	0.874
β_1 Serious Conflicts	0.009817	2.48e-14	
β_0 Collisions	-0.128648	4.19e-05	0.8149
β_1 Collisions	0.006251	5.57e-12	

TABLE 15 – LINEAR REGRESSION: REACTION TIMES ONLY

While the linear component of the rates of safety-critical interactions with respect to speed limit is significant, the magnitude of these effects for reaction times included in isolation is small relative to the stochasticity of the simulation. The linear fit does not account for a great deal of the variation.

	S-W p-value	F value	p-value
Conflicts	0.06101	665.79	< 2e-16
Serious Conflicts	0.482	862.75	< 2e-16
Collisions	0.0004721	389.75	< 2e-16
Collisions (30kph ≤ SL ≤ 60kph)	0.08049	108.97	1.95e-12

TABLE 16 – ANOVA: REACTION TIMES VS. BASELINE

Even when reaction times are the only human factor included in the simulation, their influence is significant in all severities of safety-critical interactions. However, when reaction times are included in isolation, outliers in the rate of collisions can be produced at high speeds that violate the normality assumption of ANOVA. This is an artifact of assigning individual drivers' reaction times from the derived distribution. If the number of outliers generated becomes significant (without the other human factors present to drive up the mean and mask the effect) the results can be skewed. This occurred for collisions when the posted speed limit of the simulation reached 70 kph. To prevent this from invalidating the ANOVA, simulations at this speed were also excluded in the analysis that excluded low speeds for the reasons discussed in the introduction to this *Human Factors Analysis* section.

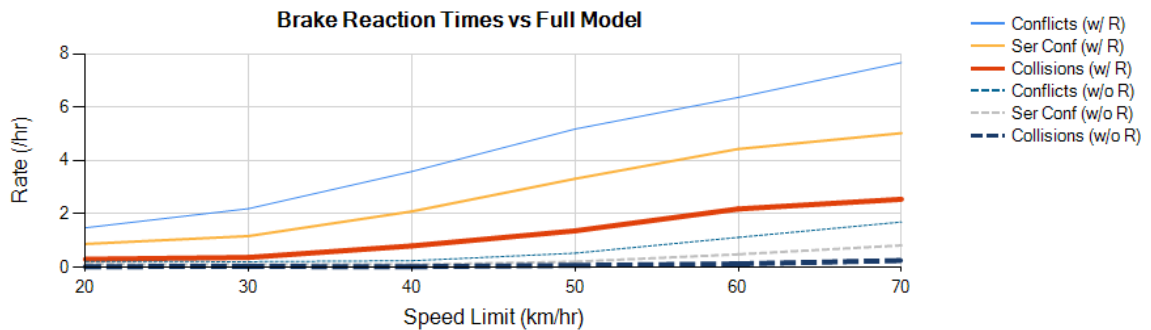


FIGURE 21 – SIMULATION RESULTS: REACTION TIMES VS. FULL MODEL

	Value	p-value	Adjusted R ²
β_0 Conflicts	-0.67105	4.84e-05	0.7818
β_1 Conflicts	0.02977	5.66e-11	
β_0 Serious Conflicts	-0.30990	0.000542	0.6998
β_1 Serious Conflicts	0.01366	5.17e-09	
β_0 Collisions	-0.104648	0.00612	0.5294
β_1 Collisions	0.004251	3.15e-06	

TABLE 17 – LINEAR REGRESSION: FULL MODEL EXCLUDING REACTION TIMES

	S-W p-value	F value	p-value
Conflicts	0.538	4495.4	<2e-16
Serious Conflicts	0.3557	5878.5	<2e-16
Collisions	0.01265	1868.7	<2e-16
Collisions (SL \geq 30kph)	0.3171	1370.1	<2e-16

TABLE 18 – ANOVA: REACTION TIMES VS. FULL MODEL

The decrease in conflicts, serious conflicts, and crashes from excluding finite reaction times from the otherwise-complete model is significant; and this effect is readily evident at all speeds.

DISTRACTION

There are nearly infinite objects and ideas competing for a driver's attention. Despite the vigilance necessary to remain aware of the ever-changing road around them, drivers will often divert their attention to secondary tasks. This behavior was modeled using an alternating renewal process. The durations of distracted periods are generated by Equation [34], and inter-distraction intervals are generated using Equation [38].

During distracted periods, drivers are not aware of actions by other road users. This means they do not yield to pedestrians or adjust the measurements used in their car-following behavior (z in Equation [10] and a_{CAH} in Equation [11]) from pre-distraction values. For the baseline, this distracted state is never entered.

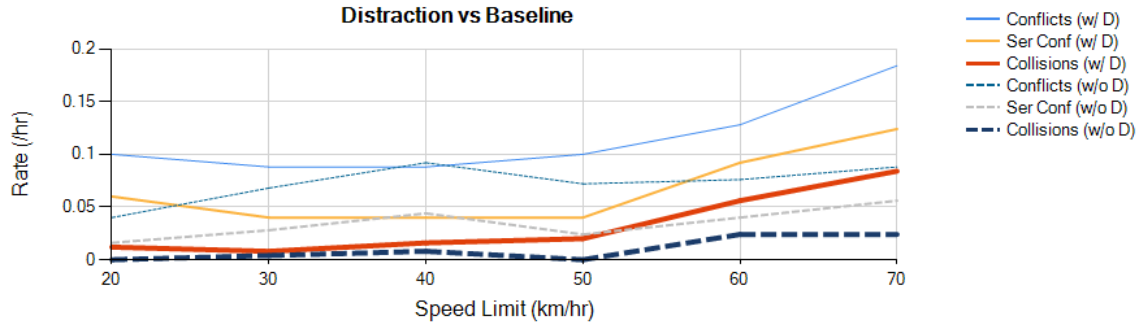


FIGURE 22 – SIMULATION RESULTS: DISTRACTION VS. BASELINE

	Value	p-value	Adjusted R ²
β_0 Conflicts	0.043695	0.04573	0.296
β_1 Conflicts	0.001577	0.00111	
β_0 Serious Conflicts	0.00480	0.81039	0.2547
β_1 Serious Conflicts	0.00136	0.00262	
β_0 Collisions	-0.032648	0.00223	0.6368
β_1 Collisions	0.001451	7.74e-08	

TABLE 19 – LINEAR REGRESSION: DISTRACTION ONLY

The effects of distraction, in isolation, are barely distinguishable from noise.

While the relationship to speed is significantly linear, the derived linear model is insufficient to describe the variation in safety-critical interactions.

	S-W p-value	F value	p-value
Conflicts	0.4486	19.600	5.49e-05
Serious Conflicts	0.5067	19.991	4.75e-05
Collisions	0.1076	56.390	1.22e-09

TABLE 20 – ANOVA: DISTRACTION VS. BASELINE

Using a driver model that includes distraction alone produces significantly different results than without; however, the resulting increases are small.

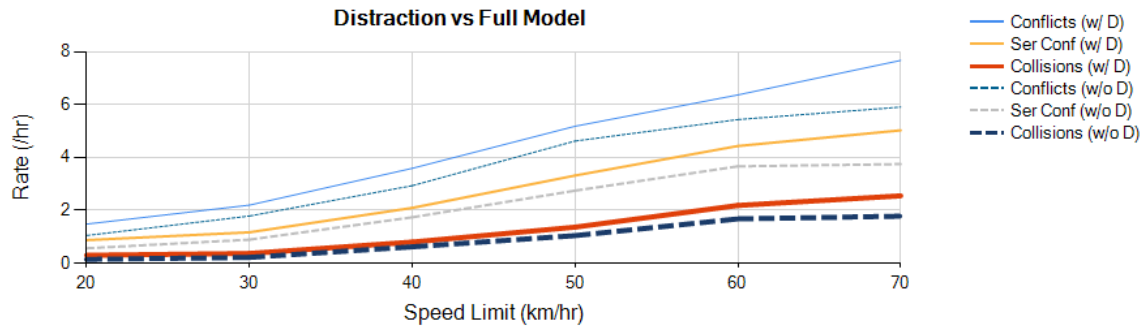


FIGURE 23 – SIMULATION RESULTS: DISTRACTION VS. FULL MODEL

	Value	p-value	Adjusted R ²
β_0 Conflicts	-1.1258	3.47e-05	0.9442
β_1 Conflicts	0.1055	< 2e-16	
β_0 Serious Conflicts	-1.02099	1.24e-06	0.9376
β_1 Serious Conflicts	0.07213	< 2e-16	
β_0 Collisions	-0.75383	6.24e-09	0.9284
β_1 Collisions	0.03706	< 2e-16	

TABLE 21 – LINEAR REGRESSION: FULL MODEL EXCLUDING DISTRACTION

	S-W p-value	F value	p-value
Conflicts	0.992	92.953	8.26e-13
Serious Conflicts	0.9965	134.982	1.49e-15
Collisions	0.4473	105.915	9.80e-14

TABLE 22 – ANOVA: DISTRACTION VS. FULL MODEL

The linear models for safety-critical interactions are significantly different between the full model and the model excluding distraction. To explore the magnitude of this difference, the predictions of the two models were compared. In each of the following tables, the last row shows the relative underestimation between the linear model from excluding distraction. For example, for 20kph, the full model predicts 1.196 conflicts per hour. Excluding distraction brings this down to 0.984 conflicts per hour – a difference of 0.212 conflicts per hour, or approximately 17.7% of the original 1.196.

Speed (kph)	20	30	40	50	60	70
Full Model	1.196	2.483	3.77	5.057	6.344	7.631
Excluding Distraction	0.984	2.039	3.094	4.149	5.204	6.259
Relative Difference	17.7%	17.9%	17.9%	18%	18%	18%

TABLE 23 – PREDICTED CONFLICTS FROM LINEAR MODELS

Speed (kph)	20	30	40	50	60	70
Full Model	0.548	1.456	2.363	3.271	4.179	5.087
Excluding Distraction	0.422	1.143	1.864	2.586	3.307	4.028
Relative Difference	23.0%	21.5%	21.1%	21.0%	20.9%	20.8%

TABLE 24 – PREDICTED SERIOUS CONFLICTS FROM LINEAR MODELS

Speed (kph)	20	30	40	50	60	70
Full Model	0.028	0.521	1.014	1.507	2.0	2.493
Excluding Distraction	0 ⁵¹	0.358	0.729	1.099	1.47	1.84
Relative Difference	100%	31.3%	28.2%	27.1%	26.5%	26.2%

TABLE 25 – PREDICTED COLLISIONS FROM LINEAR MODELS

These data show the underestimation of conflicts and collisions that result from excluding distraction from the driver model. The predicted collisions and serious conflicts are rather high, especially at low speeds. There are two significant reasons for this. First, the linear fits of these distributions are heavily influenced by the higher speeds. The relatively low R^2 values quantify this behavior, and a piece-wise linear fit would be more appropriate. More fundamentally, however, the pedestrian model used

⁵¹ The linear model actually predicts negative collisions (-0.013) at this speed.

in the simulation is overly simplistic. This limitation is discussed in more detail in the *Summary* subsection below.

SUMMARY

As demonstrated by the results of the Shapiro-Wilks tests, the results of the simulation (the generated rates of safety-critical interactions) are normally distributed for each combination of model parameters. Each of the human factors analyzed, with the exception of visual obstruction, has a small impact on simulated safety-critical interactions when included in isolation. However, using an incomplete combination of any three of the four significantly underestimates the resulting safety-critical interactions of all severities.

It should be noted that the produced rates of *serious conflicts* and *collisions* are artificially inflated by the simplicity of the models implemented. First, drivers are incapable of lateral evasive actions. The only method available to the simulated drivers for avoiding a collision is braking. Furthermore, once a simulated pedestrian has decided to cross, no more information about oncoming vehicles is processed. This behavior lends itself to a much greater ratio of severe interactions to simple conflicts. While the linear models produce ratios of as many as 300 pedestrian collisions per 1000 conflicts, real values are closer to .4 to 6.4 per 1000. [Davis, Sanderson, and Davuluri 2002]

OTHER AGENT RATES

Since the agent generation rates used in the initial analysis above are somewhat arbitrary, the responses of the rates of safety-critical interactions to changes in agent volumes were explored. To avoid the edge effects of congestion at low speeds (the linear behavior in the produced rates discussed above seems to break down below 30kph for many parameter sets), a second linear fit was performed using only the results for speeds greater than or equal to 30kph. This analysis was performed for the original rates of pedestrian and vehicle generation, as well as a few other illustrative combinations. Those results are presented here. Each features the full model with all four human factors.

	Value	p-value	Adjusted R ²
β_0 Conflicts	-1.8608	5.49e-09	0.9425
β_1 Conflicts	0.1372	< 2e-16	
β_0 Serious Conflicts	-1.8196	7.66e-11	0.9776
β_1 Serious Conflicts	0.1005	< 2e-16	
β_0 Collisions	-1.41920	6.24e-11	0.9598
β_1 Collisions	0.05744	< 2e-16	

TABLE 26 – LINEAR REGRESSION: FULL MODEL – PPH=400, VPH=800 (SPEED ≥ 30KPH)

	Value	p-value	Adjusted R ²
β_0 Conflicts	-1.566	1.59e-06	0.9561
β_1 Conflicts	0.1078	< 2e-16	
β_0 Serious Conflicts	-1.3064	1.71e-08	0.9625
β_1 Serious Conflicts	0.07408	< 2e-16	
β_0 Collisions	-0.9052	7.31e-10	0.9542
β_1 Collisions	0.03892	< 2e-16	

TABLE 27 – LINEAR REGRESSION: FULL MODEL – PPH=300, VPH=800 (SPEED ≥ 30KPH)

	Value	p-value	Adjusted R ²
β_0 Conflicts	-0.9076	4.02e-07	0.9698
β_1 Conflicts	0.0694	< 2e-16	
β_0 Serious Conflicts	-0.80640	1.06e-06	0.9461
β_1 Serious Conflicts	0.04856	2.68e-16	
β_0 Collisions	-0.66960	3.31e-10	0.9524
β_1 Collisions	0.02704	< 2e-16	

TABLE 28 – LINEAR REGRESSION: FULL MODEL – PPH=200, VPH=800 (SPEED ≥ 30KPH)

	Value	p-value	Adjusted R ²
β_0 Conflicts	-0.79480	0.000203	0.9623
β_1 Conflicts	0.08588	< 2e-16	
β_0 Serious Conflicts	-0.9120	1.33e-06	0.9565
β_1 Serious Conflicts	0.0624	< 2e-16	
β_0 Collisions	-0.7604	2.73e-09	0.9497
β_1 Collisions	0.0334	< 2e-16	

TABLE 29 – LINEAR REGRESSION: FULL MODEL – PPH=400, VPH=600 (SPEED ≥ 30KPH)

	Value	p-value	Adjusted R ²
β_0 Conflicts	0.22040	0.00941	0.9565
β_1 Conflicts	0.03444	< 2e-16	
β_0 Serious Conflicts	-0.15200	0.0311	0.949
β_1 Serious Conflicts	0.02696	<2e-16	
β_0 Collisions	-0.34120	0.000118	0.8486
β_1 Collisions	0.01652	4.05e-11	

TABLE 30 – LINEAR REGRESSION: FULL MODEL – PPH=400, VPH=400 (SPEED ≥ 30KPH)

	Value	p-value	Adjusted R ²
β_0 Conflicts	0.24840	2.23e-05	0.5612
β_1 Conflicts	0.00508	9.91e-06	
β_0 Serious Conflicts	0.15520	0.000223	0.4913
β_1 Serious Conflicts	0.00336	5.73e-05	
β_0 Collisions	0.0444	0.0216	0.6208
β_1 Collisions	0.0022	1.78e-06	

TABLE 31 – LINEAR REGRESSION: FULL MODEL – PPH=400, VPH=200 (SPEED ≥ 30KPH)

These data demonstrate that the results for other flow rates are also significantly linear ($p < .05$). Further, the R^2 values suggest that the linear fit is sufficient to describe the behavior in all cases except that in Table 31]. At low vehicular volumes, while there is a significant linear aspect with respect to changes in PSL, the stochasticity of the model introduces variations on the same order of magnitude as the linear trend.

Presented below are comparisons of the predicted rates of safety-critical interactions at each speed. The top row in each of the following tables shows the linear predictions for the rate combination used in the initial analysis (excluding the values for PSL of 20kph, since the linear behavior often broke down at these low speeds). The remaining lines show the predicted rates of safety-critical interactions for the other examined agent generation rates. Also presented is the proportion of the initial fitted rates at that speed. For example, in Table 32 the results show approximately 2.26 conflicts per hour for 400 PPH and 800 VPH traveling at 30kph. When the pedestrian rate is halved to 200 PPH, the linear fit predicts approximately 1.17 conflicts per hour for the same speed limit, or 52.1% of the original.

Agent Rates	30	40	50	60	70	
PPH = 400 VPH = 800	2.2552	3.6272	4.9992	6.3712	7.7432	100%
PPH = 300 VPH = 800	1.668	2.746	3.824	4.902	5.98	
	74	75.7	76.5	76.9	77.2	%
PPH = 200 VPH = 800	1.1744	1.8684	2.5624	3.2564	3.9504	
	52.1	51.1	51.3	51.1	51	%
PPH = 400 VPH = 600	1.7816	2.6404	3.4992	4.358	5.2168	
	79	72.8	70	68.4	67.4	%
PPH = 400 VPH = 400	1.2536	1.598	1.9424	2.2868	2.6312	
	55.6	44.1	38.9	35.9	34	%
PPH = 400 VPH = 200	0.4008	0.4516	0.5024	0.5532	0.604	
	17.8	12.5	10	8.7	7.8	%

TABLE 32 – PREDICTED CONFLICTS FROM LINEAR MODELS AT VARIOUS AGENT RATES

Agent Rates	30	40	50	60	70	
PPH = 400 VPH = 800	1.1954	2.2004	3.2054	4.2104	5.2154	100%
PPH = 300 VPH = 800	0.916	1.6568	2.3976	3.1384	3.8792	
	76.6	75.3	74.8	74.5	74.4	%
PPH = 200 VPH = 800	0.6504	1.136	1.6216	2.1072	2.5928	
	54.4	51.6	50.6	50	49.7	%
PPH = 400 VPH = 600	0.96	1.584	2.208	2.832	3.456	
	80.3	72	68.9	67.3	66.3	%
PPH = 400 VPH = 400	0.6568	0.9264	1.196	1.4656	1.7352	
	54.9	42.1	37.3	34.8	33.3	%
PPH = 400 VPH = 200	0.256	0.2896	0.3232	0.3568	0.3904	
	21.4	13.2	10.1	8.5	7.5	%

TABLE 33 – PREDICTED SERIOUS CONFLICTS FROM LINEAR MODELS AT VARIOUS AGENT RATES

Agent Rates	30	40	50	60	70	
PPH = 400 VPH = 800	0.304	0.8784	1.4528	2.0272	2.6016	100%
PPH = 300 VPH = 800	0.2624	0.6516	1.0408	1.43	1.8192	
	86.3	74.2	71.6	70.5	69.9	%
PPH = 200 VPH = 800	0.1416	0.412	0.6824	0.9528	1.2232	
	46.6	46.9	47	47	47	%
PPH = 400 VPH = 600	0.2416	0.5756	0.9096	1.2436	1.5776	
	79.5	65.5	62.6	61.3	60.6	%
PPH = 400 VPH = 400	0.1544	0.3196	0.4848	0.65	0.8152	
	50.8	36.3	33.4	32.1	31.3	%
PPH = 400 VPH = 200	0.1104	0.1324	0.1544	0.1764	0.1984	
	36.3	15.1	10.6	8.7	7.6	%

TABLE 34 – PREDICTED COLLISIONS FROM LINEAR MODELS AT VARIOUS AGENT RATES

These results show an approximately linear response in all severities of safety-critical interactions at all speeds to changes in pedestrian volumes, q_p (given a constant vehicular flow rate, q_v). This implies that the results can simply be scaled linearly for changes in pedestrian volumes. (This linear trend is considerably weaker for collision rates, but – as discussed previously – these rates are of limited applicability due to their being inflated by the limitations of the simulation model.) Defining Y as the rate of an individual severity of safety-critical interaction, and Y_0 as the rate of those interactions at 400 PPH, we have

$$Y = \frac{q_p}{400} * Y_0 \quad [54]$$

Changes in vehicular volumes, on the other hand, do not produce a uniform, linear response curve in the rates of safety-critical interactions across all speeds. However, for a given PSL, changes in vehicular volumes do have an approximately linear effect on rates of safety-critical interactions. This is shown in the following figures.

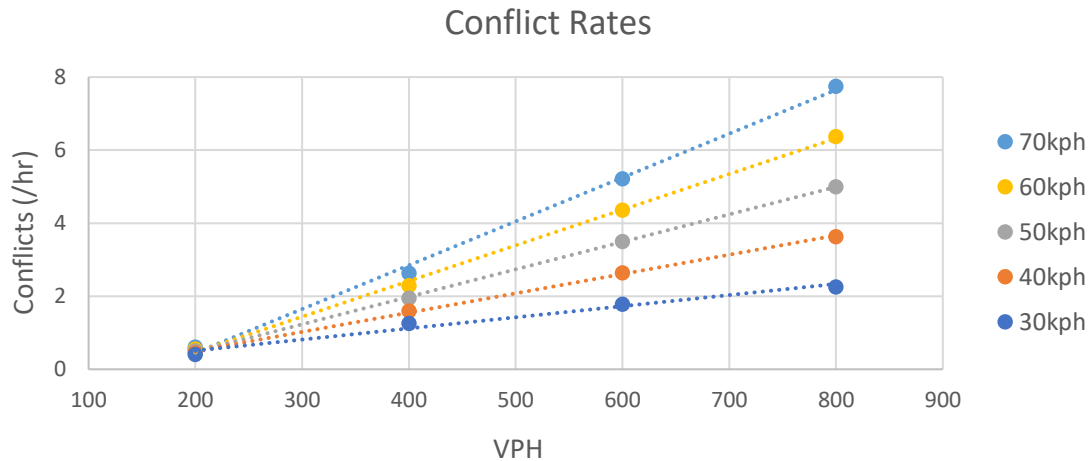


FIGURE 24 – LINEAR RESPONSE OF CONFLICT RATES TO CHANGES IN VEHICLE VOLUMES

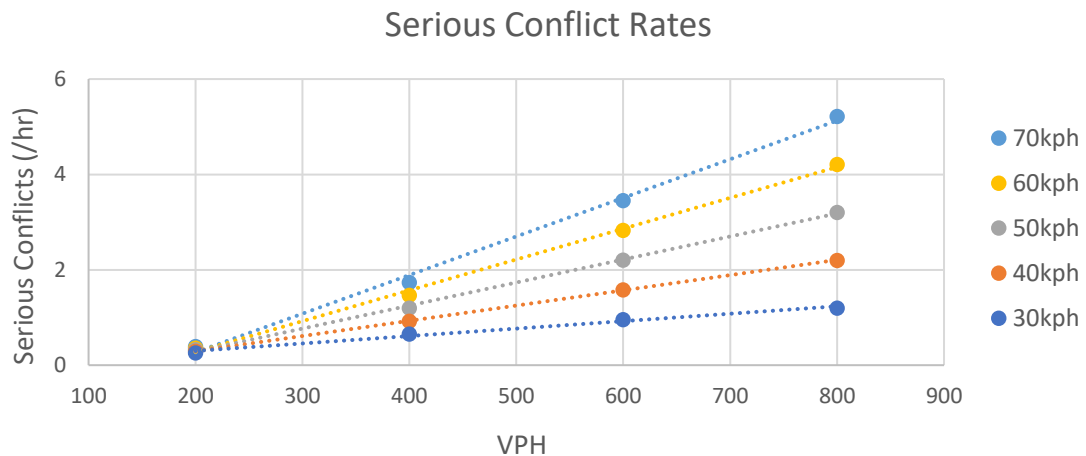


FIGURE 25 – LINEAR RESPONSE OF SERIOUS CONFLICT RATES TO CHANGES IN VEHICLE VOLUMES

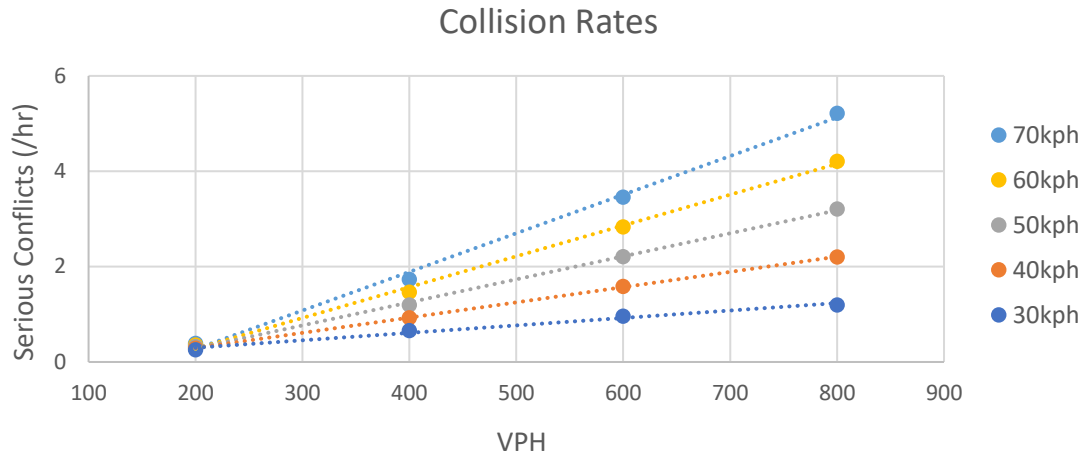


FIGURE 26 – LINEAR RESPONSE OF COLLISION RATES TO CHANGES IN VEHICLE VOLUMES

For a given PSL, the rates of safety-critical interactions with changing vehicular volumes, q_v , are substantially linear. Defining the rates of the various severities of these interactions, Y , this relationship can be represented by the following equation:

$$Y = \beta_0 + \beta_1 \cdot q_v \quad [55]$$

Here, Y is in units of *safety-critical interactions per hour*, and q_v is in *vehicles per hour*.

While there may be a small quadratic component to the curves (noticeable in the figures above at high speeds), the linear component dominates. The results of these linear fits are given in Table 35

	30kph	40kph	50kph	60kph	70kph
β_0 Conflicts	-0.1	-0.563	-1.026	-1.489	-1.952
β_1 Conflicts	0.003046	0.005285	0.007524	0.009763	0.012
Adjusted R^2 Conflicts	0.9686	0.9983	0.9997	0.9978	0.9958
β_0 Serious Conflicts	-0.0133	-0.3474	-0.6815	-1.0156	-1.3497
β_1 Serious Conflicts	0.001561	0.003195	0.004829	0.006464	0.008098
Adjusted R^2 Serious Conflicts	0.9791	0.9998	0.9984	0.9962	0.9945
β_0 Collisions	0.0356	-0.147	-0.3296	-0.5122	-0.6948
β_1 Collisions	0.000334	0.001247	0.00216	0.003073	0.003986
Adjusted R^2 Collisions	0.9791	0.9839	0.982	0.981	0.9804

TABLE 35 – LINEAR FITS OF RESPONSE CURVES TO CHANGES IN VEHICLE VOLUMES

The coefficients of these linear fits (β_0 and β_1) are, themselves, linearly related to PSL. These linear fits are exact ($R^2 = 1$), since the data that led to the coefficients in Table 35 were produced from linear fits to the raw simulation output.

$$\begin{aligned}\phi_{0,0} + \phi_{1,0} \cdot PSL &= \beta_0 \\ \phi_{0,1} + \phi_{1,1} \cdot PSL &= \beta_1\end{aligned}\tag{56}$$

	$\phi_{0,0}$	$\phi_{1,0}$	$\phi_{0,1}$	$\phi_{1,1}$
Conflicts	1.289	-0.0463	-0.0036694	0.0002239
Serious Conflicts	0.989	-0.03341	-0.0033421	0.0001634
Collisions	0.5834	-0.01826	-0.0024050	0.0000913

TABLE 36 – PARAMETERS DESCRIBING HOW LINEAR FITS TO VPH CHANGE WITH PSL

Combining Equations [55] and [56], we have an equation for predicting rates of safety-critical interactions at the initial rate of 400 pedestrians per hour:

$$Y = \phi_{0,0} + \phi_{1,0} \cdot PSL + (\phi_{0,1} + \phi_{1,1} \cdot PSL)q_v$$

Combining this result with that for changes in pedestrian flows, Equation [54], we have the final predictive equation for safety-critical interactions:

$$Y = \frac{q_p}{400} [\phi_{0,0} + \phi_{1,0} \cdot PSL + (\phi_{0,1} + \phi_{1,1} \cdot PSL)q_v] \quad [57]$$

To use Equation [57], q_p is in units of PPH; PSL is in kph; and q_v is in VPH. The results are returned in safety-critical interactions per hour. For most combinations of parameters, Equation [57] is accurate to within 10% of the simulation results.

PERCEPTION DISTANCE

Arguably the most arbitrary parameter chosen in the development of the simulation model is that of the maximum distance at which a driver can perceive a pedestrian intending to cross or crossing. The value used in the above analyses for this d_{perc} , 100m, is essentially an upper-bound since it is based on driver observation of safety workers wearing full-body fluorescent safety attire and standing next to conspicuous equipment. [Helman and Palmer 2010] The literature on the topic is sparse, and this was the only study found that hadn't explicitly instructed drivers to look out for pedestrians. The simulation results for lower values of this perception distance are presented below, along with the original results ($d_{perc} = 100m$) for reference.

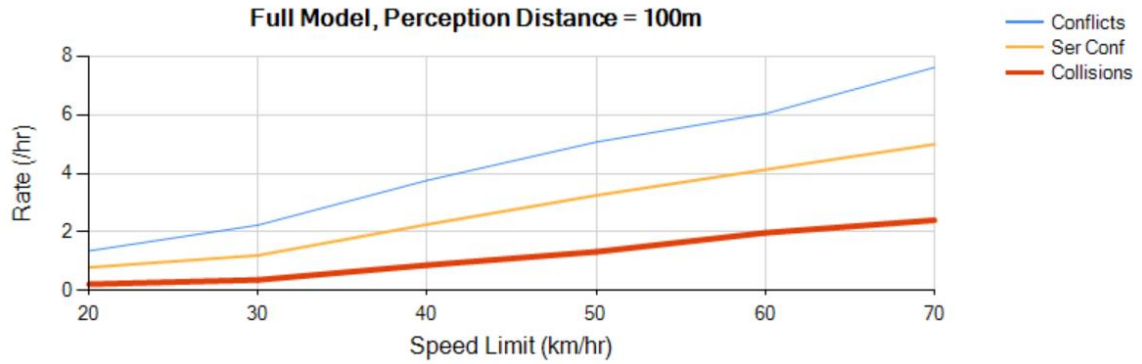


FIGURE 27A – MODEL RESULTS, PERCEPTION DISTANCE = 100M

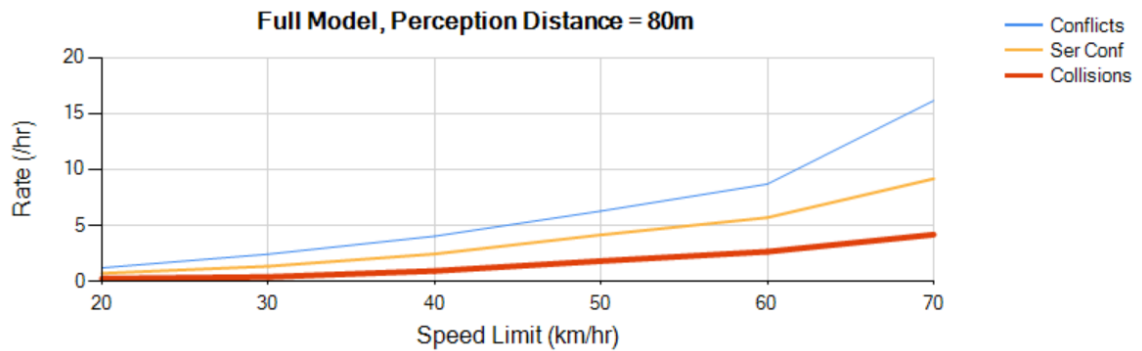


FIGURE 27B – MODEL RESULTS, PERCEPTION DISTANCE = 80M

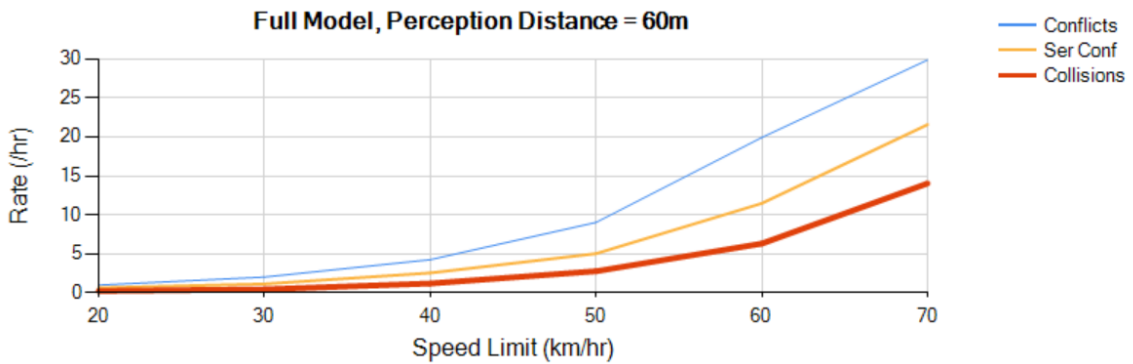


FIGURE 27C – MODEL RESULTS, PERCEPTION DISTANCE = 60M

From these data, it is obvious that the linear behavior for the rates of safety-critical interactions with changes in speed breaks down for lower perception distances. When the value of this parameter is low enough, an approximately quadratic behavior seems to dominate the results. This pattern arises when even attentive drivers simply do not have enough space in which to stop. To demonstrate this, heatmaps were

generated to visualize the locations of both driver decisions to yield and the incidence of emergency braking (defined here as yield decelerations equal to the minimum allowed: -7.4 m/s). The full human factors model is used in the simulations run to generate these graphs, so the decision points plotted have taken distraction and BRT into account. The x-axis shows meters from the end of the road. The center of the crosswalk is located at 750m.

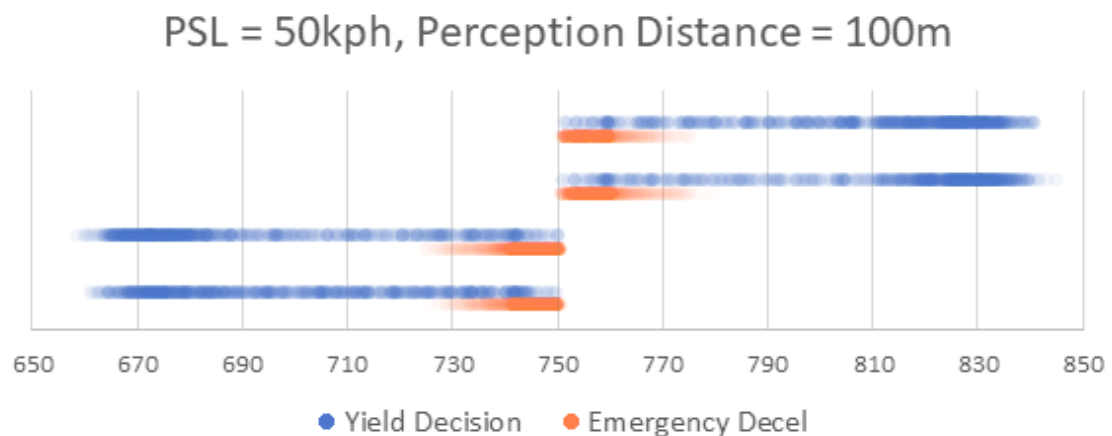


FIGURE 28A – HEATMAP: PSL = 50KPH, PERCEPTION DISTANCE = 100M

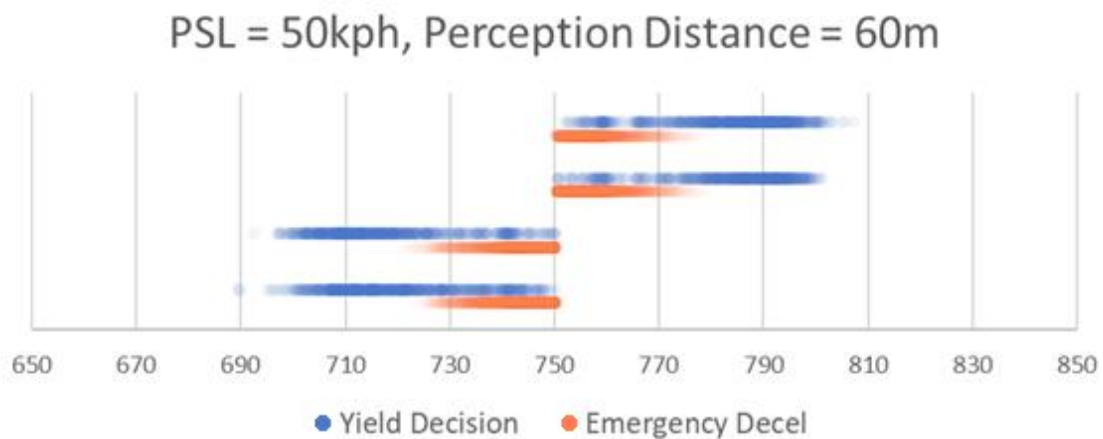


FIGURE 28B – HEATMAP: PSL = 50KPH, PERCEPTION DISTANCE = 60M

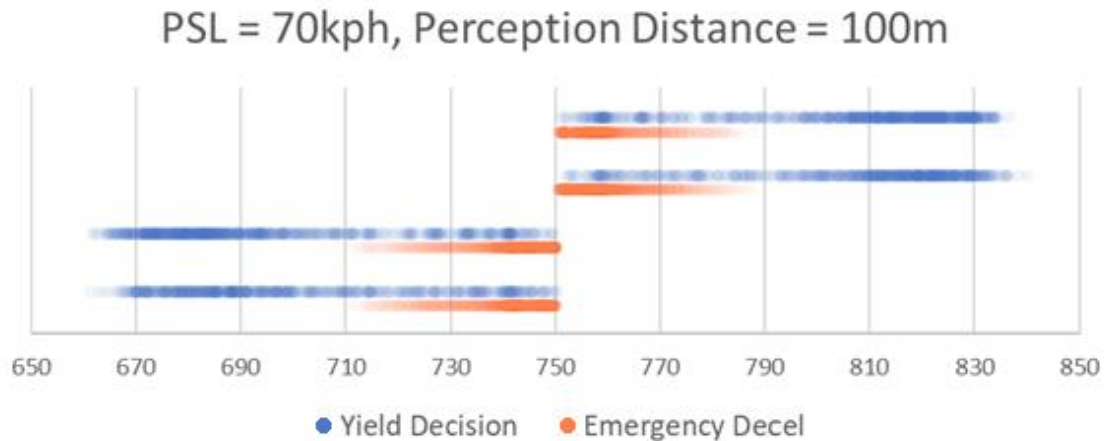


FIGURE 29A – HEATMAP: PSL = 70KPH, PERCEPTION DISTANCE = 100M

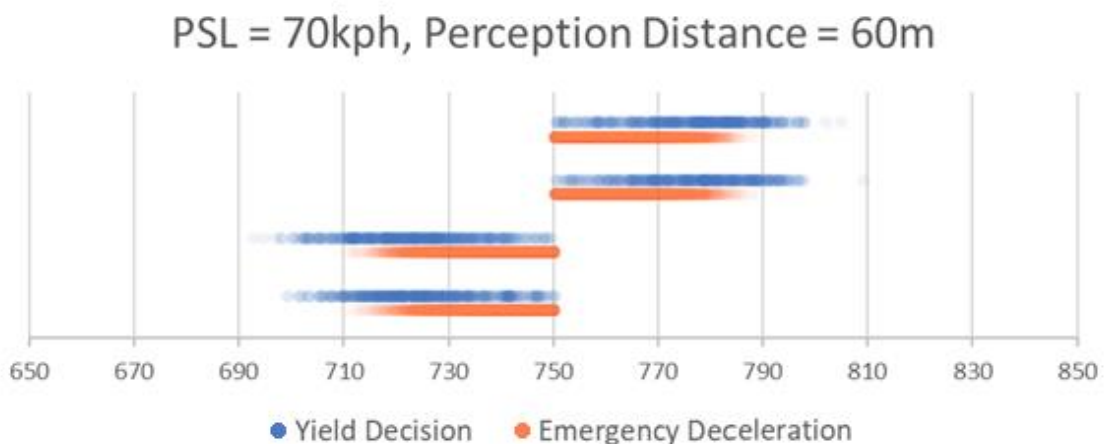


FIGURE 29B – HEATMAP: PSL = 70KPH, PERCEPTION DISTANCE = 60M

When the perception distance is 100m (Figure 28a and Figure 29a), the vast majority of driver decisions to yield happen before emergency braking is required, even at high speeds. When this value is lowered to 60m, however, (Figures Figure 28b and Figure 29b) higher speeds mean a higher percentage of drivers recognize too late that a yield is necessary, leading to higher rates of conflicts and collisions.

Counter-intuitively, at still higher speeds the rates of safety-critical interactions reach maxima and begin to decrease. For a given flow rate, increasing speeds have two

effects. First, a higher percentage of vehicles are unable to yield in time to avoid a conflict. On the other hand, vehicular densities steadily decrease. When the former effect reaches a maximum of all approaching vehicles being involved in safety-critical encounters, the latter takes over and conflicts steadily decrease. This can be seen more clearly by running the simulation with unrealistically high speeds, as shown in Figure 30.

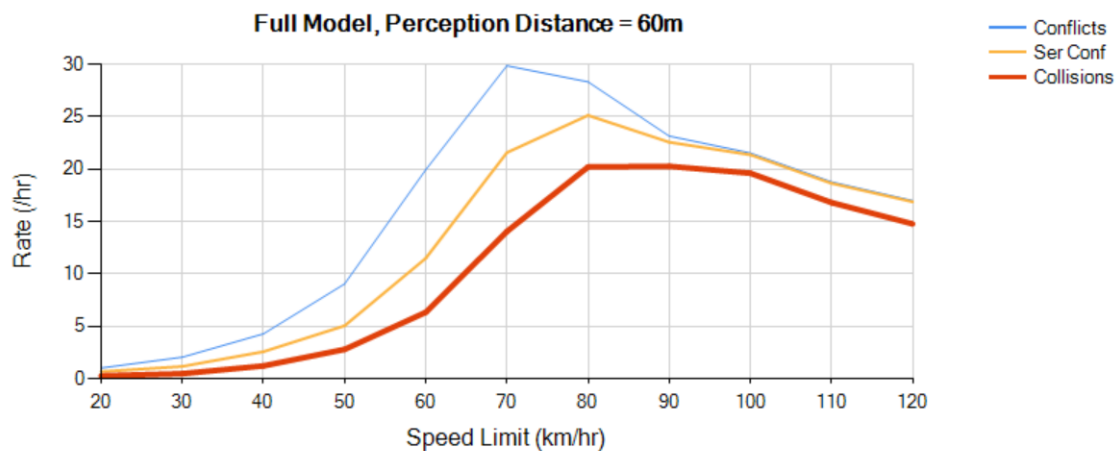


FIGURE 30 – SIMULATION RESULTS EXTENDED TO HIGH PSL, PERCEPTION DISTANCE = 60M

Doing the same for the original perception distance value of 100m shows that, while higher speeds are required, the same phenomenon occurs:

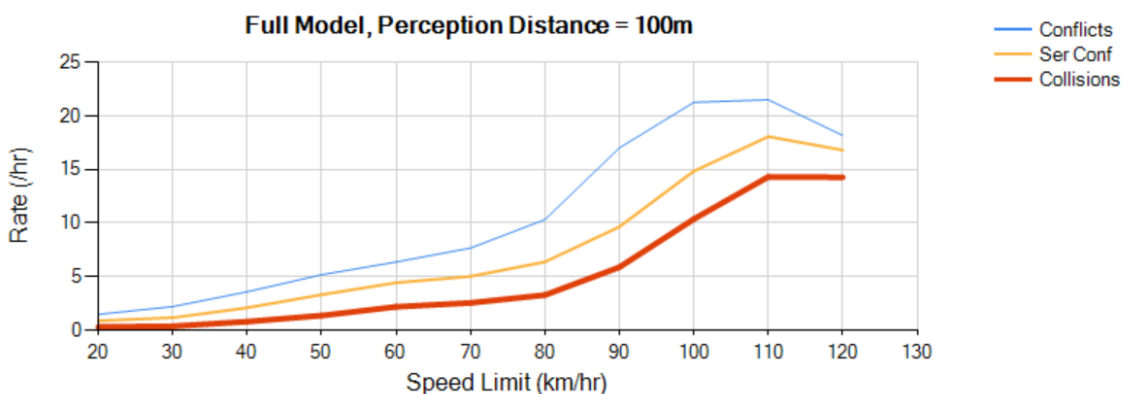


FIGURE 31 – SIMULATION RESULTS EXTENDED TO HIGH PSL, PERCEPTION DISTANCE = 100M

If a model such as this is to be used for predicting pedestrian safety, the most important parameter for calibration is this value of perception distance, d_{perc} . There is a need for both more empirical evidence on this value and further exploration of the patterns hinted at above.

CONCLUSIONS

The work presented herein developed a data-driven approach to including distraction in a microsimulation driver model. Microscopic simulation of traffic is an essential tool for studying road safety. However, no currently-available simulation platforms explicitly include driver distraction. An agent-based model of a mid-block pedestrian crossing was constructed to analyze the effects of excluding distraction or other human factors from the models underlying microscopic traffic safety simulations.

Four human factors were analyzed: distraction, finite reaction times, visual obstruction, and errors in estimation of distance and relative velocity. The safety performance of driver models incorporating various combinations of these factors were analyzed. This safety performance was quantified using traffic conflict analysis, based on minimum time-to-collision.

Each human factor was included in the model in isolation, and the resulting conflict rates were compared to a baseline of no human factors. Next, each was excluded in isolation, and these results were compared to the full driver model that incorporated all four. Each of these combinations was run for a range of speeds.

None of the human factors in isolation generated more than a single conflict per hour, even at 70kph. However, exclusion from the full model was statistically significant for all four. While previous research comparing simulations to empirical data is scant, it has been shown that commercially available traffic simulation software platforms tend

to under-predict vehicle-pedestrian conflicts, as well as vehicle-vehicle conflicts away from intersections. The possibility of this being partially remedied through the explicit inclusion of driver distraction should be explored further in future research.

These analyses of including or excluding individual human factors were all performed on simulation results featuring pedestrian and vehicle arrival rates of 400 PPH and 800 VPH, respectively. (These are combined rates for both travel directions, e.g. 400 PPH represents 200 pedestrians per hour from each side of the road.) The rates of safety-critical interactions were approximately linear for a given set of agent arrival rates. This linear relationship was used to derive Equation [57] for predicting the rates of safety-critical interactions for reasonable agent arrival rates and any posted speed limit between 30kph and 70kph. For most rate combinations, this equation is accurate to within 10% of the safety-critical interactions generated by the simulation. It should be noted that, due to the simplicity of the pedestrian model and the driver yielding model utilized, these predictions break down for severe congestion.

While the conflict rates produced by the simulation are reasonable, the ratios of these to the more serious interactions are unrealistic. A substantial portion of this is the result of the simulated model not providing any means of evasive action beyond driver braking. Further, incorporation of a pedestrian model that allows pedestrians to change speed or return to the curb as the situation unfolds would be necessary for the rates of serious conflicts or collisions to be more realistic.

The foundation of Equation [57] rests in the observation that, for a given combination of vehicle and pedestrian arrival rates, the rates of safety-critical interactions are linearly related to the posted speed limit. However, this relationship breaks down when the distance at which drivers are able to perceive waiting or crossing pedestrians decreases significantly below 100m. Further research is needed to empirically quantify this parameter; and any applications of the presented model should first attempt to calibrate this value.

It should also be noted that the eye-glance dataset used for development of the distribution of distraction events is based on less than three hours of combined recordings. A more substantial set of data that explores how this distribution changes based on external factors (scene complexity, time of day, etc.) would make the derived model more robust. Fortunately, these eye-glance behavior data are now typical in traffic studies, and the form of the data employed requires limited manual coding.

The limitations of human drivers have received increased attention over recent years with the introduction of driver assistance systems that are approaching full automation. Automated vehicles have the potential of removing driver distraction as a factor in transportation safety. Until that day comes, however, it is important that our tools for evaluating and predicting road safety incorporate this fundamentally human factor.

REFERENCES

- AASHTO – American Association of State Highway and Transportation Officials. (2001) A Policy on Geometric Design of Highways and Streets
- Absil N. (2008) Driver Behaviour Model for the Multi-Agent Real-time Simulation. Master's thesis. Department of Electrical Engineering, Mathematics, and Computer Science. Delft University of Technology. The Netherlands.
- Akçelik R, Seidel E, and Hill A. (2001) An Investigation of Pedestrian Movement Characteristics at Mid-Block Signalised Crossings. (Technical Report) Akçelik & Associates Pty Ltd. Balwyn, Australia
- Allen B, Shin BT, and Cooper P. (1978) Analysis of Traffic Conflicts and Collisions. TRR: Journal of the TRB 667:67-72
- Almodfer R, et al. (2016) Quantitative Analysis of Lane-Based Pedestrian-Vehicle Conflict at a Non-Signalized Marked Crosswalk. Transportation Research Part F: Traffic Psychology and Behaviour 42:468-478
- Ambros J, Turek R, and Paukrt J. (2014) Road Safety Evaluation Using Traffic Conflicts: Pilot Comparison of Micro-Simulation and Observation. Presented at: International Conference on Traffic and Transport Engineering. Belgrade, Serbia.
- Amundsen F and Hydén C. (Eds.) (1977) Proceedings from the First Workshop on Traffic Conflicts. 26-27 September 1977, Oslo, Norway. (in Norwegian, cited in van der Horst 1990)
- Angell, et al. (2006). Driver workload metrics task 2 final report and appendices (DOT HS 810 635). Final research report. Washington, DC: National Highway Transportation Safety Administration, U.S. Department of Transportation.

Anvari B, Bell M, Sivakumar A, and Ochieng W. (2015) Modelling Shared Space Users via Rule-Based Social Force Model. *Transportation Research Part C: Emerging Technologies* 51:83-103

Archer J and Kosonen I. (2000) The Potential of Micro-Simulation Modelling in Relation to Traffic Safety Assessment. *Proceedings of the European Simulation Symposium*. Hamburg, Germany

Ariza A. (2011) Validation of Road Safety Surrogate Measures as a Predictor of Crash Frequency Rates on a Large-Scale Microsimulation Network. Master's Thesis, University of Toronto School of Civil Engineering

Austrroads. (2011) The Nature of Errors Made by Drivers. Austrroads Ltd. Sydney, Australia. AP-R378/11

Axelrod R. (2005) Advancing the Art of Simulation in the Social Sciences, in *Handbook of Research on Nature Inspired Computing for Economy and Management* (Rennard JP ed.) Idea Group. Hersey, PA

Balmer M, Nagel K, and Raney B. (2004) Large-Scale Multi-Agent Simulations for Transportation Applications. *Intelligent Transportation Systems* 8:205-2011

Bärgman J, et al. (2015) How does glance behavior influence crash and injury risk? A 'what-if' counterfactual simulation using crashes and near-crashes from SHRP2. *Transportation Research Part F: Traffic Psychology and Behaviour* 35:152-169

Batty M. (2001) Agent-Based Pedestrian Modeling. *Environment and Planning B: Planning and Design* 28:321-326

Bazzan A and Klügl F. (2013) A Review on Agent-Based Technology for Traffic and Transportation. *The Knowledge Engineering Review* 29(3):375-403

- Bella F and Silvestri M. (2016) Driver's Braking Behavior Approaching Pedestrian Crossings: A Parametric Duration Model of the Speed Reduction Times. *Journal of Advanced Transportation* 50:630-646
- Bellman R and Cooke K. (1963) Differential-Difference Equations. The RAND Corporation. US Airforce Project Rand: R-374-PR
- Benenson R, Mathias M, Timofte R, and van Gool L. (2012) Pedestrian Detection at 100 Frames per Second. *IEEE Conference on Computer Vision and Pattern Recognition*
- Berger W. (1975) Urban Pedestrian Accident Countermeasures Experimental Evaluation – Vol. 1: Behavioral Evaluation Studies. National Highway Traffic Safety Administration. DOT-HS-190-2-480
- Berry D and Belmont D. (1951) Distribution of Vehicle Speeds and Travel Times. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. pp:589-602
- Bertulis T and Dulaski D. (2014) Driver Approach Speed and its Impact on Driver Yielding to Pedestrian Behavior at Unsignalized Crosswalks. *TRR: Journal of the TRB* 2464:46-51
- Bevrani K and Chung E. (2012) A Safety Adapted Car Following Model for Traffic Safety Studies. In: Stanton N (Ed.) *Advances in Human Aspects of Road and Rail Transportation*. CRC Press. San Francisco. pp:550-559
- Bratman M. (1987) *Intention, Plans, and Practical Reason*. Harvard University Press. Cambridge, MA
- Brewer M, et al. (2006) Exploration of Pedestrian Gap-Acceptance Behavior at Selected Locations. *TRR: Journal of the TRB* 1982:132-140
- Brockfeld E, Kühne R, Skabardonis A, and Wagner P. (2003) Toward Benchmarking of Microscopic Traffic Flow Models. *TRR: Journal of the TRB* 1852:124-129

Brockfeld E, Kühne R, and Wagner P. (2004) Calibration and Validation of Microscopic Traffic Flow Models. TRR: Journal of the TRB 1876:62-70

Burstedde C, Klauck K, Schadschneider A, and Zittartz J. (2001) Simulation of Pedestrian Dynamics Using a Two-Dimensional Cellular Automaton. Physica A: Statistical Mechanics and its Applications 295(3):507-525

Caird JK, Willness C, Steel P, and Scialfa C. (2008) A Meta-Analysis of Cell Phone Use on Driver Performance. Accident Analysis & Prevention 40:1282-1293

Campbell B, Zegeer C, Huang H, and Cynecki M. (2004) A Review of Safety Research in the United States and Abroad: Final Report. Federal Highway Administration Report FHWA-RD-03-042

Campbell R and King LE. (1970) The Traffic Conflicts Technique Applied to Rural Intersections. Accident Analysis and Prevention 2:209-221

Chae K. (2005) Simulation of Pedestrian-Vehicle Interactions at Roundabouts. PhD Dissertation. North Carolina State University.

Chandler R, Herman R, and Montroll E. (1958) Traffic dynamics: studies in car following. Oper. Res. 6(2):165-184

Chang D. (2008) National Pedestrian Crash Report. National Highway Traffic Safety Administration: National Center for Statistics and Analysis. Washington, DC. DOT-HS-810-968

Chang MS, Messer C, and Santiago A. (1985) Timing Traffic Signal Change Intervals Based on Driver Behavior. TRR: Journal of the TRB 1027:20-30

Chen B and Cheng H. (2010) A Review of the Applications of Agent Technology in Traffic and Transportation Systems. IEEE Transactions on Intelligent Transportation Systems 11(2):485-497

- Cheng G, Wang Y, and Li D. (2013) Setting Conditions of Crosswalk Signal on Urban Road Sections in China. 2013 International Conference on Transportation. Xianning, China
- Ciuffo B, Punzo V, and Montanino M. (2014) Global sensitivity analysis techniques to simplify the calibration of traffic simulation models. Methodology and application to the IDM car-following model. IET Intelligent Transport Systems 8(5):479-489
- Cody D, Tan S, Caird J, Lees M, and Edwards C. (2008) The Naturalistic Driver Model: Development, Integration, and Verification of Lane Change Maneuver, Driver Emergency and Impairment Modules. California PATH Working Paper UCB-ITS-PWP-2008-8
- Cohen AC. (1965) Maximum Likelihood Estimation in the Weibull Distribution Based on Complete and on Censored Samples. Technometrics 7(4):579-588
- Cole B and Hughes P. (1984) A Field Trial of Attention and Search Conspicuity. Human Factors 26(3):299-313
- Collier N. (2001) RePast: An extensible framework for agent simulation. Natural Resources and Environmental Issues 8: Article 4
- Cox D. (1967) *Renewal Theory*. Chapman and Hall Ltd. London, UK
- Cremer M and Ludwig J. (1986) A fast simulation model for traffic flow on the basis of Boolean operations. Mathematics and Computers in Simulation 28(4):297-303
- Crompton D. (1979) Pedestrian Delay, Annoyance and Risk: Preliminary Results from a 2 Years Study. In Proceedings of the 7th Annual Planning and Transport Research and Computation Seminar. pp:275-299
- Daamen W. (2004) Modelling Passenger Flows in Public Transport Facilities (PhD thesis). Delft University of Technology: Faculty of Civil Engineering and Geosciences, Department Transport and Planning. Delft, Netherlands

- Dahl OJ and Nygaard K. (1968) SIMULA 67 common base language. Norwegian Computing Center
- Davis G, Sanderson K, and Davuluri S. (2002) Analysis of Pedestrian Safety Using Microsimulation and Driving Simulator. Minnesota Department of Transportation Report MN/RC – 2002-23
- Dingus T, et al. (2006) The 100-Car Naturalistic Driving Study: Phase II – Results of the 100-Car Field Experiment. National Highway Traffic Safety Administration. DOT-HS-810-593
- Distraction.gov (unknown) Official US Government Website for Distracted Driving: Facts and Statistics. Available: <https://www.distraction.gov/stats-research-laws/facts-and-statistics.html>. Accessed: 2017/03/25
- Dowling R, Skabardonis A, and Alexiadis V. (2004) Traffic Analysis Toolbox Volume III: Guidelines for Applying Traffic Microsimulation Software. US DoT Federal Highway Administration. FHWA-HRT-040-040
- Dumbaugh E, Li W, and Joh K. (2013) The Built Environment and the Incidence of Pedestrian and Cyclist Crashes. *Urban Design International* 18(3):217-228
- Elias W and Shiftan Y. (2014) Analyzing and Modeling Risk Exposure of Pedestrian Children to Involvement in Car Crashes. *Accident Analysis and Prevention* 62:397-405
- Epstein J and Axtell R. (1996) *Growing Artificial Societies*. The Brookings Institute. Washington, DC
- FARS – Fatality Analysis Reporting System. (online) <https://www-fars.nhtsa.dot.gov>. Queried by author 2017/03/25 (year = 2014)
- FARS. (online) <https://www-fars.nhtsa.dot.gov>. Queried by author 2017/03/24 (year = 2015; pedestrian position = 2; non-motorist location = 10)

FARS. (online) <https://www-fars.nhtsa.dot.gov>. Queried by author 2017/12/31 (year = 2016; person type = 5)

FHWA – Federal Highway Administration (2004) NGSIM Task E.1-1: Core Algorithms Assessment, Final Report. U.S. DOT. Washington D.C. [cited in Chae 2005]

FHWA (2009) Manual on Uniform Traffic Control Devices, Chapter 4C Warrant 4.

Available: <https://mutcd.fhwa.dot.gov/htm/2009/part4/part4c.htm>. Accessed

2018/4/15

Figliozi M and Tipagornwong C. (2016) Pedestrian Crosswalk Law: A Study of Traffic and Trajectory Factors that Affect Non-Compliance and Stopping Distance. Accident Analysis and Prevention 96:169-179

Finney DJ. (1941) On the Distribution of a Variate Whose Logarithm is Normally Distributed. Supplement to the Journal of the Royal Statistical Society 7(2):155-161

Fiorini P and Shiller Z. (1998) Motion Planning in Dynamic Environments Using Velocity Obstacles. International Journal of Robotics Research 17(7):760-772

Fitzpatrick K, Carlson P, Brewer M, and Woolridge M. (2001) Design Factors that Affect Driver Speed on Suburban Streets. TRR: Journal of the TRB 1751:18-25

Fitzpatrick K, Carlson P, Brewer M, Woolridge M, and Miaou SP. (2003) Design Speed, Operating Speed, and Posted Speed Practices. National Cooperative Highway Research Program. Report 504

Forbes T. (1957) Analysis of Near Accident Reports. Highway Research Board. HRB Bulletin 52:23-25

Fuller R. (2005) Towards a General Theory of Driver Behaviour. Accident Analysis and Prevention 37(3):461-472

GAO – Government Accountability Office. Pedestrians and Cyclists: Cities, States, and DOT are Implementing Actions to Improve Safety. GAO-16-66

Gardiner CW. (1983) Stochastic Methods: A Handbook for the Natural and Social Sciences, 2nd Ed. Springer

Gazis D, Herman R, and Rothery R. (1961) Nonlinear Follow-the-Leader Models of Traffic Flow. Operations Research 9(4):545-567

Getchell A. (2008) Agent-Based Modeling. University of CA, Davis.
<http://econ2.econ.iastate.edu/tesfatsi/AgentBasedModeling.AdamGetchell.phy250.Report.pdf>. Accessed 2017/03/04

Gettman D and Head L. (2003) Surrogate Safety Measures from Traffic Simulation Models, Final Report. Federal Highway Administration report FHWA-RD-03-050

Gettman D, Pu L, Sayed T, and Shelby S. (2008) Surrogate Safety Assessment Model and Validation: Final Report. US Department of Transportation: Federal Highway Administration. FHWA-HRT-08-051

Gillespie D. (1996) Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral. Physical Review E 54(2):2084-2091

Ginos BF. (2009) Parameter Estimation for the Lognormal Distribution (thesis). Brigham Young University: All Theses and Dissertations. Paper 1928

Gipps P. (1981) A Behavioural Car-Following Model for Computer Simulation. Transportation Research Part B: Methodological 15(2):105-111

Gipps P. (1987) Simulation of Pedestrian Traffic in Buildings. Universitat Karlsruhe: Institute fur Verkehrswesen. Schriftenreihe 35

Grayson G, Hydén C, Kraay J, Muhlrads N and Oppe S. (1984) The Malmö Study: A Calibration of Traffic Conflict Techniques. Institute for Road Safety Research. R-84-12

- Green M. (2000) "How Long Does It Take to Stop?" Methodological Analysis of Driver Perception-Brake Times. *Transportation Human Factors* 2(3):195-216
- Greenshields B. (1935) A Study of Traffic Capacity. *Highway Research Board Proceedings* 14:448-477
- Greibe, P. (2007) Braking Distance, Friction and Behaviour. Trafitec. Denmark.
- Grimm V, et al. (2005) Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology. *Science* 310:987-991
- Grover C, Knight I, Okoro F, Simmons I, Couper G, Massie P, Smith B. (2008) Automated Emergency Braking Systems: Technical Requirements, Costs, and Benefits. TRL Limited. Published Project Report 227
- Guzzella L and Onder C. (2004) Introduction to Modeling and Control of Internal Combustion Engine Systems. Springer-Verlag. Berlin, Germany.
- Habibovic A, Tivesten E, Uchida N, Bärgrman J, and Aust ML. (2013) Driver behavior in car-to-pedestrian incidents: An application of the Driving Reliability and Error Analysis Method (DREAM). *Accident Analysis and Prevention* 50:554-565
- Hacohen S, Shvalb N, and Shoval S. (2018) Dynamic Model for Pedestrian Crossing in Congested Traffic Based on Probabilistic Navigation Function. *Transportation Research Part C: Emerging Technologies* 86:78-96
- Hamdar S, Mahmassani H, and Treiber M. (2014) From Behavioral Psychology to Acceleration Modeling: Calibration, Validation, and Exploration of Drivers' Cognitive and Safety Parameters in a Risk-Taking Environment. *Transportation Research Part B: Methodological* 78:32-53
- Hammond, W. (1876) A Treatise on the Nervous System (6th ed). Appleton. New York, NY

Haque M and Washington S. (2014) A Parametric Duration Model of the Reaction Times of Drivers Distracted by Mobile Phone Conversations. *Accident Analysis & Prevention* 62:42-53

Harrell WA. (1993) The Impact of Pedestrian Visibility and Assertiveness on Motorist Yielding. *The Journal of Social Psychology* 133(3):353-360

Hayward J. (1971) Near Misses as a Measure of Safety at Urban Intersections. Thesis, Dept. of Civil Engineering. The Pennsylvania State University, Pennsylvania

Hayward J. (1972) Near Miss Determination Through Use of a Scale of Danger. The Pennsylvania State University, Pennsylvania. Report TTSC-7115

Heaviside O. (1892) On Operators in Physical Mathematics: Part I. *Proceedings of the Royal Society of London* 52:504-529

Heaviside O. (1899) *Electromagnetic Theory: Vol II. "The Electrician"* Printing and Publishing Company. London.

Helbing D. (1996) Derivation and Empirical Validation of a Refined Traffic Flow Model. *Physica A: Statistical Mechanics and its Applications* 233(1-2):253-282

Helbing D, Molnár P, Farkas I, and Bolay K. (2001) Self-Organizing Pedestrian Movement. *Environment and Planning B: Planning and Design* 28:361-383

Helbing D and Treiber M. (1998) Enskog equations for traffic flow evaluated up to Navier-Stokes order. *Granular Matter* 1:21-3199

Helbing D, Farkas I, and Vicsek T. (2000) Simulating Dynamical Features of Escape Panic. *Nature* 407:487-490

Helbing D, Hennecke A, Shvetsov V, and Treiber M. (2002) Micro- and Macro-Simulation of Freeway Traffic. *Mathematical and Computer Modeling* 35:517-547

Helman S and Palmer M. (2010) Road Worker Conspicuity: Daytime & Night Time. Transportation Research Laboratory. Client Project Report 1001 – 241(1308)HALC

Hendricks DL, Fell JC, and Freedman M. (2001) The Relative Frequency of Unsafe Driving Acts in Serious Traffic Crashes. NHTSA. Report DOT-HS-809-205

Hidas P. (2005) Modelling Individual Behavior in Microsimulation Models. 28th Australasian Transport Research Forum. Sydney, Australia

Hoogendoorn S. (2001) Normative Pedestrian Flow Behavior Theory and Applications. Social Science Research Council of the Netherlands Organization of Scientific Research. Delft, Netherlands

Hoogendoorn S and Bovy P. (2004) Pedestrian Route-Choice and Activity Scheduling Theory and Models. Transportation Research Part B 38:169-190

Hoogendoorn S and Hoogendoorn P. (2010) Generic Calibration Framework for Joint Estimation of Car-Following Models by Using Microscopic Data. TRR: Journal of the TRB 2188:37-45

Hope M. (2010) An Introduction to TRANSIMS: Core Components and New Developments. (presentation) USDOE: Transportation Research and Analysis Computing Center

Horni A, Nagel K, and Axhausen K eds. (2016) The Multi-Agent Transportation Simulation MATSim. Ubiquity Press, London

van der Horst A. (1984). The ICTCT Calibration Study at Malmo: A Quantitative Analysis of Video-Recordings. Report IZF 1984-37, TNO Institute for Perception, Soesterberg, Netherlands.

van der Horst A. (1990) A Time-Based Analysis of Road User Behaviour in Normal and Critical Encounters. Dissertation, TU Delft, Netherlands.

- Houten R, Malenfant J, and McCusker D. (2001) Advance Yield Markings: Reducing Motor Vehicle-Pedestrian Conflicts at Multilane Crosswalks with Uncontrolled Approach. TRR: Journal of the TRB 1773:69-74
- Hurts K, Angell L, and Perez M. (2011) The Distracted Driver: Mechanisms, Models, and Measurement. Reviews of Human Factors and Ergonomics 7(3)
- Hydén C. (1987) The development of a method for traffic safety evaluation: The Swedish Traffic Conflict Technique. (dissertation) Department of Traffic Planning and Engineering, Lund University. Scania, Sweden
- Ismail K, Sayed T, Saunier N, and Lim C. (2009) Automated Analysis of Pedestrian-Vehicle Conflicts Using Video Data. TRR: Journal of the TRB 2140:44-54
- Janssen W, Michon J, and Harvey L. (1976). The Perception of Lead Vehicle Movement in Darkness. Accident Analysis and Prevention 8:151-166
- Jennings N. (1999) On agent-based software engineering. Artificial Intelligence 117:277-29
- Johansson G and Rumar K. (1971) Drivers' Brake Reaction Times. Human Factors 13(1):23-27
- Jost D and Nagel K. (2003) Probabilistic Traffic Flow Breakdown in Stochastic Car-Following Models. TRR: Journal of the TRB 1852:152-158
- Kachroo P, Al-nasur SA, Wadoo A, and Shende A. (2008) Pedestrian dynamics: feedback control of crowd evacuation. Springer. Berlin
- Kahneman D and Tversky A. (1979) Prospect Theory: An Analysis of Decision Under Risk. Econometrica 47(2):263-292
- Kalos MH. (1962) Monte Carlo Calculations of the Ground State of Three- and Four-Body Nuclei. Physical Review 128(4):1791-1795

Kantowitz BH, et al. (2004) Development of Critical Gaps and Research Efforts: in Support of the Safety R & T Partnership Agenda. FHWA Contract # DTFH61-01-C-00049, Task 10

Kesting A. (2008) Microscopic Modeling of Human and Automated Driving: Towards Traffic-Adaptive Cruise Control. Doctoral Thesis: Faculty of Traffic Sciences, Technische Universität Dresden (Germany)

Kesting A and Treiber M. (2008a) Calibrating car-following models by using trajectory data: Methodological study. *Transport Res Record* 2088:148-156

Kesting A and Treiber M. (2008b) How Reaction Time, Update Time, and Adaptation Time Influence the Stability of Traffic Flow. *Computer-Aided Civil and Infrastructure Engineering* 23:125-137

Kesting A, Treiber M, and Helbing D. (2009) Agents for Traffic Simulation, in *Multi-Agent Systems: Simulation and Applications* (Uhrmacher A. and Weyns D eds.) Taylor & Francis Group. Boca Raton, FL. pp: 325-356

Kesting A, Treiber M, and Helbing D. (2010) Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity. *Philos. Trans. Royal Soc. A* 368:4585-4605

Kikuchi S and Chakroborty P. (1992) Car-Following Model Based on Fuzzy Inference System. *TRR: Journal of the TRB* 1365:82-91

Kim J and Mahmassani H. (2011) Correlated Parameters in Driving Behavior Models: Car-Following Examples and Implications for Traffic Microsimulation. *TRR: Journal of the TRB* 2249:62-77

Kim JK, Wang Y, and Ulfarsson G. (2007) Modeling the Probability of Freeway Rear-End Crash Occurrence. *Journal of Transportation Engineering* 133(1):11-19

Kim T and Zhang HM. (2011) Interrelations of Reaction Time, Driver Sensitivity, and Time Headway in Congested Traffic. TRR: Journal of the TRB 2249:52-61

Kittelson & Associates. (2006) Transportation System Simulation State-of-the-Practice Report. Available: sites.kittelson.com/TSSM/Downloads/Download/39734. Accessed: 2018/04/22

Knoop V, Hoogendoorn S, and van Zuylen H. (2009) Empirical Differences Between Time Mean Speed and Space Mean Speed, in *Traffic and Granular Flow '07* (Appert-Rolland C, et al. eds.) Springer-Verlag Berlin. pp: 351-356

Kovács T, Bolla K, Gil R, et al. [2016] Parameters of the Intelligent Driver Model in Signalized Intersections. Tehnički Vjesnik 23 (5):1469-1474

Kraay J, van der Horst A, and Oppe S. (1986). Handleiding voor de conflictobservatietechniek DOCTOR (Dutch Objective Conflict Technique for Operation and Research). Report R-86-3, Institute for Road Safety Research SWOV, Leidschendam, Netherlands. (in Dutch)

Kwan I and Mapstone J. (2004) Visibility Aids for Pedestrians and Cyclists: A Systematic Review of Randomised Controlled Trials. Accident Analysis and Prevention 36:305-312

Laberge J, Scialfa C, White C, and Caird J. (2004) Effects of Passenger and Cellular Phone Conversations on Driver Distraction. TRR: Journal of the TRB 1899:109-116

Langham M and Moberly N. (2003) Pedestrian Conspicuity Research: A Review. Ergonomics 46(4):354-363

Laureshyn A, Svensson Å, and Hydén C. (2010) Evaluation of traffic safety, based on micro-level behavioural data: Theoretical framework and first implementation. Accident Analysis and Prevention 42:1637-1646

Lee D. (1976) A Theory of Visual Control of Braking Based on Information about Time-to-Collision. *Perception* 5:437-459

Lee J, Regan M, and Young K. (2008) Defining Driver Distraction, in *Driver Distraction: Theory, Effects, and Mitigation* (Regan M, Lee J, Young K eds.) CRC Press. Boca Raton, FL. pp:31-40

Lee YC, Lee J, and Boyle L. (2009) The Interaction of Cognitive Load and Attention-Directing Cues in Driving. *Human Factors* 51(3):271-280

Li SS, Qian DL, and Luo Y. (2012) Microscopic Dynamic Simulation Model for Pedestrian at Signalized Intersection. *Journal of Central South University* 19(11):3351-3362

Li X, Yan X, Li X, and Wang J. (2012) Using Cellular Automata to Investigate Pedestrian Conflicts with Vehicles in Crosswalk at Signalized Intersection. *Discrete Dynamics in Nature and Society*. Volume 2012, Article 287502

Liang Y, Lee J, and Reyes M. (2007) Nonintrusive Detection of Driver Cognitive Distraction in Real Time Using Bayesian Networks. *TRR: Journal of the TRB* 2018:1-8

Liang Y, Lee J, and Yekhshatyan L. (2012) How Dangerous is Looking Away from the Road? Algorithms Predict Crash Risk from Glance Patterns in Naturalistic Driving. *Human Factors* 54(6):1104-1116

Lighthill MJ and Whitham GB (1955) On kinematic waves II: A theory of traffic flow on long crowded roads. *Proc. R. Soc. Lond. A* 229:317-345

Limpert E, Stahel W, and Abbt M. (2001) Log-normal Distributions across the Sciences: Keys and Clues. *BioScience* 51(5):341-352

Lindorfer M, Backfrieder C, Ostermayer G, and Mecklenbräuker C. (2018) A Stochastic Driver Distraction Model for Microscopic Traffic Simulations. 31st European Simulation and Modeling Conference. Lisbon, Portugal

Lister RD. (1950) The Reaction Time of Drivers in Moving and Stationary Vehicles. Road Research Laboratory: Technical Report RN/1324/RDL (as cited in Johansson and Rumar 1971)

Liu M, Zeng W, Chen P, and Wu X. (2017) A Microscopic Simulation Model for Pedestrian-Pedestrian and Pedestrian-Vehicle Interactions at Crosswalks. PLoS ONE 12(7):e0180992

Liu Y, Feyen R, and Tsimhoni, O. Queueing Network-Model Human Processor (QN-MHP): A Computational Architecture for Multitask Performance. ACM Transactions on Human-Computer Interaction 13(1):37-70

Lun C, Savage S, Jeffrey D, and Chepurniy N. (1984) Kinetic theories for granular flow: inelastic particles in Couette flow and slightly inelastic particles in a general flowfield. Journal of Fluid Mechanics 140:223-256

Malamuth N, Shayne E, and Pogue B. (1978) Infant Cues and Stopping at the Crosswalk. Personality and Social Psychology Bulletin 4(2):334-336

Malibongwe M. (2017) Adaptive Cruise Control System. Durban University of Technology. Technical Report. DOI: 10.13140/RG.2.2.22762.93122

Mathewson JH, Trautman DL, Gerlough DL (1955) *Study of Traffic Flow by Simulation*. Institute of Transportation and Traffic Engineering, UC Berkeley. Berkeley, CA

Matúš Š. (2014) Road Users' Strategies and Communication: Driver-Pedestrian Interaction. Presented at Transport Research Arena 5th Conference: Transport Solutions from Research to Deployment. Paris, France

Mehmood A and Easa S. (2009) Modeling Reaction Time in Car-Following Behaviour Based on Human Factors. International Journal of Civil, Environmental, Structural, Construction and Architectural Engineering 3(9):325-333

Metropolis N. (1987) The Beginning of the Monte Carlo Method. Los Alamos Science Special Issue 15:125-130

Metropolis N and Ulam S (1949) The Monte Carlo Method. Journal of the American Statistical Association 44(247):335-341

Michon J. (1971) Psychonie Onderweg (inaugural lecture). University of Groningen. Groningen, Netherlands (in Dutch)

Michon J. (1979) Dealing with Danger. Traffic Research Center of the University of Groningen. Groningen, The Netherlands. Technical Report VK 79-01

Migletz D, Glauz W, and Bauer K. (1985) Relationships Between Traffic Conflicts and Accidents Volume 2 – Final Technical Report. Federal Highway Administration. FHWA/RD-84/042

Moses R and Mtoi E. (2013) Evaluation of Free Flow Speeds on Interrupted Flow Facilities. Florida Department of Transportation Project. BDK83 977-18: Final Report

Mwakalonge J, Siuhi S, and White J. (2015) Distracted Walking: Examining the Extent to Pedestrian Safety Problems. Journal of Traffic and Transportation Engineering (English Edition) 2(5):327-337

Nagel K and Schreckenberg M. (1992) A cellular automaton model for freeway traffic. J Phys I France 2:2221-2229

von Neumann J and Morgenstern O. (1944) Theory of Games and Economic Behavior. Princeton University Press. Princeton, NJ

Newell G. (1961) Nonlinear Effects in the Dynamics of Car Following. Operations Research 9(2):209-229

NGSIM – Next Generation Simulation (2006) Federal Highway Administration. FHWA-HRT-06-137

NHTSA – National Highway Traffic Safety Administration (2017) Traffic Safety Facts: 2015 Data. NHTSA's National Center for Statistics and Analysis. Washington, DC. DOT-HS-812-375

NHTSA (2018) Traffic Safety Facts: 2016 Data – Pedestrians. NHTSA's National Center for Statistics and Analysis. Washington, DC. DOT-HS-812-493

Niazi M and Hussain A. (2011) Agent-based computing from multi-agent systems to agent-based models: a visual survey. *Scientometrics*: 1-21

OECD – Organization for Economic Co-operation and Development. (1998) Safety of Vulnerable Road Users. Paris, France. DSTI/DOT/RTR/RS7(98)1/FINAL

Ossen S and Hoogendoorn. (2010) Heterogeneity in car-following behavior: Theory and empirics. *Transportation Research Part C: Emerging Technologies* 19(2):182-195

Pancer S, Adams D, Mollard D, Solsberg D, and Tammen L. (1979) Perceived Distinctiveness of the Handicapped. *Journal of Social Psychology* 108:275-276

Parunak H, Savit R, and Riolo R. (1998) Agent-Based Modeling vs. Equation-Based Modeling: A Case Study and Users' Guide in *Proceedings of Multi-Agent Systems and Agent-Based Simulation Workshop*. Paris. pp:10-25

Perkins S and Harris J. (1967) Criteria for Traffic Conflict Characteristics: Signalized Intersections. General Motors Research Laboratories. Warren, MI. GMR-632

Phillips W. (1978) A kinetic model for traffic flow with continuum implications. *Transportation Planning and Technology* 5(3):131-138

Piccoli B and Tosin A. (2011) Time-Evolving Measures and Macroscopic Modeling of Pedestrian Flow. *Archive for Rational Mechanics and Analysis* 199(3):707-738

Pipes L. (1953) An Operational Analysis of Traffic Dynamics. *Journal of Applied Physics* 24:274-281

Ploeg J, Scheepers B, van Nunen E, van de Wouw N, and Nijmeijer H. (2011) Design and Experimental Evaluation of Cooperative Adaptive Cruise Control. 14th International IEEE Conference on Intelligent Transportation Systems. Washington, D.C.

Premebida C, Carreira J, Batista J, and Nunes U. (2014) Pedestrian Detection Combining RGB and Dense LIDAR Data. 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems. pp:4112-4117. Chicago, USA

Prigogine I and Herman R. (1971) *Kinetic Theory of Vehicular Traffic*. Elsevier. New York, NY

Przybyla J, Taylor J, Jupe J, and Zhou X. (2012) Simplified, Data-Driven, Errorable Car-Following Model to Predict the Safety Effects of Distracted Driving. 15th International IEEE Conference on Intelligent Transportation Systems. Anchorage AK.

PTV – Planung Transport Verkehr AG (2011) VISSIM 5.40 – User Manual

PTVGroup.com (online) PTV Vissim. Available: <http://vision-traffic.ptvgroup.com/en-us/products/ptv-vissim/>. Accessed: 2018/04/12

Pulugurtha S, Vasudevan V, Nambisan S, and Dangeti M. (2012) Evaluating Effectiveness of Infrastructure-Based Countermeasures for Pedestrian Safety. TRR: Journal of the TRB 2299:100-109

Punzo V and Simonelli F. (2005) Analysis and Comparison of Microscopic Traffic Flow Models with Real Traffic Microscopic Data. TRR: Journal of the TRB 1934:53-63

Quensel CE. (1947) The Validity of the Z-Criterion When the Variates are Taken from Different Normal Populations. Scandinavian Actuarial Journal 30:44-45

Quistberg DA, et al. (2015) Multilevel Models for Evaluating the Risk of Pedestrian-Motor Vehicle Collisions at Intersections and Mid-Blocks. Accident Analysis and Prevention 84:99-111

- Rajamani R and Shladover SE. (2001) An Experimental Comparative Study of Autonomous and Co-operative Vehicle-Follower Control Systems. *Transportation Research Part C: Emerging Technologies* 9:15-31
- Railsback S, Lytinen S, and Jackson S (2005) Agent-based Simulation Platforms: Review and Development Recommendations. *Simulation* 82(9):609-623
- Reason J. (1990) *Human Error*. Cambridge University Press. Cambridge, UK
- Reason J, et al. (1990) Errors and violations on the roads: a real distinction? *Ergonomics* 33(10-11):1315-1332
- Reichardt D. (2008) Approaching Driver Models Which Integrate Models of Emotion and Risk. Presented at 2008 IEEE Intelligent Vehicles Symposium. Eindhoven, The Netherlands
- Richards P. (1956) Shock Waves on the Highway. *Operations Research* 4(1):42-51
- Ronald N, Sterling L, and Kirley M. (2007) An Agent-Based Approach to Modelling Pedestrian Behaviour. *International Journal of Simulation* 8(1):25-38
- Rootzén H and Zholud D. (2016) Tail Estimation for Window-Censored Processes. *Technometrics* 53(1):95-103
- Roth M, Flohr F, and Gavrila D. (2016) Driver and Pedestrian Awareness-Based Collision Risk Analysis. Presented at: 2016 IEEE Intelligent Vehicles Symposium. Gothenburg, Sweden
- Rouphail N and Chae K. (2002) Evaluation and Application of Pedestrian Modeling Capabilities using Computer Simulation, Final Report. Prepared for HSRC and NIH
- Rouphail N, et al. (1998) Recommended Procedures: Chapter 13, "Pedestrians," of the Highway Capacity Manual (technical report). US DoT. FHWA-RD-98-107

- Saifuzzaman M and Zheng Z. (2014) Incorporating Human-Factors in Car-Following Models: A Review of Recent Developments and Research Needs. *Transportation Research Part C: Emerging Technologies* 48:379-403
- Saifuzzaman M, Zheng Z, Haque M, and Washington S. (2015) Revisiting the Task–Capability Interface Model for Incorporating Human Factors into Car-Following Models. *Transportation Research Part B: Methodological* 82:1-19
- Salvucci D. (2005) Modeling Tools for Predicting Driver Distraction. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 49:1149-1152
- Salvucci D. (2006) Modeling Driver Behavior in a Cognitive Architecture. *Human Factors* 48(2):362-380
- Løvås G. (1994) Modeling and Simulation of Pedestrian Traffic Flow. *Transportation Research Part B: Methodological* 28(6):429-443
- Sayer J and Buonarosa M. (2008) The Roles of Garment Design and Scene Complexity in the Daytime Conspicuity of High-Visibility Safety Apparel. *Journal of Safety Research* 39:281-286
- Schaap T. (2012) Driving Behaviour in Unexpected Situations (dissertation). University of Twente. Twente, Netherlands. TRAIL Thesis Series T2012/1
- Schaap T, van der Horst A, van Arem B, and Brookhuis K. (2013) The Relationship Between Driver Distraction and Mental Workload in *Driver Distraction and Inattention: Advances in Research and Countermeasures: Volume 1* (Regan M, Lee J, and Victor T ed.). CRC Press. Surrey UK. pp:63-82
- Schakel W, Knoop V, and van Arem B (2012) Integrated Lane Change Model with Relaxation and Synchronization. *TRR: Journal of the TRB* 2316:47-57
- Schelhorn T, O'Sullivan D, Haklay M, and Thurstain-Goodwin M. (1999) STREETS: An Agent-Based Pedestrian Model. Center for Advanced Spatial Analysis, University College London. London

- Schroeder B and Roupail N. (2011) Event-Based Modelling of Driver Yielding Behavior at Unsignalized Crosswalks. *Journal of Transportation Engineering* 137(7):455-465
- Shinar D. (2007a) Theories and Models of Driver Behavior, in *Traffic Safety and Human Behavior*. Emerald Group Publishing. Bingley, UK. pp:53-90
- Shinar D. (2007b) Vision, Visual Attention, and Visual Search, in *Traffic Safety and Human Behavior*. Emerald Group Publishing. Bingley, UK. pp:91-129
- Short M, Pont M, and Huang Q. (2004) Simulation of Longitudinal Vehicle Dynamics. Safety and Reliability of Distributed Embedded Systems. Technical Report ESL 04-01
- Singh S. (2015) Critical reasons for crashes investigated in the National Motor Vehicle Crash Causation Survey. National Highway Traffic Safety Administration. Washington, DC. Report No. DOT-HS-812-115
- Shoham Y. (1993) Agent-oriented programming. *Artificial Intelligence* 60(1):51-92
- Scholz F. (1996, revised 2001) Maximum Likelihood Estimation for Type I Censored Weibull Data Including Covariates. Boeing Phantom Works. ISSTECH-96-022
- Shvetsov V and Helbing D. (1999) Macroscopic Dynamics of Multi-Lane Traffic. *Physical Review E* 59(6):6328-6339
- Schweitzer N, Apter Y, Ben-David G, Liebermann DG, and Parush A. (1995) A Field Study on Braking Responses During Driving II: Minimum Driver Braking Times. *Ergonomics* 38(9):1903-1910
- Silva AW, Alves G, Marques Jr W, and Kremer G. (2008) Enskog's kinetic theory of dense gases for chemically reacting binary mixtures, I: Reaction rate and viscosity coefficients. *Physica A* 387:1733-1749
- Siuhi S and Kaseko M. (2016) Incorporating Vehicle Mix in Stimulus-Response Car-Following Models. *Journal of Traffic and Transportation Engineering* 3(3):226-235

de Smedt F, van Beeck K, Tuytelaars T, and Goedemé T. (2013) Pedestrian Detection at Warp Speed: Exceeding 500 Detections per Second. IEEE Conference on Computer Vision and Pattern Recognition Workshops

Snyder M. (1972) Traffic Engineering for Pedestrian Safety: Some New Data and Solutions. Highway Research Record (406):21-27

Still GK. (2000) Crowd Dynamics (thesis). University of Warwick. Coventry, England

Stutts J, et al. (2003) The Causes and Consequences of Distraction in Everyday Driving. Annual Proceedings: Association for the Advancement of Automotive Medicine 47:235-251

Sun D, Ukkusuri S, Benekahal R, and Waller ST. (2002) Modeling of Motorist-Pedestrian Interaction at Uncontrolled Mid-Block Crosswalks. TRR: Proceedings of the 2003 Annual Meeting. Washington, DC

Swarm Development Group. 1999. Swarm. www.swarm.org. Swarm Development Group. Santa Fe, NM

Tanaboriboon Y, Hwa SS, and Chor CH. (1986) Pedestrian Characteristics Study in Singapore. Journal of Transportation Engineering 112(3):229-235

Taylor D. (1964) Drivers' Galvanic Skin Response and the Risk of Accident. Ergonomics 7(4):439-451

Tesfatsion L. (2016) General Software and Toolkits: Agent-Based Computational Economics (ACE), Agent-Based Modeling (ABM), and Complex Adaptive Systems (CAS). Online: <http://www2.econ.iastate.edu/tesfatsi/acecode.htm> Accessed: 2017/03/04

Tesfatsion L and Judd K. (2006) Handbook of Computational Economics Volume 2: Agent-Based Computational Economics. Elsevier – North Holland. Amsterdam, Netherlands

Tobias R and Hofmann C. (2004) Evaluation of free Java-libraries for social-scientific agent based simulation. *Journal of Artificial Societies and Social Simulation* 7(1):online

Treat J, et al. (1979) Tri-Level Study of the Causes of Traffic Accidents: Final Report – Executive Summary. National Highway Traffic Safety Administration. Washington, DC. DOT-HS-805-099

Treiber M., Hennecke A., and Helbing D. (2000) Congested traffic states in empirical observations and microscopic simulations. *Phys Rev E* 62:1805-1824

Treiber M and Helbing D. (2003) Memory Effects in Microscopic Traffic Models and Wide Scattering in Flow-Density Data. *Physical Review E* 68 046119:1-8

Treiber M and Kesting A. (2010) Validation of Traffic Flow Models with respect to the Spatiotemporal Evolution of Congested Traffic Patterns. *Transportation Research Part C: Emerging Technologies* 21(1):31-41

Treiber M and Kesting A. (2013a) *Traffic Flow Dynamics: Data, Models and Simulation*. Springer. Heidelberg, Germany

Treiber M and Kesting A. (2013b) Microscopic Calibration and Validation of Car-Following Models – A Systematic Approach. *Procedia - Social and Behavioral Sciences* 80:922-939

Treiber M, Kesting A, and Helbing D. (2006) Delays, Inaccuracies and Anticipation in Microscopic Traffic Models. *Physica A* 360(1):71-88

Treiber M, Kesting A, and Helbing D. (2007) Influence of Reaction Times and Anticipation on the Stability of Vehicular Traffic Flow. *TRR: Journal of the TRB* 1999:23-29

Victor T. (2005) Keeping Eye and Mind on the Road. *Acta Universitatis Upsaliensis*. Digital Comprehensive Summaries of Uppsala Dissertation from the Faculty of Social Sciences 9

Victor T, et al. (2015) Analysis of Naturalistic Driving Study Data: Safer Glances, Driver Inattention, and Crash Risk. The Second Strategic Highway Research Program. Transportation Research Board. S2-S08A-RW-1

Vogel K. (2002) What Characterizes a “Free Vehicle” in an Urban Area? Transportation Research Part F: Psychology and Behaviour 5(1):15-29

van Wageningen-Kessels F, van Lint H, Vuik K, and Hoogendoorn S. (2015) Genealogy of traffic models. EURO Journal on Transportation and Logistics 4(4):445-473

Wagenmakers EJ and Brown S. (2007) Theoretical Note: On the Linear Relation Between the Mean and the Standard Deviation of a Response Time Distribution. Psychological Review 114(3):830-841

Waizman G, Shoval S, and Benenson I. (2015) Micro-Simulation Model for Assessing the Risk of Vehicle–Pedestrian Road Accidents. Journal of Intelligent Transportation Systems 19(1):63-77

Wang H, Jia L, Qianyong C, and Daiheng N. (2009) Speed-Density Relationship: From Deterministic to Stochastic. Transportation Research Board 88th Annual Meeting Compendium of Papers DVD

Wang M, Hoogendoorn S, Daamen W, van Arem B, Shyrokau B, and Happee R. (2016) Delay-Compensating Strategy to Enhance String Stability of Adaptive Cruise Controlled Vehicles. Transportmetrica B: Transport Dynamics

Weidmann U. (1993) Transporttechnik der Fussgänger: Transporttechnische Eigenschaften des Fussgängerverkehrs (Literaturauswertung). Schriftenreihe des Institute für Verkehrsplanung, Transporttechnik, Strassen, und Eisenbahnbau 90. Zürich

Wickens CD. (1981) Processing Resources in Attention and Workload. Engineering-Psychology Research Laboratory, University of Illinois at Urbana-Champaign. Report EPL-81-3/ONR-81-3

- Wiedemann R. (1974) Simulation des Strassenverkehrsflusses. Institute for Traffic Engineering, University of Karlsruhe (in German)
- Wierwille W and Tijerina L. (1988) Modelling the Relationship Between Driver In-Vehicle Visual Demands and Accident Occurrence, in *Vision in Vehicles – VI* (Gale A ed.) Elsevier. Oxford, UK. pp:233-243
- Wilde G. (1982) The Theory of Risk Homeostasis: Implications for Safety and Health. *Risk Analysis* 2(4):209-225
- Wilenski U and Rand W. (2015) An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo. MIT Press. Cambridge, MA
- Woodyard C. (2007) Vehicles keep inching up and putting on pounds. USA Today: Money: Cars. http://usatoday30.usatoday.com/money/autos/2007-07-15-little-big-cars_N.htm Accessed: 2017/02/10
- Wu J. (2017) Analysis of Pedestrian Safety Using Microsimulation and Driving Simulator. PhD Dissertation University of Central Florida College of Engineering and Computer Science
- Wu J, Radwan E, and Abou-Senna H. (2016) Pedestrian-Vehicle Conflict Analysis at Signalized Intersections Using Micro-Simulation. Presented at 17th Road Safety on Five Continents Conference. Rio de Janeiro, Brazil
- Xiang X, Kennedy R, and Madey G. (2005) Verification and Validation of Agent-based Scientific Simulation Models. Presented at the 2005 Agent-Directed Simulation Symposium. San Diego, California.
- Xin W, Hourdos J, Michalopoulos P, and Davis G. (2008) The Less-Than-Perfect Driver: A Model of Collision-Inclusive Car-Following Behavior. *TRR: Journal of the TRB* 2088:126-137

- Yang H and Peng H. (2009) Development of an Errorable Car-Following Driver Model. *Vehicle System Dynamics* 48(6):751-773
- Yang Y, Mingwu R, and Jingyu Y. (2014) Obstacles and Pedestrian Detection on a Moving Vehicle. *International Journal of Advanced Robotic Systems* 11(4)
- Yannis G, Papadimitriou E, and Theofilatos A. (2013) Pedestrian Gap Acceptance for Mid-Block Street Crossing. *Transportation Planning and Technology* 36(5):450-462
- Zębala J, Ciępka P, and Reza A. (2012) Pedestrian Acceleration and Speeds. *Problems of Forensic Science* 91:227-234
- Zegeer CV, et al. (2005) Safety Effects of Marked versus Unmarked Crosswalks at Uncontrolled Locations: Final Report and Recommended Guidelines. US DoT FHWA-HRT-04-100
- Zegeer J, et al. (2008) Default Values for Highway Capacity and Level of Service Analyses. National Cooperative Highway Research Program Report 599
- Zhang X and Chang GL. (2014) A Mixed-Flow Simulation Model for Congested Intersections with High Pedestrian-Vehicle Traffic Flows. *Simulation: Transactions of the Society for Modeling and Simulation International* 90(5):570-590
- Zhang X, Cheng L, Li B, and Hu HM. (2017) Too Far to See? Not Really! — Pedestrian Detection with Scale-aware Localization Policy. (submitted for review – arXiv:1709.00235v1 [cs.CV])
- Zheng L, Ismail K, and Meng X (2014) Traffic Conflict Techniques for Road Safety Analysis: Open Questions and Some Insights. *Canadian Journal of Civil Engineering* 41(7):633-641
- Zheng Y, et al. (2015) Modeling Vehicle-Pedestrian Interactions Outside of Crosswalks. *Simulation Modelling Practice and Theory* 59:89-101

Zhuang X and Wu C. (2014) Pedestrian Gestures Increase Driver Yielding at Uncontrolled Mid-Block Road Crossings. *Accident Analysis & Prevention* 70:235-244

APPENDIX

APPENDIX A – VARIABLE VALUES USED IN SIMULATIONS

Variable	Name	Units	Parameters								
			Constant	Normal		Lognormal			Weibull		
				<i>M</i>	<i>S</i>	γ	μ	σ	γ	λ	<i>k</i>
<i>A</i>	Pedestrian repulsive interaction force constant	$kg \cdot m/s^2$	2000	-	-	-	-	-	-	-	-
<i>a_{max}</i>	IDM maximum comfortable acceleration	m/s^2	-	-	-	-	0.132	0.6461	-	-	-
<i>B</i>	Pedestrian repulsive interaction space constant	<i>m</i>	0.08	-	-	-	-	-	-	-	-
<i>b</i>	IDM maximum comfortable deceleration	m/s^2	-	-	-	-	0.5372	0.7246	-	-	-
<i>b_{yield}</i>	Comfortable yielding deceleration	m/s^2	3	-	-	-	-	-	-	-	-
<i>c₁</i>	Pedestrian gap acceptance constant	-	6.2064	-	-	-	-	-	-	-	-
<i>c₂</i>	Pedestrian gap acceptance constant	-	0.942	-	-	-	-	-	-	-	-
<i>D_i</i>	Duration of distraction event	<i>s</i>	-	-	-	-	-0.455	0.6107	-	-	-
<i>d_{stop}</i>	Stopping distance from crosswalk	<i>m</i>	9	-	-	-	-	-	-	-	-
δ	IDM free acceleration exponent	-	4	-	-	-	-	-	-	-	-
λ_{ID}	Arrival rate of distraction events	s^{-1}	0.352	-	-	-	-	-	-	-	-
<i>m_i</i>	Pedestrian mass	<i>kg</i>	80	-	-	-	-	-	-	-	-

TABLE 37 – VARIABLE DISTRIBUTIONS USED IN SIMULATIONS

Variable	Name	Units	Parameters								
			Constant	Normal		Lognormal			Weibull		
				M	S	γ	μ	σ	γ	λ	k
r	Pedestrian radius	m	0.87	-	-	-	-	-	-	-	-
s_0	IDM jam headway	m	-	-	-	-	0.6517	0.4979	-	-	-
σ_r	Standard deviation of relative approach rate	s^{-1}	0.01	-	-	-	-	-	-	-	-
T	IDM time-headway	s	-	1.266	0.507	-	-	-	-	-	-
Δt	Simulation time-step	s	0.1	-	-	-	-	-	-	-	-
$\tilde{\tau}$	Persistence time of perception errors	s	20	-	-	-	-	-	-	-	-
τ_{ADRT}	Acceleration/deceleration reaction time	s	-	-	-	0.25	-0.60692	1.19376	-	-	-
τ_{BRT}	Brake perception reaction time	s	-	-	-	-	-	-	0.25	1.20	2.435
τ_{DR}	Braking device response time	s	0.2	-	-	-	-	-	-	-	-
τ_M	Braking movement time	s	0.15	-	-	-	-	-	-	-	-
τ_{ped}	Characteristic time of pedestrian acceleration	s	0.5	-	-	-	-	-	-	-	-
v_t^0	Pedestrian natural walking speed	m/s	-	1.4	0.26	-	-	-	-	-	-
v_f	IDM free-flow speed	km/hr	-	PSL + 3.22	7.24	-	-	-	-	-	-
V_s	Variation coefficient of distance perception errors	-	0.1	-	-	-	-	-	-	-	-

TABLE 37 (CONTINUED) – VARIABLE DISTRIBUTIONS USED IN SIMULATIONS

APPENDIX B – MACROSCOPIC EQUIVALENT TO IDM

The authors of the IDM have developed a macroscopic equivalent that can be classified as a modified gas-kinetic model. [Helbing et al. 2002] This model still utilizes Equations [1] and [6], but instead of the equilibrium velocity equation of the Lighthill-Whitham model, it uses a dynamic velocity equation of a form similar to other macroscopic models:

$$\frac{\partial V}{\partial t} + V \frac{\partial V}{\partial x} + \frac{1}{\rho} \frac{\partial P}{\partial x} = \frac{1}{\tau_{macro}} (V_e - V) \quad [58]$$

Here, using terminology from fluid dynamics, the second term is a *transport term*, the third is a *pressure term*, and the right-hand side is a *relaxation term*. This model differs from others in that it takes into account the finite space occupied by vehicles (as opposed to assuming point-like particles), and the *equilibrium velocity* on the right-hand side is dynamic and non-localized (allowing drivers to react to the traffic situation ahead of them):

$$V_e = V_f - \tau_{macro} [1 - p(\rho')] \chi(\rho') \rho' \mathcal{B}(\Delta V, S) \quad [59]$$

In this equation, the quantity subtracted from the free-flow speed on the right-hand side is a *braking term*, describing the interaction with vehicles ahead. A prime indicates that the variable is taken at the *interaction point*, $x' = x + s_{macro}$. For simplicity, the safe headway, s_{macro} , is assumed to vary linearly with velocity:

$$s_{macro} = \gamma \left(\frac{1}{\rho_{max}} + T_{avg} V \right) \quad [60]$$

Here, ρ_{max} is the density at a standstill, and γ is an anticipation factor for future velocity changes. (Since the location of the next car is the minimum distance ahead that can affect driver behavior, $\gamma \geq 1$. However, it is typically quite small.)

In Equations [58] and [59], τ_{macro} represents the acceleration time, or *relaxation term*⁵² (as can be seen by substituting Equation [59] into Equation [58] if we ignore the braking interaction and set the spatial derivatives to 0: $\frac{\partial V}{\partial t} = \frac{1}{\tau_{macro}}(V_f - V)$.) [Helbing and Treiber 1998] The second term on the right side of Equation [59] is an adaptation of the Enskog theory of particle interactions within a dense gas, [Silva et al. 2008] incorporating methods from granular flow. [Lun et al. 1984] It includes the density-dependent probability that a vehicle will overtake its leader,⁵³ $p(\rho)$; and a pair correlation function, $\chi(x) = 1/[1 - \rho(x, t)s_{macro}]$. Together, these define the *effective cross-section* for the passing interaction:

$$[1 - p(\rho)]\chi(\rho) = \frac{v_f T_{avg}^2}{\tau_{macro} \alpha(\rho_{max})} \frac{\rho}{(1 - \rho/\rho_{max})^2} \quad [61]$$

α is a *structure factor* that, based on empirical data, relates the squared average velocity at a point in the road to the velocity variance, ϕ , at that point: $\phi = \alpha(\rho)V^2$.

[Phillips 1978] ϕ is higher in dense traffic than free-flowing traffic.⁵⁴

⁵² The value is density dependent: $\tau_{macro}(\rho) \approx 8s/(0.97e^{-\rho/(16 \text{ km}^{-1})} + 0.03)$ [Helbing & Treiber 1998]

⁵³ $p(\rho) \approx e^{-\rho/(16 \text{ km}^{-1})}$

⁵⁴ The data can be approximated by a Fermi function. For a detailed analysis and derivation of this and the macroscopic theory it is part of, see [Helbing 1996]

The final portion of equation [59] is a Boltzmann factor, $\mathcal{B}(\Delta V, S)$. For vehicle interactions, this can be written: [Shvetsov and Helbing 1999]

$$\mathcal{B}(\Delta V, S) = S\{\Delta V N(\Delta V) + [1 + (\Delta V)^2]E(\Delta V)\} \quad [62]$$

where

$$N(w) = \frac{e^{-y^2/2}}{\sqrt{2\pi}} \quad [63]$$

is the normal distribution and $E(y) = \int_{-\infty}^y N(z)dz$, the Gaussian error function, depends on the *effective dimensionless velocity difference*, ΔV between the velocities at x and x' : $\Delta V = \frac{v-v'}{\sqrt{S}}$. Here, S ($\neq s$) is the difference in velocity variance, φ , between the point in question and the interaction point: $S = \varphi - \varphi'$.

Returning to equation [58], the pressure relation must be defined. Ignoring differences in the average velocity of adjacent lanes, this is simply $P = \rho \langle \varphi_i \rangle$, where $\langle \varphi_i \rangle$ is the weighted average velocity variance across all I lanes:

$$\langle \varphi_i \rangle = \sum_{i=1}^I \frac{\rho_i}{I\rho} \varphi_i \quad [64]$$