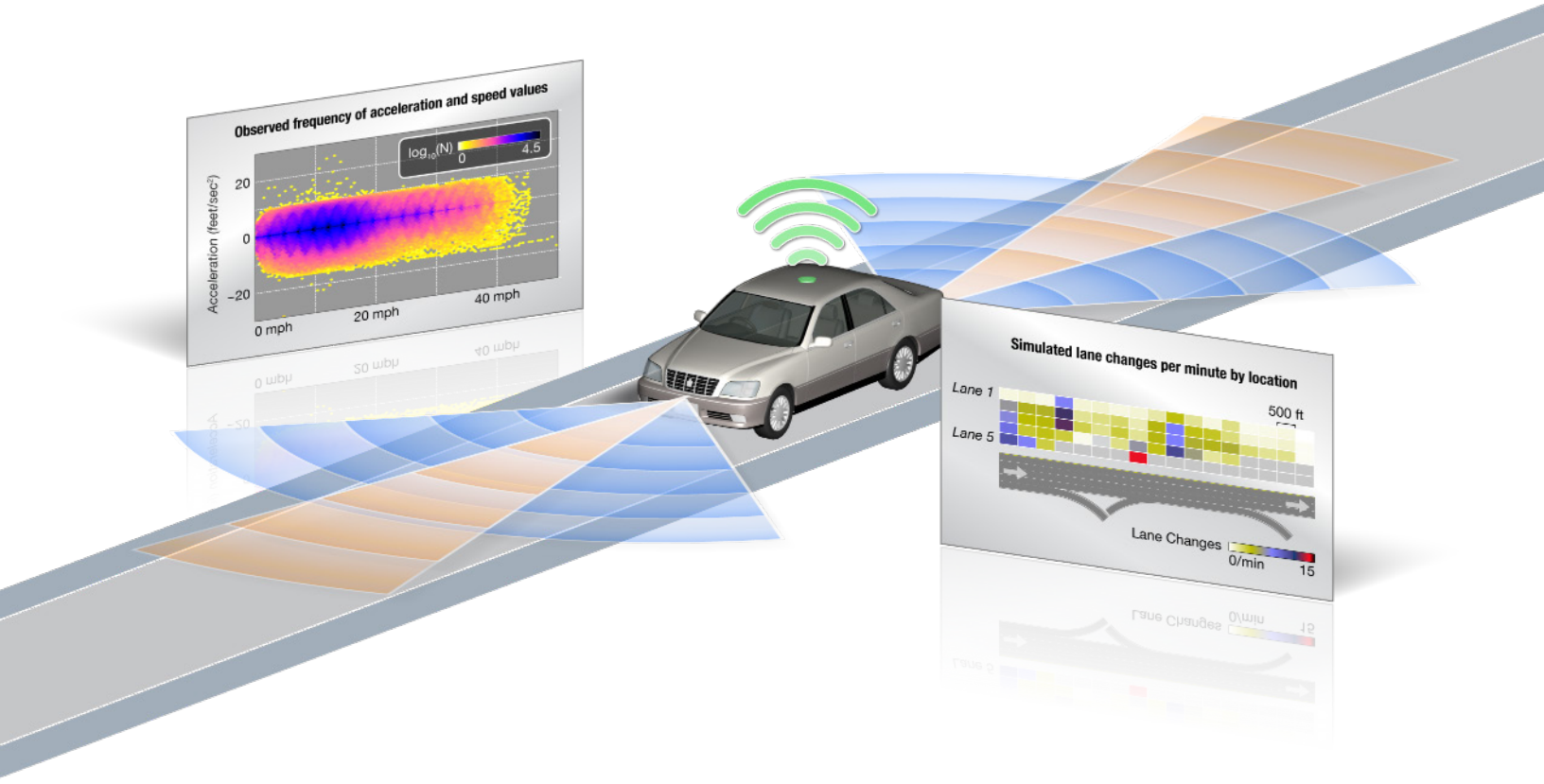
Proof of Concept for the Trajectory-Level Validation Framework for Traffic Simulation Models

www.its.dot.gov/index.htm

Final Report—October 2017
FHWA-JPO-16-407



U.S. Department of Transportation

Produced by Cambridge Systematics, Inc.
U.S. Department of Transportation
Office of the Assistant Secretary for Research and Technology
Intelligent Transportation System (ITS) Joint Program Office (JPO)

Notice

Cover Image Source: Cambridge Systematics, Inc.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-JPO-16-407		

4. Title and Subtitle	5. Report Date
Proof of Concept for the Trajectory-Level Validation Framework for Traffic Simulation Models	October 2017
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Michalis Xyntarakis, Vassili Alexiadis, Vincenzo Punzo, Robert Campbell, Alex Skabardonis, Erin Flanigan	

9. Performing Organization Name And Address	10. Work Unit No. (TRAIS)
Cambridge Systematics 3 Bethesda Metro Center, Suite 1200 Bethesda, MD 20814	11. Contract or Grant No. DTFH61-12-D-00042

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Office of Operations Federal Highway Administration 1200 New Jersey Avenue, S.E. Washington, D.C. 20590	Final Report
	14. Sponsoring Agency Code

15. Supplementary Notes

FHWA GTM: James Colyar, Office of Transportation Management

16. Abstract

Based on current practices, traffic simulation models are calibrated and validated using macroscopic measures such as 15-minute averages of traffic counts or average point-to-point travel times. For an emerging number of applications, including connected vehicles, the realism of simulated driver dynamics at the second-by-second or subsecond trajectory level plays an important role.

This report presents a Proof of Concept Application (PCA) for the Trajectory Level Validation Framework, which is presented in a separate report. The structure of this document parallels that of the Validation Framework report, enabling the reader to easily cross-reference descriptions and discussion details between the two. This PCA document contains the validation performance measures and validation tests presented in the Validation Framework, including those related to: safety, acceleration limits and comfort, traffic flow, and lane changing. This document illustrates how practitioners can use the Validation Framework to assess the realism of the simulated vehicle dynamics in a model. By analogously applying this Validation Framework to their own data sets and simulation models, analysts and practitioners can more precisely measure and improve the performance of their traffic simulation models.

17. Key Words	18. Distribution Statement
Traffic simulation, data collection, vehicle trajectories, validation	No restrictions.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	73	N/A

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
	NOTE: volumes greater than 1000 L shall be shown in m^3			
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	$candela/m^2$	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
°C	Celsius	1.8C+32	Fahrenheit	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	$candela/m^2$	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. (Revised March 2003)

Source: Federal Highway Administration.

Acknowledgments

A The Federal Highway Administration team that oversaw and contributed to the development of this project was led by James Colyar and consisted of John Halkias, James Sturrock, Paul Heishman, and Chris Melson.

The project team would like to thank many members of the transportation community who have shared their thoughts and advice during the development of this project. Specifically, the project team would like to thank Mark Brackstone (AIMSUN), Michael Mahut (INRO), Dan Morgan (Caliper), Ben Coifmann (Ohio State), Jordi Casas (TSS), Keir Opie (Cambridge Systematics), Karl Wunderlich (Noblis), Xuesong Zhou (University of Arizona), Jorge Laval (Georgia Tech), Lily Elefteriadou (University of Florida), and Kaan Ozbay (NYU) among many others who participated in public outreach webinars and provided feedback. Finally, the project team would like to thank the Transportation Research Board (TRB) Joint Simulation Subcommittee and the Traffic Flow Theory and Characteristics Committee for their invitations to present project objectives and findings at subcommittee meetings and at the TRB annual meeting. It should be noted that the views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of any agency of the U.S. government or the project advisors.

Table of Contents

List of Figures

List of Tables

Executive Summary

Traffic microsimulation models are designed to emulate vehicle dynamics at the microscopic subsecond level. Realistic vehicle dynamics are a prerequisite for the successful application of simulation models in a variety of existing and emerging fields. Modeling the benefits of safety and mobility of emerging transportation technologies, such as connected vehicles, is not possible unless the simulation models can realistically capture how drivers accelerate, decelerate, select speeds and change lanes.

This report presents a proof-of-concept application based on a framework of microscopic and macroscopic tests to validate simulated vehicle dynamics in terms of safety, comfort, feasibility, and compatibility with known traffic flow properties. The validation tests presented in this report have been categorized into three major application areas: a) **safety-related** that quantify driver aggressiveness and collision risk; b) **acceleration-based** tests associated with mechanical feasibility and driver comfort; and c) **traffic flow** modeling tests linked to microscopic and macroscopic properties of traffic flow. Taken in total, the validation tests have been designed to cover all aspects of simulated vehicle dynamics, including longitudinal (car-following) and lateral (lane-changing) movements.

The Next Generation Simulation (NGSIM) data for the I-80 corridor in Berkeley California is the primary field trajectory dataset used for the proof of concept application. Field trajectories for a 30-minute time period between 5:00 p.m. and 5:30 p.m. were used as the basis for all trajectory-level analysis and simulation model building. Simulated trajectories from a model of the I-80 corridor that had been previously been built were extracted and validated using the tests and procedures of the validation framework. It should be stressed that the results of any of the trajectory tests in this document serve the purpose to demonstrate the application of the trajectory validation framework and the details of its constituent tests. None of the validation tests should be considered representative of the simulation tool used, the capabilities of simulation tools in general, or the properties of a well calibrated model at the macroscopic level. For this reason, this report does not mention the particular simulation tool deployed. The study team would like to thank the software vendor for the provided support.

For each validation test, insights from naturalistic and other trajectory datasets are provided as references. If local trajectory data exist, the validation tests can be used to statistically test observed and simulated distributions of trajectory measures such as the time to collision. If local trajectory data do not exist, the insights from naturalistic or other trajectory datasets can be used as a point of insight but not as the basis for statistical comparisons. This is because congestion level and location-specific parameters can have a significant impact in the range and distribution of a particular trajectory measure. Additional research and analytics on trajectory datasets is required to establish invariants and norms. As more and more trajectory datasets become available it is expected that a growing body of knowledge on the properties and distribution of trajectory measures will be of further assistance to the simulation analyst.

Observed and simulated trajectory measures used for comparison include point estimates such as the number of safety events or lane changes per vehicle mile, one-dimensional distributions such as the time-to-collision distribution, and two dimensional distributions such as the speed versus acceleration relationship, or the distribution of lane change rate by time and space. The chi-squared statistical test is used to compare observed and simulated point estimates and one-dimensional distributions. Two-dimensional distributions are compared only qualitatively as the statistical tools for such comparisons require additional assumptions and caveats. Empirical comparisons also are used when a statistical comparison is not applicable or the results are orders of magnitude different.

In terms of safety tests, it was found that the time-to-collision distributions from 0 to 15 seconds are statistically different in the model and field data. However, the positional error in the field data may have

obfuscated this test. For example, time to collision less than one second was much more frequently observed in raw NGSIM data compared to the reconstructed NGSIM or the naturalistic data. The outcomes of the rest of the safety tests were more revealing about vehicles engaging in risky maneuvers. In this POC test safety events such as collisions and near-crashes happen three orders of magnitude more frequently in the model compared to the observed data. Lane change maneuvers on the other hand were executed with comparable risk in the observed and modeled trajectories.

In terms of vehicle limits and driver comfort, it was found that the maximum simulated deceleration significantly exceeds what currently is mechanically possible. Feasible maximum deceleration under favorable conditions is close to 1g (32 fpss) and never exceeds 2g. In the model, however, accelerations greater than 2g occurred approximately once every 10 vehicle miles. *The relationship between speed and acceleration at different speed levels showed significant differences that are likely to impact emissions calculations.* Specifically, the maximum acceleration increased by speed until approximately 20 miles per hour even though field data and mechanical charts indicate that maximum acceleration decreases by speed. Increased simulated acceleration variability was also measured at different speed levels. The acceleration root mean square test revealed that acceleration in the model fluctuates significantly more compared to field data.

The number of lane changes in the model is significantly higher than what is observed. Furthermore, the distribution of lane change intensity by time and space reveals that the same high intensity of lane-change activity that is observed at the on-ramp also is observed on the rest of the lanes. In the field data the majority of the drivers stay on the same lane while in the model the majority of the drivers change at least one lane. The speed versus density relationship of the fundamental diagram also shows significant differences. For densities greater than 50 vehicles per mile that correspond to congested conditions the space mean speed of the model is approximately half of what is observed on the field.

Chapter 1. Introduction

Traffic simulation models are designed to emulate vehicle dynamics or the behavior of individual drivers in traffic at a detailed level, typically at a second or subsecond resolution for microsimulation models. Specifically, microsimulation models emulate how vehicles accelerate, decelerate, and change lanes in response to travel and driving goals, and to surrounding traffic. Operational needs related to connected, autonomous, and active transportation and demand management (ATDM) applications rely on the realism of the simulated traffic dynamics to provide robust policy recommendations and sound transportation investment decisions. In ATDM applications, for example, realistic vehicle dynamics are required for model calibration. In connected and automated vehicle (CAV) applications, realistic emulation of vehicle dynamics is a prerequisite for measuring the impacts of human-assistive technologies such as collision warning systems. More generally, to meaningfully evaluate a technology or scenario, the analyst should ensure realistic driving dynamics are included in the model with and without the technology being tested.

Current methodologies to develop, calibrate and validate simulation tools are based on approaches that are using macroscopic or aggregate-level field data such as 15-minute averages of volumes and speeds. When using aggregate data only, the model calibration and validation space becomes an arbitrary multiparametric field that often provides too many options to achieve the same aggregate goodness-of-fit. For example, to improve the modeled travel time on a link, one can change the desired speed, driver aggressiveness for overtaking, or even the travel demand, all of which may yield a better macroscopic goodness-of-fit. However, if these changes are not done systematically, the result can be an overfitting of the model to observed data, which reduces the descriptive power of the model at both the macroscopic and microscopic level.

Validation of simulated vehicle dynamics at the subsecond trajectory level is not frequently performed in practice for two main reasons. First, there is no conceptual framework that defines validation tests at the microscopic trajectory level. This is in part due to the fact that a limited number of trajectory datasets exist to function as a baseline of comparison for vehicle dynamics measurements. Second, analysts often do not have access to local trajectory data for the particular location, time period, and driver population to be modeled. Trajectory data (e.g., sequences of time-stamped positions) have limited transferability and can vary significantly on a day-to-day basis even for the same driver and the same location. Although a holistic picture of driver behavior under different traffic, weather, incident, and information provision situations may still be elusive, the increasing number of trajectory datasets becoming available can provide further insights into the patterns and norms of driver behavior.

1.1 Purpose of this Document

This Proof of Concept Application (PCA) report illustrates the application of the Framework for Validating a Traffic Simulation Models and the Vehicle Trajectory Level. (Xyntarakis, Alexiadis, Punzo, Campbell, Skabardonis, and Flanigan. *A Framework for Validating Simulation Models at the Vehicle Trajectory Level.* Federal Highway Administration, Office of Operations. Final Report. November 2017. FHWA-JPO-16-405.) The framework hereafter it will be referred to as the Validation Framework. This report is designed to parallel the organization and content of the Validation Framework document as closely as possible, to facilitate cross referencing between the two. The tests covered in this report are listed in Table 1-1.

Table 1-1. Overview of validation measures and tests.

Validation Measures and Tests	Description
Safety	
Time to collision (TTC)	Quantifies risk of collision and driver behavior in terms of aggressiveness. TTC is a widely used safety surrogate measure.
Number of rear-end safety events	Events include rear-end crashes, near-misses, and warning messages. The number of such events per mile quantifies driver safety and aggressiveness.
Time Gap	The time until the subject vehicle reaches the current position of the principal other vehicle.
Lane Change Urgency (LCU)	Quantifies lane-changing risk from the perspective of the subject vehicle. It also quantifies tailgating risk.
Lane Change Severity (LCS)	Quantifies lane-changing risk from the perspective of the vehicle that can be cut off on the adjacent lane.
Vehicle Limits and Driver Comfort Levels	
Acceleration range (AR)	Determines feasible vehicle dynamics using ranges. Comfortable acceleration limits are defined in this document.
Acceleration jerk (AJ)	Determines feasible vehicle dynamics and characterizes traveler discomfort.
Acceleration root mean square (ARMS)	Determines driver comfort based on International Organization for Standardization (ISO) guidelines.
Traffic Flow	
Lane change type (LT)	Mandatory versus discretionary lane changes.
Lane change rate (LCR)	Reveals lane changing intensity over time and space.
Fundamental Diagram (FD)	Identifies macroscopic properties of flow and dispersion around ideal traffic flow theory assumptions.

Source: Cambridge Systematics, Inc.

For each test, this report describes how the corresponding measure was derived from observed and simulated data. Results from the simulated trajectories are also compared to either observed trajectories from the same site, or to reference data/distributions presented in the Validation Framework, as appropriate.

This report presents the results associated with applying the Validation Framework to the following three data sets. In addition to the three datasets that were processed as part of this study, publicly available summary tables from naturalistic data collected from multiple areas also were used.

1. Field data on I-80 that cover the entire traffic stream and collected under Next Generation Simulation (NGSIM) program.
2. Simulated trajectories that emulate the aggregate properties of the NGSIM dataset.
3. Field data recently obtained from a single instrumented vehicle driving multiple times on I-80 and nearby arterials.

These three data sets are described in greater detail in Chapter 2. In Chapter 3, safety related tests such as the number of near-crashes per vehicle mile are presented. Chapter 4 contains three tests specific to

vehicle limits and driver comfort. Chapter 5 presents four tests related to lane changing and macroscopic properties of traffic flow. A final chapter at the end of this report presents a summary of overall conclusions based on these tests.

1.2 The Validation Framework

The validation framework, and accordingly this PCA report, is divided into three major application areas or groups:

1. Safety-related tests that quantify driver aggressiveness and collision risk.
2. Tests associated with mechanical limitations, physical constraints, and driver comfort.
3. Traffic flow modeling tests linked to the microscopic and macroscopic properties of traffic flow.

The validation tests have been designed to cover all aspects of vehicle dynamics, including both longitudinal (car-following) and lateral (lane-changing) movements. Although all tests are related to each other, each of them highlights a different aspect of vehicle dynamics.

In a given simulation model, the level of realism in vehicle dynamics may not be accurately captured by conventional measures of model calibration and reasonableness alone, such as agreement with macroscopic measures (e.g., traffic counts and travel times) or queue lengths. While the tests presented in the Validation Framework can be used to assess the reliability and accuracy of a model on a finer scale, it is equally important to recognize that this framework alone is not a sufficient condition for determining whether a model accurately portrays real-world conditions accurately and reliably, and that a proper calibration/validation effort would include consideration of both the conventional macroscopic model considerations and the microscopic validation measures presented herein.

The framework and the associated tests can assist the analyst in producing better models of driver behavior and traffic flow at the microscopic and macroscopic levels. For example:

- **Practitioners** can use the framework to assess the realism of the resulting vehicle dynamics of any model or scenario.
- **Researchers** can use the framework to document, analyze, and assess the properties of different microsimulation models.
- **Software developers** can use the framework to assess and document the properties of their software packages for different types of applications.

Additional detail about the Validation Framework is available in the full Validation Framework report.

Chapter 2. Trajectory Datasets

A simulation model was built based on the Next Generation Simulation (NGSIM) I-80 (Emeryville, CA) trajectory data from 5:00 p.m. to 5:30 p.m. Network geometry and travel demand were obtained from the NGSIM project. Congestion in the NGSIM corridor is exogenous; the source of congestion is outside the boundaries of the collected data. To emulate this, a zone of slow speed was introduced right after the end of the NGSIM corridor. The composition of the vehicle fleet in terms of cars, trucks, and motorcycles was obtained from the NGSIM data. A number of calibration techniques based on macroscopic variables such as average vehicle speeds were employed. It should be noted that:

The results of any of the trajectory validation tests in this document serve the purpose to demonstrate the validation framework and the application of trajectory tests. Any of the outcomes should NOT be considered representative of:

- **One or more simulation tools or the capabilities of simulation tools in general.**
- **The properties of a well-calibrated model at the macroscopic level.**

The following three data sets were analyzed in this report:

1. Field data on I-80 that cover the entire traffic stream and collected under the NGSIM program.
2. Simulated trajectories that emulate the aggregate properties of the first dataset.
3. Field data recently obtained from a single instrumented vehicle driving multiple times on I-80 and nearby arterials.

No naturalistic trajectory datasets were analyzed but summary tables that have been previously published have been used in this work. This chapter provides additional background, context, and descriptive details about each of these data sets.

2.1 Simulated Trajectories for Next Generation Simulation I-80 Site

Trajectory data from the simulation model of the I-80 NGSIM site were used to illustrate the application of the validation framework. The input volume by lane in the model matches what has been observed in the NGSIM data. In addition to matching the inputs, the model was calibrated using macroscopic methods (e.g., aggregate volumes and speeds) to recreate to the extent possible by the available resources the traffic conditions and facility performance as captured in the NGSIM field data set for I-80. The macroscopic goodness-of-fit of the model is discussed in detail in Chapter 3.

Although some validation measures are influenced by the level of macroscopic model calibration performed (such as the frequency of lane changes), others may reflect vehicle-to-vehicle dynamics and may be independent of the model calibration effort (such as the maximum acceleration values observed in simulation). To establish which tests are sensitive to calibration and which are not, additional research is required to take into account the sensitivity of vehicle dynamics.

In addition to the portion that overlaps with the NGSIM site the simulation model also includes the following downstream freeway junctions to the north of the NGSIM site: University Ave, Gilman St, and Buchanan St. However, even though the surrounding spatial detail is included in the model, only the traffic demands for the NGSIM segment are simulated (e.g., ramp volumes for junctions downstream on the corridor are set to zero). The simulation outputs and trajectories used in the PCA are based only on the NGSIM portion of the corridor. Downstream congestion is emulated by reducing vehicle speeds at the end of the NGSIM corridor to what is observed in the NGSIM data.

2.2 Next Generation Simulation Field Data from I-80

As a real-world comparison baseline for simulation model performance, trajectories from the FHWA Next Generation Simulation (NGSIM) datasets were used in this PCA. Specifically, this proof of concept test focused on the I-80 NGSIM site in Emeryville, California, which spans the eastbound side of I-80 between approximately Powell St and Ashby Avenue. (https://www.fhwa.dot.gov/publications/research/operations/06137/index.cfm.) The NGSIM dataset is available through the Research Data Exchange (RDE) Website. (https://www.its-rde.net/.) Figure 2.1 presents the I-80 corridor segment where NGSIM data were collected.

The NGSIM data sets, including the I-80 dataset, were designed to capture the trajectories of the entire traffic stream, such that a complete awareness of all vehicle movements on the roadway segments is achieved. The intent of these data, whose collection was sponsored by FHWA, was to provide researchers with the data resources to investigate, analyze, evaluate, and better understand freeway lane selection, cooperative merging, and simulation modeling of oversaturated freeway conditions. The trajectories have been extracted from overhead camera recordings, with equipment mounted on the roof of the nearby Pacific Park Plaza building for the I-80 site (at a height of 97 meters above the roadway). Conversion of the video footage into trajectories was achieved using the NGSIM-VIDEO software at resolutions of 10 Hz, and was performed as part of a separate FHWA-sponsored effort. (http://www.webpages.uidaho.edu/niatt/Internal/directors_notes/UIdaho%200406_NGSIM%20and%20Simulation_JColyar.pdf (accessed 4-6-2017).)

These data were collected on April 13, 2005, and of the 10 hours of footage recorded for this 500-meter section of freeway, 45 minutes of it were processed for the PM Peak, yielding 5,648 vehicle trajectories. (https://arxiv.org/pdf/0804.0108.pdf (accessed 4-6-2017).) For the purposes of this PCA, data for the 30-minute period between 5:00 p.m. and 5:30 p.m. were used.

It should be noted that although observed trajectories can provide valuable site-specific information on driver behavior that is always preferable, guidance from the Framework can be used to qualitatively validate models even if observed trajectory data are not readily available by using naturalistic and other datasets as a reference.

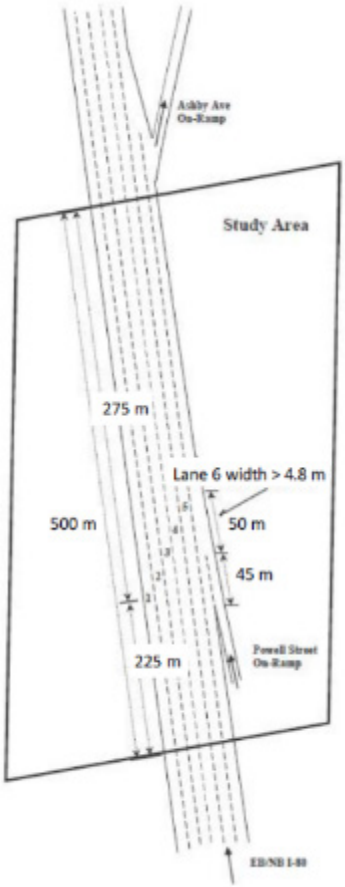


Figure 2-1. Schematic. I-80 corridor segment where Next Generation Simulation data were collected.

(Source: Cambridge Systematics, Inc.)

2.3 Instrumented Vehicle Data from I-80 Site and Vicinity

As part of an earlier task in this project, a new data set was collected from the I-80 NGSIM site and other nearby roadway segments using an instrumented vehicle and several research assistants who drove the vehicle over a period of 20 days. The data collection route followed by each of the drivers is shown in Figure 2-2. Altogether, a total of 118 data collection circuits were completed in April and May 2016, covering a range of different congestion levels, roadway types, and weather conditions.

Figure 2-2. Map. Data collection route and its relationship to the Next Generation Simulation I-80 site.

(Source: Cambridge Systematics, Inc.)

The data were collected using an integrated Real Time Kinematic GPS/IMU sensor that provided highly precise vehicle position data. Two multimode radar units provided wide-angle, midrange and acute-angle, long-range detection for vehicles in front of the instrumented vehicle, behind the instrumented vehicle, and in adjacent lanes on either side. Figure 2-3 provides a summary of the instrumented vehicle equipment.

The results from this instrumented vehicle data set must be interpreted with care, however, as instrumented vehicle datasets are constrained by the number of drivers operating the vehicle and generally are not representative of the full traffic stream or driver population in a given area. Furthermore, a driver that is using the instrumented vehicle may take time to adjust to the new vehicle and may not drive normally due to the awareness that every action is being recorded. However, the data from other vehicles immediately in front of or behind the instrumented vehicle can provide additional insight in that they capture a much wider range of driver types and are associated with subjects that were unaware that their movements were being recorded.

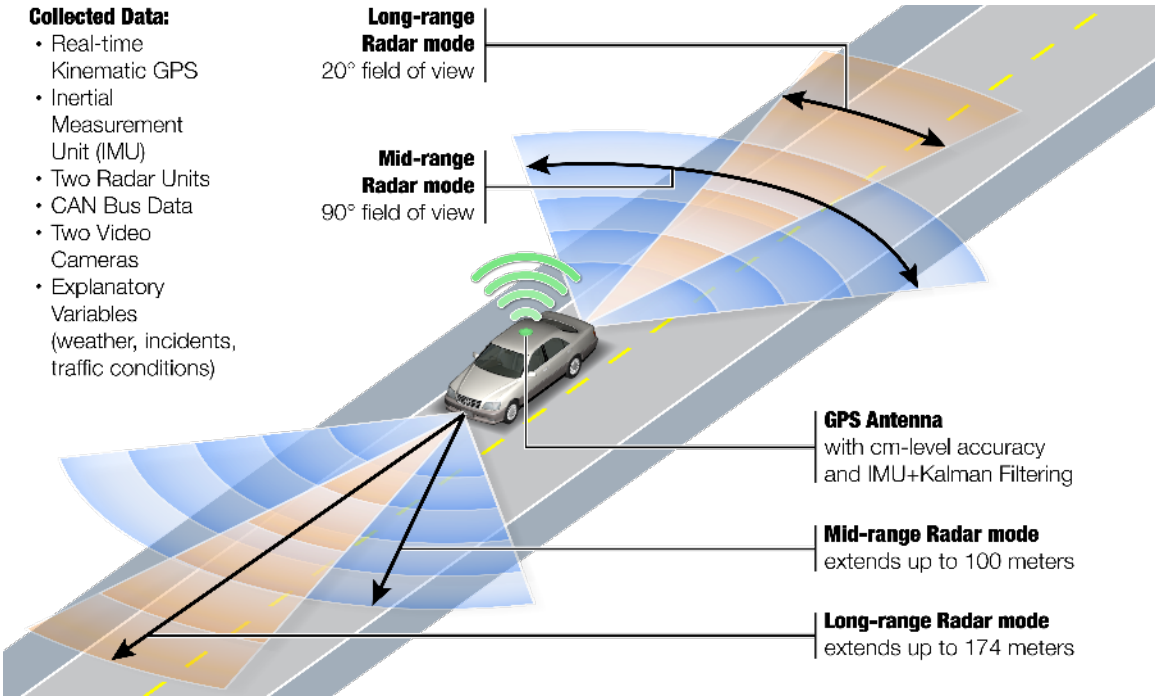


Figure 2-3. Diagram. Data collection equipment installed on the instrumented vehicle.
(Source: Cambridge Systematics, Inc.)

In summary, the instrumented vehicle was used to collect the following data:

- Position data for the instrumented research vehicle, to an accuracy of a few centimeters.
- Position data for the vehicles immediately in front of and behind the instrumented vehicle, when present, to an accuracy of approximately three feet.
- Position data for vehicles in adjacent lanes, when present, to an accuracy of approximately three feet.
- Additional contextual data associated with the data collection circuits, including:
  - Prevailing weather conditions, quantitatively captured using a nearby weather station at Oakland International Airport, and qualitatively described by each driver after each data collection circuit.
  - Traffic conditions, quantitatively captured using several freeway detectors (inductive loops), and qualitatively described by each driver in words after each data collection circuit.
  - Incident data, qualitatively described by each driver after each data collection circuit.

Note that the trajectories collected as part of this project do not reflect the conditions on April 13, 2005 when the NGSIM data were obtained. Therefore, the instrumented vehicle data were not compared directly to the simulated NGSIM trajectories, but were instead used to estimate ranges of values for quantitative and qualitative tests of reasonableness for the simulated data. As mentioned previously, the instrumented vehicle data will be available at the Research Data Exchange (RDE) Web site.

2.4 Naturalistic Driving Study Data

FHWA and many transportation agencies around the world have conducted a number of instrumented vehicle studies in which the sensors are as unobtrusive as possible and the drivers are encouraged to drive as usual or naturally. The primary motivation behind most of these studies is the analysis of factors that contribute to unsafe driving and crashes. A multitude of sensors such as radars and GPS devices are installed on the participant's own vehicle. Alternatively, a participant uses a study vehicle for a number of weeks as their primary car.

Naturalistic studies usually deploy the following sensors at a minimum:

- GPS for latitude and longitude.
- Forward radar.
- Accelerometer and gyroscope.
- Multiple video cameras facing outside the car and inside the car.
- Vehicle network data including brake pedal activation, steer wheel angle, gear position and many other.

In this project, the study team obtained many repeated measurements on I-80 and the surrounding arterials. As a result, the trajectory data collected in this study should not be classified as naturalistic. Even though drivers were encouraged to drive with their own style, the fact that a semi-familiar vehicle was used to drive repetitively the same circuit has likely influenced driver behavior.

In this document, all references to naturalistic data point to locations different to the one simulated on I-80. As a result, naturalistic data are not likely to reflect the conditions of the study corridor and should be used only for insight. Statistical comparisons between simulated and out-of-site Naturalistic data are not recommended unless there is supportive evidence that very similar behaviors were observed.

Chapter 3. Macroscopic Goodness Model Fit

In this chapter, macroscopic goodness-of-fit statistics are presented that compare the performance of the model to the field NGSIM trajectories. The distribution of the vehicle average speeds in the model and data is shown in Figure 3-1. The average speed is computed by dividing the total distance traversed by a vehicle by the time it took the vehicle to cross the corridor in both datasets, and then aggregating across all vehicles. The average vehicle speed in the NGSIM data is 15.2 miles per hour. In the model the corresponding average speed is 14.1 miles per hour, with a relatively small difference of 1.1 miles per hour. Judging by the average speed alone, which is one of the major aggregate calibration metrics, the model would likely be considered as sufficiently calibrated. However, the distributions of the two speeds are different. Fifty percent of model speeds are between 6.8 and 11.0 miles per hour while 50 percent of NGSIM speeds are between 9.4 and 18.6 miles per hour. The median speed in the model is about 8 miles per hour while the median speed in the data is about 12 miles per hour. This indicates that model speeds are generally lower except for about 10 percent of the vehicles that have average speeds close to 40 miles per hour. In comparison, the portion of NGSIM vehicles with speeds higher than 40 miles per hour is less than one percent.

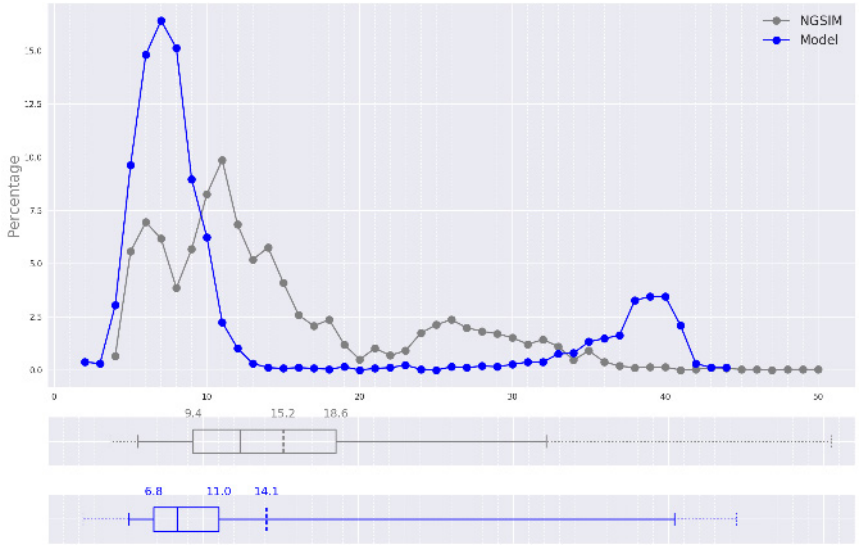


Figure 3-1. Graph. Distribution of average vehicle speed (miles per hour) in the I-80 Next Generation Simulation corridor.

(Source: Cambridge Systematics, Inc.)

The following figures show the speed heat maps for the NGSIM data and for the model. Each lane has been discretized into 50 feet segments. The space mean speed has been calculated for each of the segments using a 10 second interval. The relatively small space and time binning allows for easy identification of shockwaves and other operational phenomena. A color map that

uses seven distinct colors is used to map speeds in the 0 to 70 miles per hour range. Each lane is represented by its own heatmap to better understand lane interactions and conditions in the corridor in general. Lane one is the HOV lane while lane seven is the rightmost lane in the direction of travel.

In Figure 3-2 it is clear that lane one operates at higher speeds than the rest of the lanes. Specifically, speeds on lane one (the HOV lane) range from 20 to 50 miles per hour. In contrast, vehicle speeds on lanes two to seven range from 0 to 30 miles per hour. In lanes two to seven, the red and orange stripes of color that are tilted to the left clearly indicate the progression of shockwaves in this corridor. The slope of the stripes corresponds to the shockwave speed in the fundamental diagram. In both Figure 3-2 and Figure 3-3 white indicates no trajectory information.

Figure 3-3 shows the model heatmap of freeway speeds by time and space. In comparison to Figure 3-2, it is clear that speeds on lane one are higher and speeds on the rest of the lanes are lower than field data. Specifically, most of the time speeds in lanes two to seven in the model are less than 10 miles per hour. Furthermore, shockwaves are much harder to identify compared to field data. Due to an input error, the volume of HOV vehicles of lane one was added to lane two. As a result, the first few hundred feet of lane one do not carry vehicles as it is evident in Figure 3-3. HOV vehicles enter the HOV lane one approximately 400 feet from the beginning of the corridor. Such an error may not be acceptable in an operational simulation model.

Additional information about the macroscopic properties of the model and NGSIM trajectories such as the fundamental diagram is presented in the traffic flow validation section of this report.

Congestion levels can vary from model run to run in simulation models. Trajectories for individual vehicles cannot be averaged. As a result, the tests presented in this report can only be performed on one model run a time. It is therefore recommended that the analyst performs and documents the number of model runs defined in FHWA's Traffic Analysis Toolbox before deciding which run(s) to apply the tests described in this report.

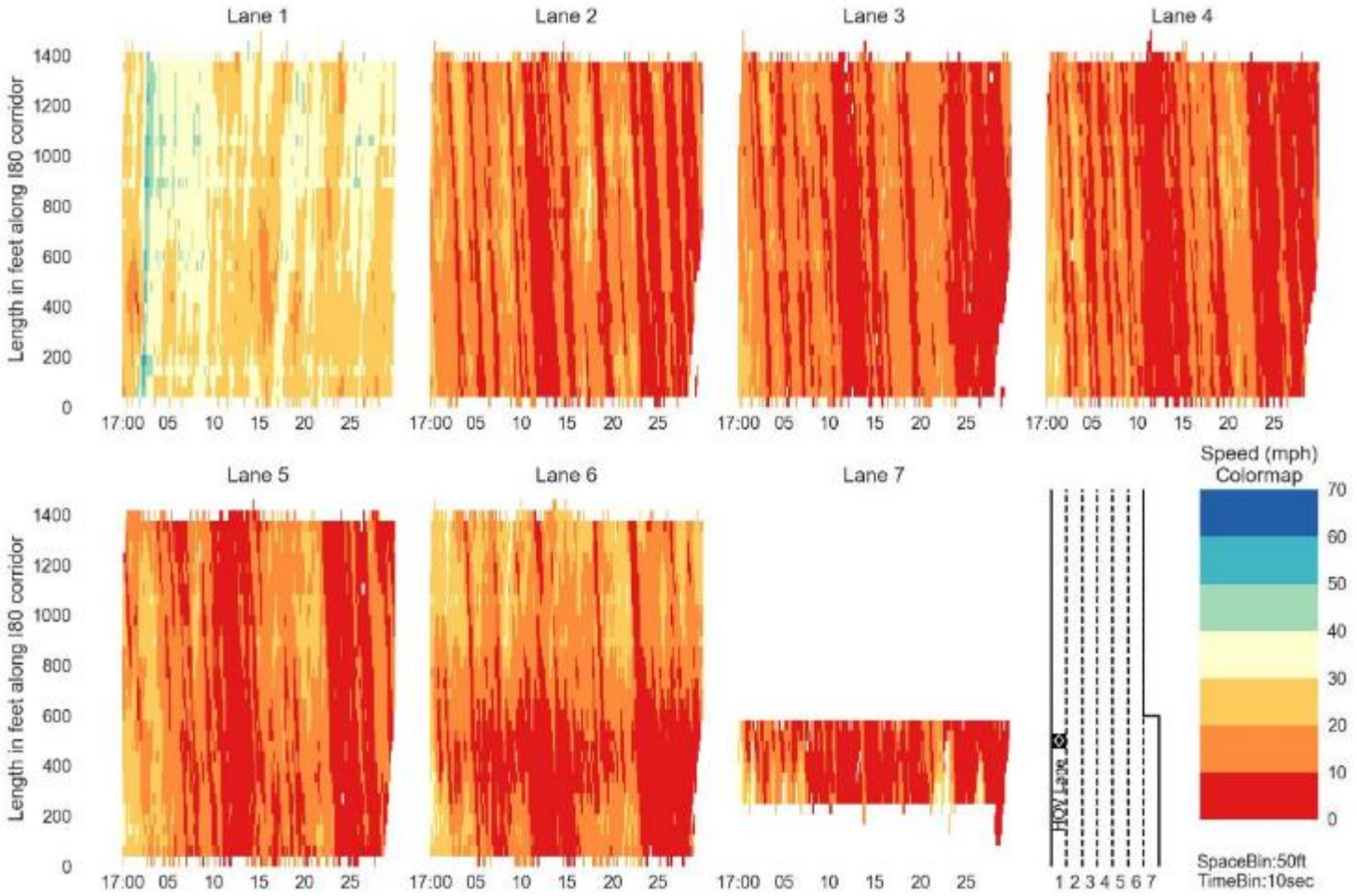


Figure 3-2. Graphs. Next Generation Simulation data speed heatmap.

(Source: Cambridge Systematics, Inc.)

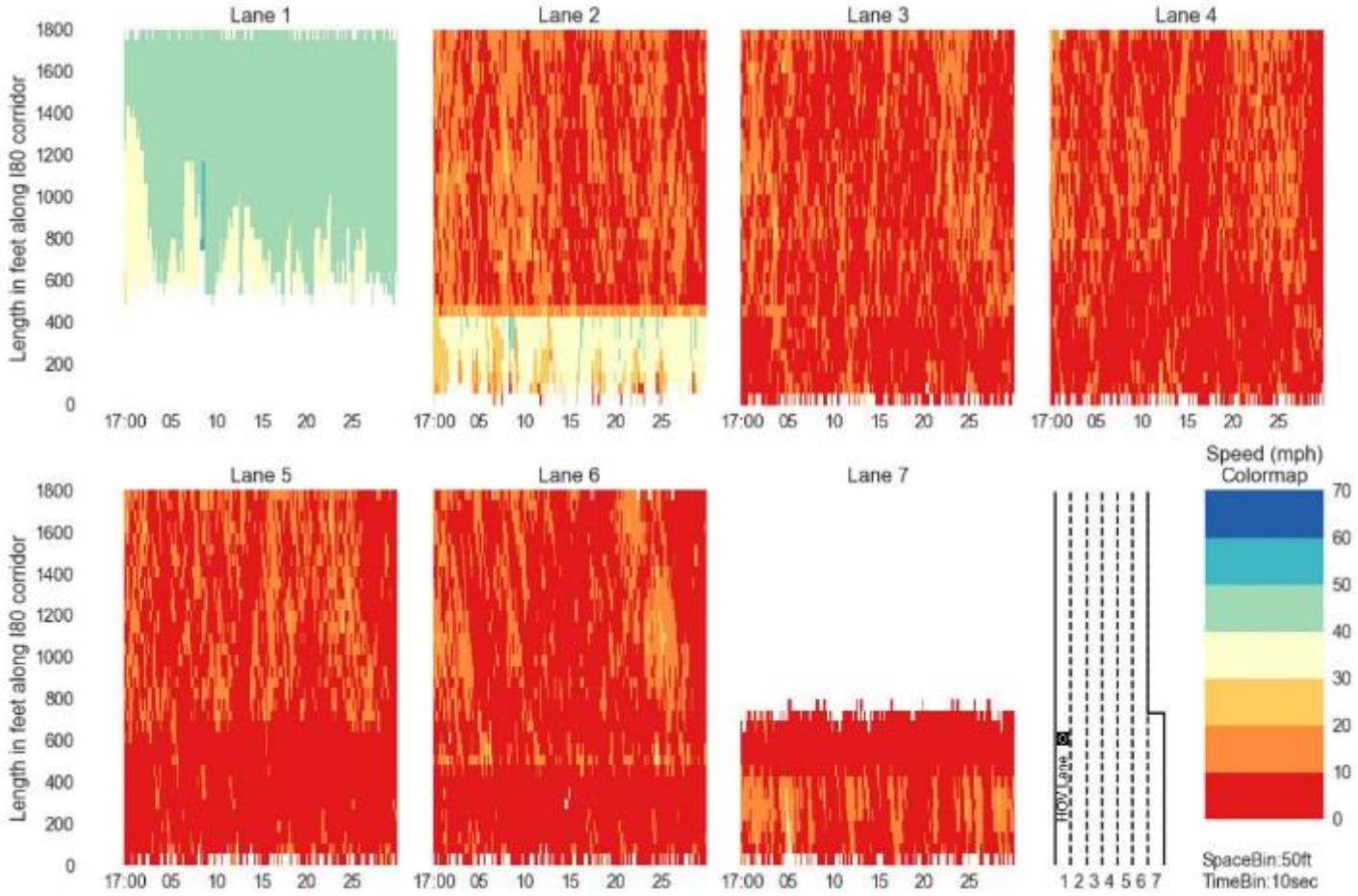


Figure 3-3. Graphs. Model speed heatmap.
(Source: Cambridge Systematics, Inc.)

Chapter 4. Safety Measures and Tests

This chapter focuses on the application of the following four trajectory validation tests related to safety:

- Time to collision (TTC).
- Number of rear-end safety events.
- Lane change severity.
- Lane change urgency.

Additional information about each of the tests is contained in the validation framework report. This section does not test the difference in time gaps since TTC is a better proxy for safety than a time gap and can be computed for both the model and field data.

For all the tests in this chapter, except the number of safety events, a statistical chi-squared test is conducted between the model and field data. The statistical tests indicate that most of the times there are significant differences between the model and the data. Measurement error in the raw NGSIM data can complicate conclusions drawn from the tests. The study team has computed the distribution of some key trajectory measures on another Next Generation Simulation (NGSIM) dataset that pertains to the same corridor but to a different time period (4:00 p.m. to 4:15 p.m.). The raw NGSIM measurements between 4:00 p.m. and 4:15 p.m. have been reconstructed by Vincenzo Punzo to reduce measurement noise. Key statistics from the reconstructed NGSIM dataset are presented in Appendix B. Comparisons between raw NGSIM data from 5:00 p.m. to 5:30 p.m. and the reconstructed NGSIM data from 4:00 p.m. to 4:15 p.m. reveal that in the raw data TTCs less than two seconds are significantly more frequent. In other measures such as the lane change severity and urgency, the results are similar.

Reference naturalistic data are also presented alongside field data for some tests. It should be noted that the reference naturalistic data presented correspond to many days or months of measurements, different traffic conditions, and for both freeways and arterials. As a result, they may show significant differences with the simulated or field data. Additional research is required to establish how trajectory-based behavioral measures change by location or traffic conditions. For example, based on the results in this chapter there is evidence that lane changing aggressiveness may vary by traffic conditions. As a result, it is recommended that naturalistic data are used primarily for insights.

4.1 Time to Collision

Time to collision is a widely used measure in safety applications and collision warning systems. For all the tests in this section TTC values using speeds were calculated in both in the NGSIM and model data at each time step. It is recommended that TTC values are computed using each dataset's native time step. However, due to measurement noise at 100 milliseconds in the raw NGSIM data, an unreasonable number of near vehicle collisions exist in the field data. For this reason, NGSIM vehicle positions, speeds and accelerations, were averaged to 1 second intervals. In all calculations, negative TTCs were set to null values since negative TTCs indicate increasing gaps that cannot lead to collision. As it is mentioned in the framework, TTC values between 0 and up to 15 seconds are of primary interest. The percentage of time in a particular

TTC interval was obtained by dividing the corresponding time spent in the TTC interval by the total vehicle time TTC is between 0 and 14 seconds in this particular test.

Table 4-1 shows the percentage of total time TTC values are within the specific 1-second intervals on the first column. The percentage of time with TTC less than 1 or 2 seconds is expected to be particularly small because it indicates high-collision risk. However, in the up-sampled NGSIM data the percentage of TTC lower than 1 second is an order of magnitude higher than the model data. It is important to note that NGSIM data that have been reconstructed to reduce measurement error (see appendix B) have a significantly lower value for this particular bin that is more in agreement with the reference data. Therefore, the raw field data in this particular test may not be accurate enough to make judgments. Regardless, the statistical test is conducted and the comparison is presented below.

The Chi Square statistical test is used to compare the model versus the field data TTC values. This test can be applied to compare the overall distributions across all TTC bins, or to look at individual TTC bins or ranges. Using an overall test, the conclusion is that the two distributions are different at a 95 percent level of statistical significance (a standard significance level for statistical analysis). Figure 4-1 shows the analysis results by individual bin, where dark gray borders indicate a statistically significant difference. As the figure shows, even when considered individually, every bin exhibited a statistically significant difference between the simulated and observed field data. It should be noted that the distribution of TTC ought to be compared for TTC values less than or equal to 15 seconds. The percentages shown in Table 4-1 pertain to driving time with TTC greater than zero and less than or equal to 14 seconds and not to total driving time. The sum of each column is 100 percent.

Table 4-1. Distribution of time spent at time to collision values between 0 and 14 seconds.

	Percentage of Time	
	Model	**NGSIM**
0 <= TTC < 1	**0.1**	**1.3**
1 <= TTC < 2	**0.4**	**1.7**
2 <= TTC < 3	**1.3**	**2.4**
3 <= TTC < 4	**4.2**	**3.9**
4 <= TTC < 5	**7.7**	**5.5**
5 <= TTC < 6	**9.7**	**7.2**
6 <= TTC < 7	**10.6**	**7.9**
7 <= TTC < 8	**10.7**	**9.0**
8 <= TTC < 9	**10.6**	**9.7**
9 <= TTC < 10	**10.0**	**10.5**
10 <= TTC < 11	**9.5**	**10.1**
11 <= TTC < 12	**8.9**	**10.4**
12 <= TTC < 13	**8.3**	**10.3**
13 <= TTC < 14	**7.8**	**10.0**

Source: Cambridge Systematics, Inc.
Note: Based on Figure 2-1 in the Validation Framework.

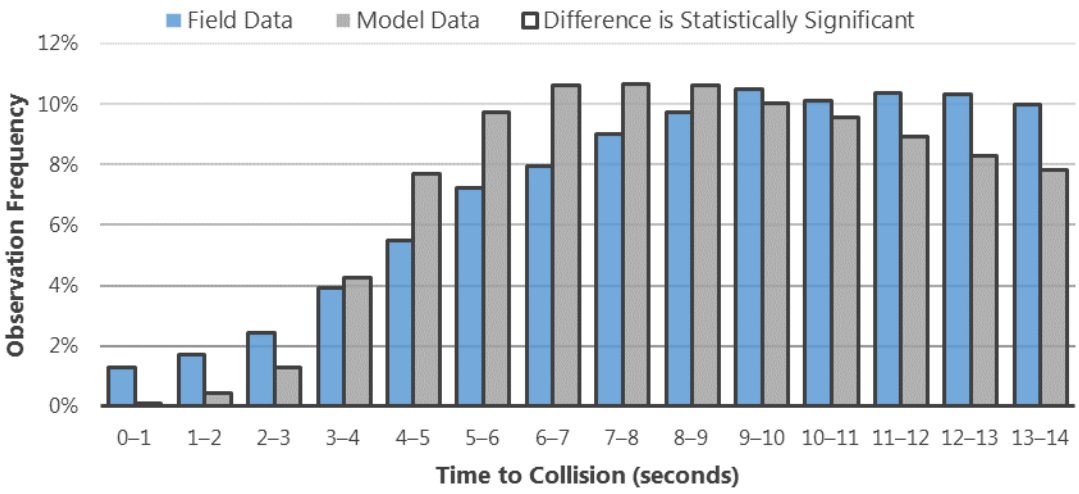


Figure 4-1. Chart. Comparisons of distributions of time to collision.
(Source: Cambridge Systematics, Inc.)

4.2 Number of Rear-End Safety Events

Calculating the number of safety events per vehicle mile provides statistics that are fairly easy to interpret and compare. Safety events in this framework constitute crashes, near crashes, and forward collision warning messages as defined in Table 2-2 of the validation framework. In general, vehicle collisions can be identified by locating negative distance gaps. However, this may not be straightforward when vehicles move at an angle.

In Table 4-2, it is shown that four collisions occur in the model between 5:00 p.m. and 5:30 p.m. In addition to the four crashes, 2,585 near-crash situations were found. These correspond to cases where deceleration is greater than 1g (32 fpss) or to cases where TTC is less than 2 seconds and deceleration is greater than 0.5g (16 fpss) at the same time step. Finally, 2,377 cases were found that correspond to a forward collision warning with TTC less than 2.4 seconds. Near-crash and collision warning events happen in the simulation approximately once every half a vehicle mile.

In Table 4-2, the number of safety events in the model has been converted to average rates per million vehicle miles traveled and are compared with guidance from Table 2-4 in the validation framework. Raw NGSIM data are not included in Table 4-2 because measurement noise results in a significant number of collisions and unrealistic number of safety events. Chi squared can be applied to test the statistical significance of the results. However, doing so is unnecessary since the numbers are orders of magnitude different in Table 4-2. It should be noted that the virtual collisions that occur in some simulation models may be an artifact of algorithm imperfections and not necessarily an accurate representation of collision risk. A simulator that does not explicitly model collision risk or collision frequency should not be used in assessing the impact of technologies that prevent collisions.

Table 4-2. Safety events.

Safety Event	Model Count	Average Rate per Million Vehicle Miles Traveled	
		Model	Naturalistic Data (Different Location)
Crashes	4	52,980	9
Near-Crashes	2,585	34,238,325	214
Forward Collision Warning	2,377	31,483,443	

Source: Cambridge Systematics, Inc.

4.3 Lane Change Severity

Lane change severity (LCS) quantifies the risk of a lane-changing maneuver on the future follower vehicle on the adjacent lane that may need to brake abruptly. Lane change severity is classified into one of seven levels as described in Table 2-5 of the Validation Framework report. Lane change severity classifications were performed using each dataset's native time step. It should be noted that measurement error in the NGSIM data may have influenced the LCS category classifications. However, the same calculations on the reconstructed NGSIM dataset presented in Appendix B show more similarity to the modeled data, a fact that indicates that LCS may be less sensitive to measurement noise compared for example to TTC or other measures in the validation framework.

In Table 4-3, the percentage of lane changes for each of the several LCS categories is provided for the model, NGSIM, and reference naturalistic driving study data. Model and NGSIM percentages follow a completely different pattern from the naturalistic data that indicate that 95 percent of the lane changes entail minimal interference from other vehicles. In the model and field data less than 10 percent of lane changes have the Principal Other Vehicle (POV) more than 5 seconds away as opposed to 95 percent in the naturalistic data. In both the model and the field NGSIM data more than 80 percent of lane changes are category 3 and higher. Even though none of these categories correspond to a safety event they do indicate a higher interference and possibly a higher cognitive load on the driver's part. As it was stated in the beginning of this section, the reference naturalistic data included in Table 4-3 correspond to entire trips from home to work and not to heavily congested weaving sections. Additional research may be required to identify how LCS depends on traffic conditions and other factors.

Table 4-3. Lane change severity validation test.

Category	Description	Percentage		
		Model	NGSIM	Naturalistic
1	Principle other vehicle (POV) in the fast approach zone with time to reach closest end of zone, Tr, > 5.0 sec, including case where there is no vehicle in the adjacent lane.	10.0	4.6	95.1
2	POV in the fast approach zone with time to reach closest end of zone in the range 3.0 < Tr < 5.0 sec.	12.8	7.2	1.2
3	POV in the fast approach zone with time to reach closest end of zone in the range 1.0 < Tr < 3.0 sec.	39.6	37.2	0.2
4	POV in the fast approach zone with time to reach closest end of zone, Tr., < 1.0 sec.	21.1	32.2	<0.1
5	POV in the proximity zone.	16.5	18.8	3.4
6	A near-miss emergency action or unplanned sudden maneuver is required to avoid a collision with a vehicle in the adjacent lane into which the driver of the SV was attempting to move.	0	0	0.1
7	A near-miss emergency action or unplanned sudden maneuver is required to avoid a collision with a vehicle in the adjacent lane into which the driver of the SV was attempting to move.	0	0	0

Source: Cambridge Systematics, Inc.

Note: Data from the "Naturalistic" column are copied from Table 2-6 of the Validation Framework. The data do not pertain to the I-80 corridor.

A statistical analysis of the Table 4-3 percentages can be performed using a Chi Square test to compare the frequency distribution NGSIM and model data. The Chi Square test can be applied to compare the overall distributions across all severity categories, or to look at individual ones separately (i.e., by considering two bins: the type of interest, and all other types combined). Using an overall test, the conclusion is that the two distributions are different at a 95 percent level of statistical significance (equivalently the probability of error is less than 5 percent). Figure 4-2 shows the analysis results by individual bin, where dark gray borders indicate a statistically significant difference. As the figure shows, when evaluated individually, the percentage of lane changes in Severity Category 3 are consistent across the two data sets, while the percentages for Severity Categories 1, 2, and 4 are different on a statistically significant level.

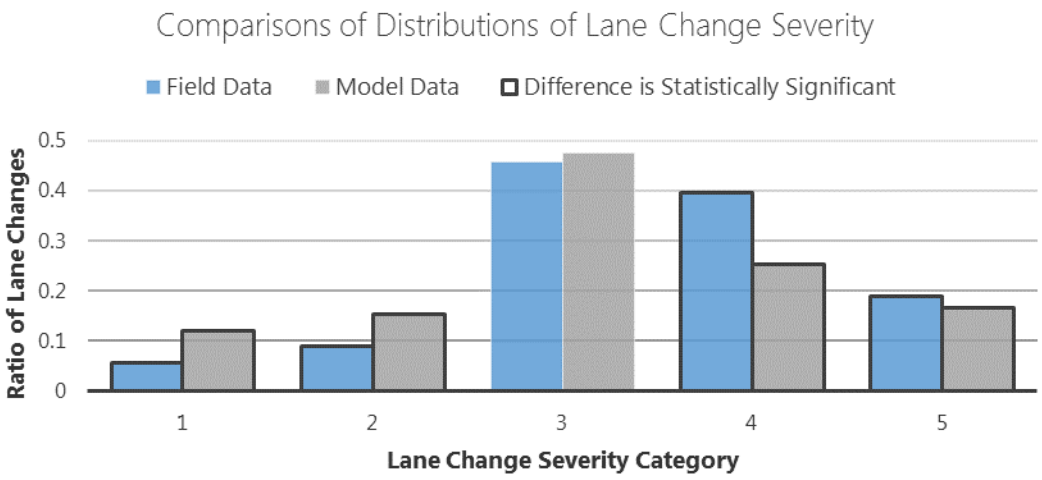


Figure 4-2. Chart. Comparisons of distributions of lane change severity.
(Source: Cambridge Systematics, Inc.)

4.4 Lane Change Urgency

Lane change urgency captures the risk associated with the vehicle executing the lane-changing maneuver. Lane change severity is classified into one of four levels as described in Table 2-6 in the Validation Framework report. As it can be seen from Table 4-4, 69 percent of the lane changes in the model are classified as non-urgent compared to 95 percent of lane changes in naturalistic data that combine freeway and arterial driving. The corresponding percentage in the NGSIM data is 83 percent. Clearly, drivers change lanes more aggressively in the model compared to NGSIM or naturalistic data. The fact that NGSIM field data deviate from general naturalistic averages pertaining to other locations and a mix of traffic conditions probably indicates that LCU is site and congestion dependent. Based on Table 4-4 there are more aggressive lane changes in the model compared to field data.

Table 4-4. Lane change urgency test.

Category	Description	Percentage		
		Model	Field Data (NGSIM)	Naturalistic (Multi-Area)
1	Nonurgent: TTC > 5.5 s to any vehicle but the future follower.	69.0	83.3	95.1
2	Urgent: 3 s < TTC < 5.5 s to any vehicle but the future follower.	13.4	4.0	3.9
3	Forced: TTC < 3 s to any vehicle but the future follower.	17.6	12.7	0.3
4	Critical incident/crash: physical contact occurs or a sudden near-miss maneuver is required to avoid collision.	0?	0?	0?

Source: Cambridge Systematics, Inc.
Note: Data from the "Naturalistic" column is copied from Table 2-6 of the Validation Framework.

A statistical analysis of these data can be performed using a Chi Square test to compare the frequency distribution of lane change urgency data between the NGSIM field data set and the reference simulation model data. This test can be applied to compare the overall distributions across all urgency categories, or to look at individual ones separately (i.e., by considering two bins). Using the overall test, the conclusion is that the two distributions are different at a 95 percent level of statistical significance. Figure 4-3 shows the analysis results by individual bin, where dark gray borders indicate a statistically significant difference. As the figure shows, even when considered individually, every bin exhibited a statistically significant difference between the simulated and observed field data.

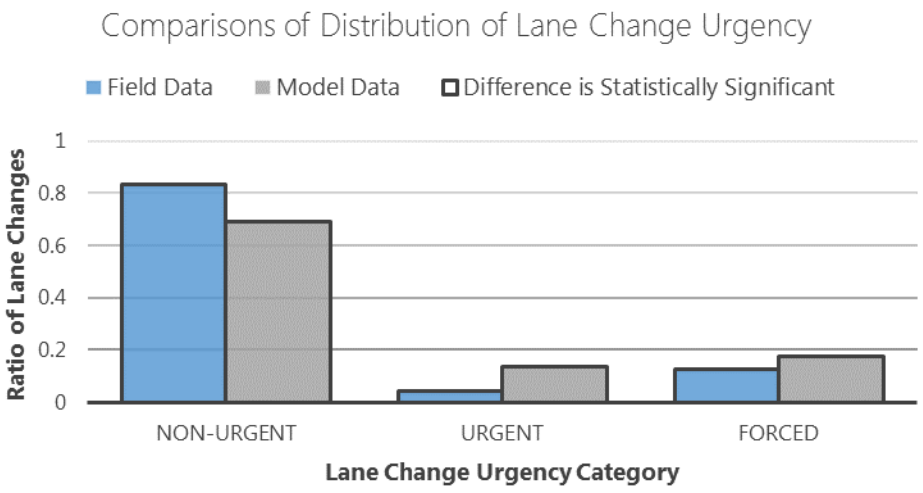


Figure 4-3. Chart. Comparisons of distributions of lane change urgency.
(Source: Cambridge Systematics, Inc.)

Chapter 5. Vehicle Limitations and Driver Comfort Validation Measures and Tests

This chapter applies the vehicle limitations and driver comfort tests described in the validation framework. The purpose of these tests is to identify cases where the modeled vehicle has physically infeasible acceleration or jerk values. The tests also provide insights on the level of driver comfort using acceleration, jerk and acceleration root mean square error (ARMS) tests. Additional background and detail about each of these tests is available in chapter three of the Validation Framework report. Validation measures in this category include:

- Acceleration Range (AR) Thresholds for Acceleration and Deceleration.
- Acceleration Jerk (AK).
- Acceleration Root Mean Squared (ARMS).

5.1 Acceleration Range Thresholds

Acceleration is formally defined as the second time derivative of position, or the rate of change in speed. The specific way speed and acceleration are calculated from positional data can have a significant impact in the number of outliers. For this reason, it is recommended that the analyst calculates speed and acceleration in both the modeled and field data using the same equations. Equations one and two in Appendix B of the Validation Framework or a valid alternative definition can be used.

In the vast majority of time, drivers use a fraction of the acceleration and deceleration capabilities of their vehicle for a variety of reasons that include safety, comfort, and fuel consumption among others. Deceleration close to 1g (acceleration less than -1g) corresponds approximately to the maximum deceleration capability of current vehicles and signifies a safety event in which the driver had to brake as hard as possible to avoid a collision. As Table 5-1 shows there are thousands of such cases where vehicle deceleration exceeds 1g in the simulation. Also, in the simulation there also are more than 100 cases with deceleration greater than 2g which is mechanically impossible with the current vehicle technology. The number of times vehicle deceleration in the model is greater than 1g is more than twice per mile (Table 5-2). In contrast, there are no such events in the instrumented vehicle field data collected for this project. Also, the number of times vehicle deceleration is between 0.5 and 1.0g is about two per vehicle mile in the model. The instrumented vehicle data contain only 14 such events on the freeway that correspond to 0.03 events per vehicle mile, a significant difference. On the San Pablo arterial, vehicle deceleration exceeded 0.5g much more frequently in the field data. On the arterial, the instrumented vehicle deceleration becomes greater than 0.5g approximately 0.6 times per vehicle mile.

Table 5-1. Deceleration limits (model).

Deceleration Range	Frequency (Model)	Comment
4g to 3g	3	Infeasible.
3g to 2g	133	Infeasible.
2g to 1g	2,447	Borderline infeasible, safety event.
1g to 0.5g	2,400	Uncomfortable, probably a safety event.

Source: Cambridge Systematics, Inc.

Table 5-2. Deceleration range events.

	Number per vehicle mile	
Deceleration Range	**Model**	**Instrumented Field Data**
More than 2g	0.12	None
2g to 1g	2.13	None
1g to 0.6g	2.08	0.03

Source: Cambridge Systematics, Inc.

In Figure 5-1 the acceleration distributions for field and modeled data are plotted together to provide a general picture. However, it is recommended that acceleration distributions are compared at different speed levels and not for all speed levels combined. To remove some of the noise in the raw field data, average acceleration values for every second are shown in Figure 5-1. The distribution of acceleration can vary by speed considerably based on the results presented in the validation framework. Therefore, even though the distribution of acceleration can be used as a comparison test, it may fail to reveal the following two important properties of acceleration, including:

- Outlier values that are close to or exceed the physical limits of current vehicles.
- The distribution of acceleration by speed level.

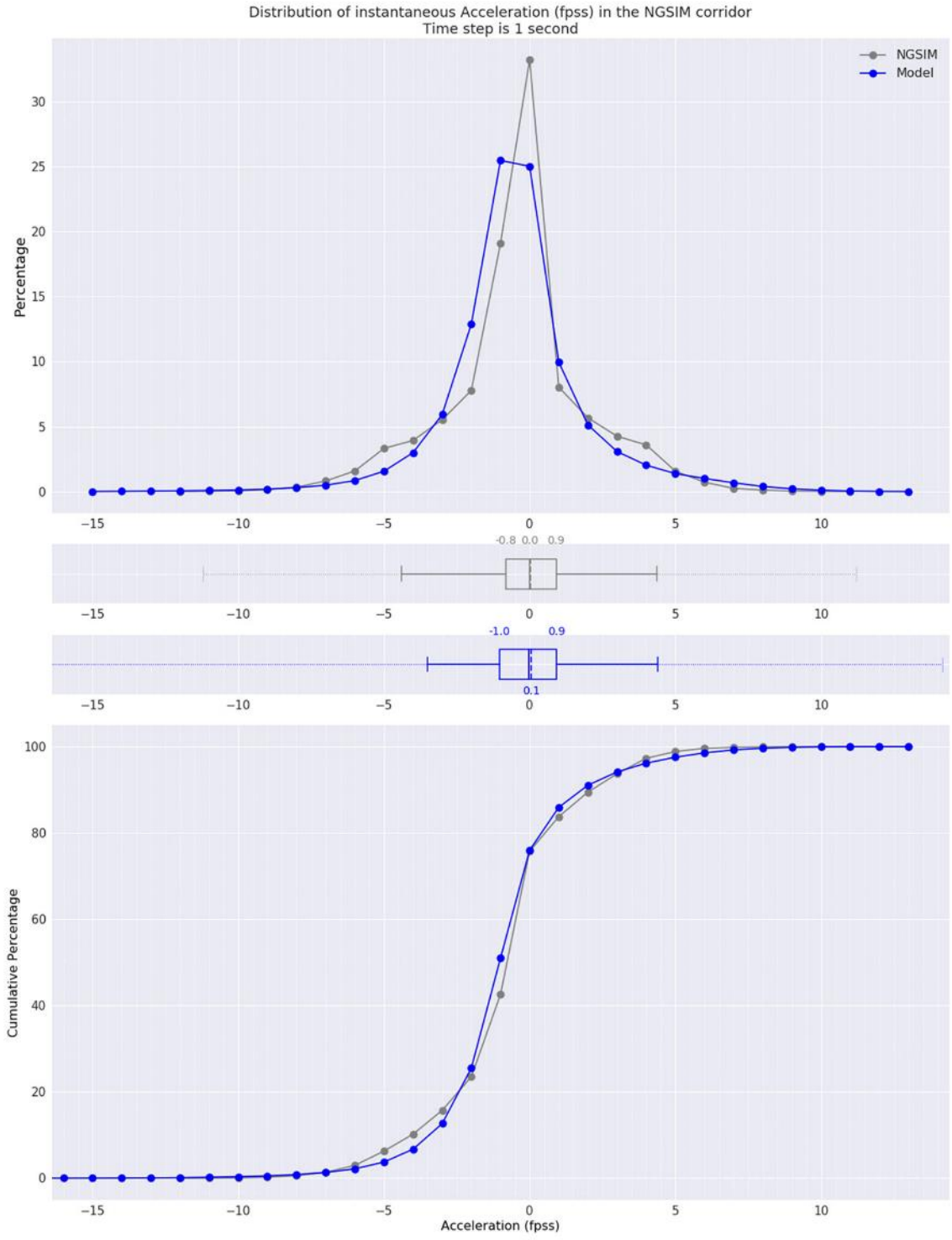


Figure 5-1. Graphs. Distribution of instantaneous acceleration (model and data).
(Source: Cambridge Systematics, Inc.)

The distribution of acceleration by speed is shown using different charts that highlight different elements of the comparison. Figure 5-2 is good at showing outliers. It shows a density scatterplot of speed versus acceleration that uses a bin of 1 mile per hour on the x-axis and a bin of 1 fpss on the y-axis. In this scatterplot all the data points are shown, including the outliers that may occur only once. This is achieved by using a logarithmic color scale that maps zero points to white, one or a few points to light green and tens of thousands of points to black. For example, it can be seen from Figure 5-2 that there are 10 or more cases of speed close to 0 and deceleration close to 1g. Decelerations greater than one g that are quantified in Table 5-1 also can be shown in this chart but have been omitted by limiting the y-axis of the figure to values greater than -1g. This was done to maintain a graphic resolution that devotes a large portion of the graph to the largest portion of the data. It also is notable in Figure 5-2 that the maximum acceleration available at low speeds is rather low, significantly less than 14 fpss which is the maximum acceleration in the modeled data. In fact, the maximum acceleration observed in the model is achieved at around 20 miles per hour which is approximately one third of the maximum speed. Observed data, in contrast, show that the maximum acceleration is achieved at low speeds (less than 10 miles per hour) and decreases by speed (figures 3-1 and 3-3 in the validation framework).

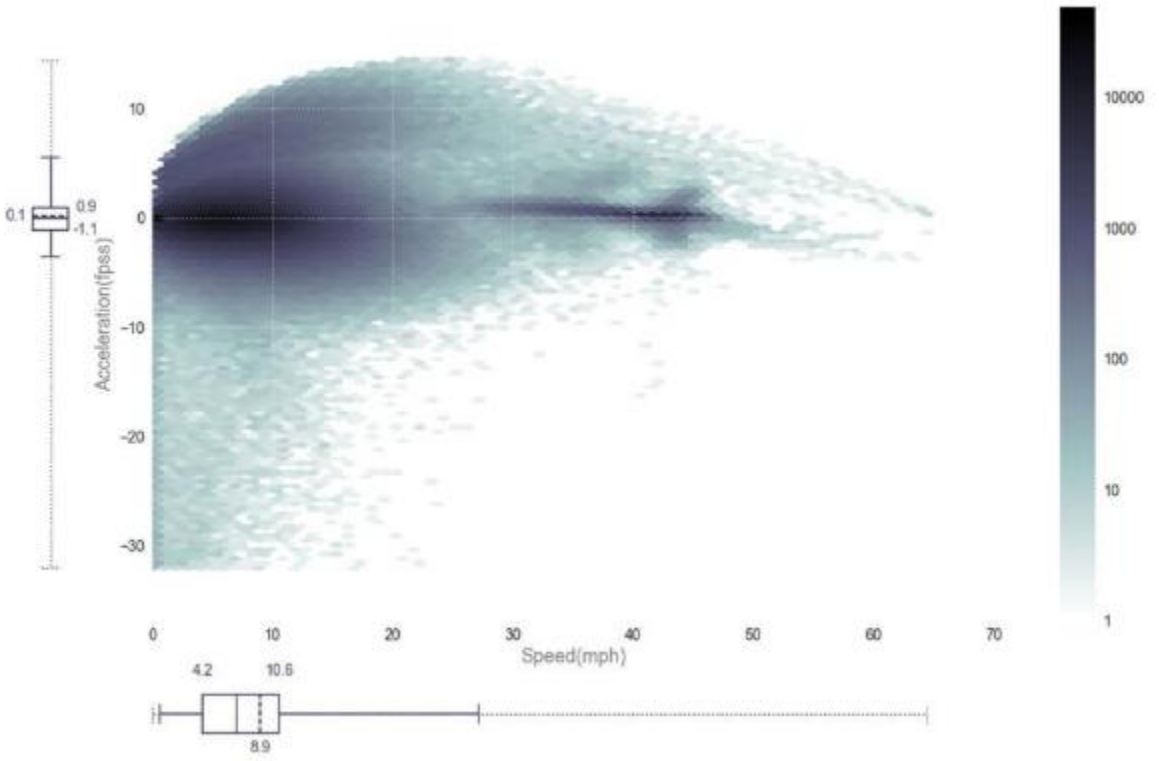


Figure 5-2. Graph. Two-dimensional instantaneous speed versus acceleration in the model.
(Source: Cambridge Systematics, Inc.)

The same data presented in Figure 5-2 also are shown in Figure 5-3 and Figure 5-5. In Figure 5-3 acceleration percentiles have been computed using a 1 mile per hour speed step. This figure shows that at low speeds the bandwidth of acceleration is narrower to the acceleration bandwidth at speeds closer to 20 miles per hour when the maximum of acceleration is achieved based on Figure 5-2. It also should be noted that the percentile acceleration bandwidths for speeds higher than 45 miles per hour are based on a relatively low number of data points and as a result they should not be given undue weight. Additional scenarios with varying levels of congestion are

recommended to investigate the speed versus acceleration relationship. For comparison purposes the x-scale in Figure 5-2, Figure 5-3, and Figure 5-5 is the same and equal to 0 to 70 miles per hour.

Comparing the percentile values at different speeds can provide significant insights on the level of acceleration variability at different speed levels. Field data shown in Figure 5-4 (same as Figure 3-3 of the validation framework) show that acceleration variability reduces by speed. Specifically, while the interquartile range of acceleration remains constant by speed the 95th and 99th percentiles of acceleration tend to reduce as speed increases. In addition, the 99th percentile of acceleration in Figure 5-4 of the framework has a maximum range of plus six to minus seven fpss. In contrast, the model data have a maximum 95th percentile value of 13 fpss approximately twice as large as it is observed. Based on Figure 5-3 vehicles have an uncomfortable or close to unfeasible deceleration one percent of the time when speed is less than one mile per hour, an unrealistic model behavior. In Figure 5-3, the variation of acceleration becomes smallest at around 42 miles per hour and widens again between 50 and 60 miles per hour. This behavior may not be realistic as well. However, such a behavior may be the result of insufficient data for speeds greater than 40 miles per hour. Multiple scenarios that simulate a wider range of traffic phenomena may be needed to obtain a complete understanding of model behavior under different traffic conditions.

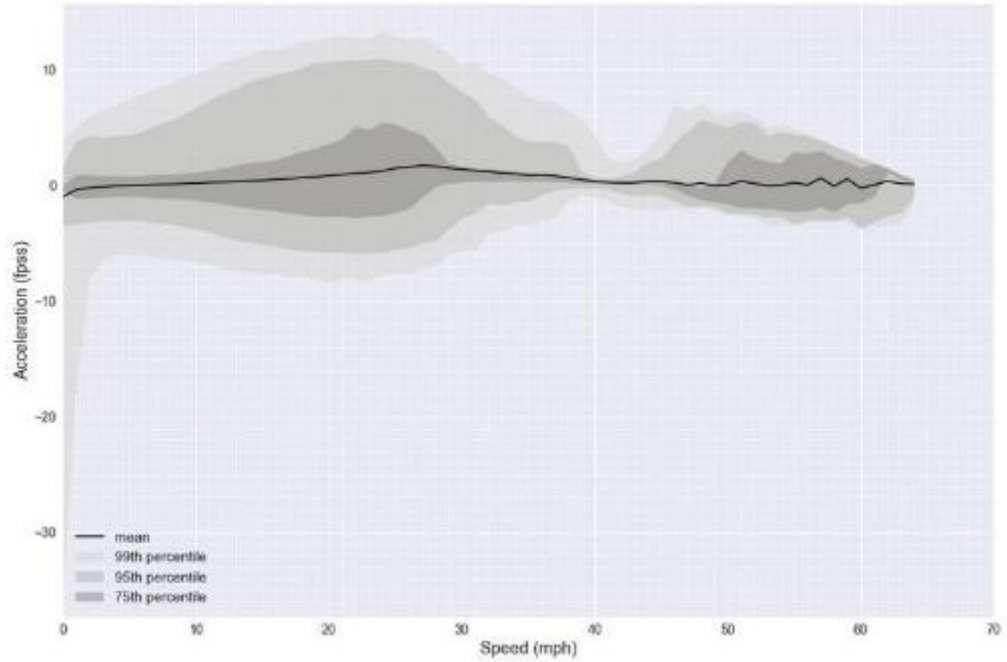


Figure 5-3. Graph. Percentile acceleration distribution by speed (model).
(Source: Cambridge Systematics, Inc.)

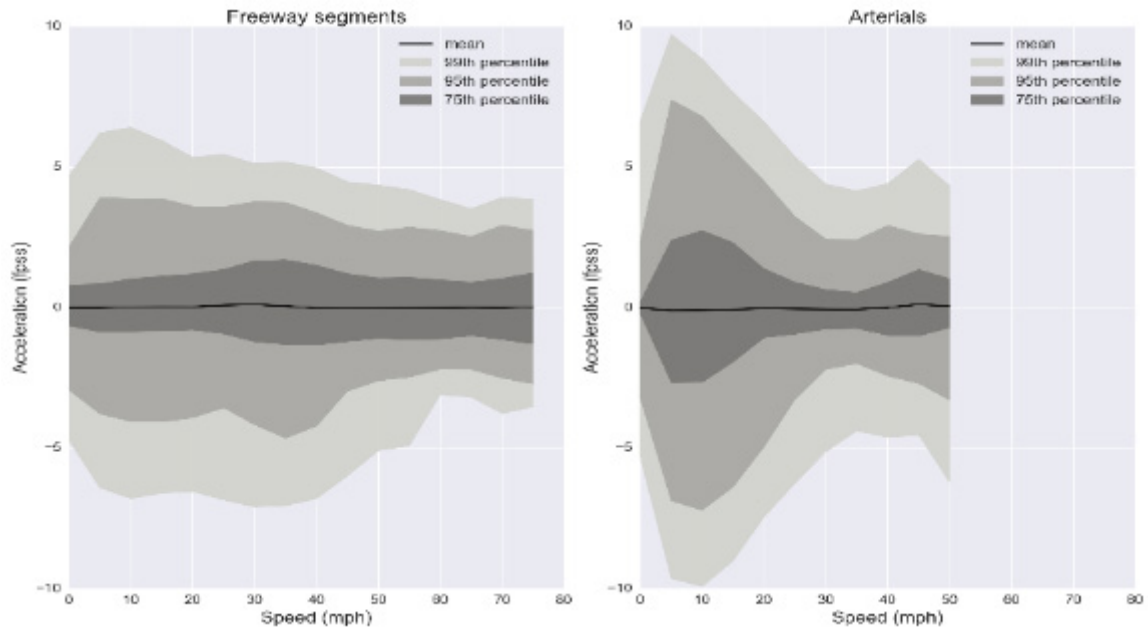


Figure 5-4. Graphs. Distribution of speed versus acceleration in the collected instrumented vehicle data (Same as Figure 3-3 in the validation framework).

(Source: Cambridge Systematics, Inc.)

Figure 5-5 shows the same data as Figure 5-2 and Figure 5-3 but uses a color map with a linear scale instead of the logarithmic one used in Figure 5-2. Speed and acceleration bins are one mile per hour by one fpss. The color of each bin is based on the probability of trajectory points that fall into the bin. Outliers, due to their very low probability cannot be seen in Figure 5-5. For example, max acceleration and deceleration or the limits of the speed versus acceleration distribution which were clearly distinguishable in Figure 5-2 are not visible. This visual, however, can quantitatively describe the distribution of the majority of the data. It can be seen for example that approximately 3.5 percent of the time vehicles have a speed between zero and one mile per hour and acceleration between zero and one fpss.

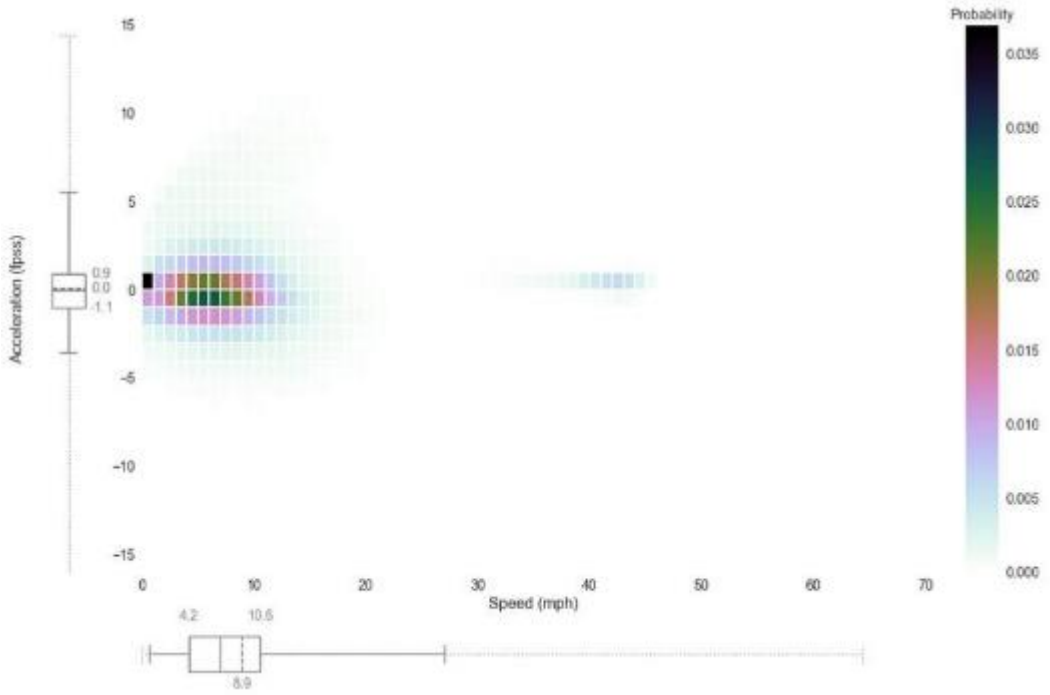


Figure 5-5. Graph. Two-dimensional instantaneous speed versus acceleration histogram (model).

(Source: Cambridge Systematics, Inc.)

Even though no statistical tests were applicable in this section, it is clear that:

- Infeasible decelerations greater than 2g exist in the modeled data at the rate of 0.12 per vehicle mile.
- Events with deceleration greater than 1g occur approximately twice per mile in the modeled data. No such event was recorded in the field data.
- A high number of uncomfortable decelerations between 0.5 and 1g that also may be a safety events exist in the modeled data. The frequency of such events is approximately two per vehicle mile and it is significantly higher than what has been observed in the field data.
- The distribution of acceleration by speed is qualitatively different between the model and the field data. Such a difference may result in emission estimates that are not realistic. In particular:
    1. For speeds less than 20 miles per hour, the maximum possible acceleration increases by speed instead of decreasing. This is in contrast with the mechanical capabilities of vehicles and the field data.
    2. The percentile distributions of acceleration versus speed are significantly different between the modeled and field data. Also, the bandwidth between the percentile curves is significantly higher in the modeled data signifying increased speed variability that also is captured in the ARMS test.

5.2 Acceleration Jerk

Acceleration jerk is a measure of the rate of change in acceleration over time. Acceleration jerk is used extensively in railway transport, elevator design and many other fields to measure human comfort under motion. Jerk can be used both as a measure of driver comfort and as a check for mechanically infeasible behavior. To calculate acceleration jerk, average acceleration values were calculated by second in both the modeled and NGSIM data. Jerk was calculated as the difference in acceleration in two subsequent time steps.

Table 5-3 shows the number and percentage of acceleration jerk values in different ranges. Of particular importance are jerk values greater than 50 f/s^3 (feet per second cubed) and less than -50 f/s^3 which are infeasible because they require acceleration and decelerations that are significantly greater than 1g. Jerk values greater than 15 f/s^3 in absolute magnitude designate significant driver discomfort. Based on Table 5-3 the model has more values greater than 15 f/s^3 in absolute terms than the field data.

The distribution of jerk in the field data and in the model are shown in Figure 5-6. Field data have a higher percentage of jerk values that are greater than 3 in absolute terms. Both the interquartile range and the 5^{th} and 95^{th} percentiles show that the jerk in the field data is more dispersed. However, model data have a higher number of outliers, jerk values greater than 15 f/s^3 in absolute terms.

Based on the analysis above:

- The model does not have infeasible jerk values.
- In terms of driver comfort, the relative frequencies in Table 5-3 shows that there are more uncomfortable events in the model compared to the field data even though the field data are more dispersed (Figure 5-6).

Table 5-3. Acceleration jerk values.

	Field Data (NGSIM)		Model	
Range(f/s^3)	**Frequency**	**Percentage**	**Frequency**	**Percentage**
Less than -50	0	0	0	0.0
-50 to -15	94	0.0	321	0.1
-15 to -3	52,903	16.4	26,086	8.8
-3 to 3	215,272	66.8	243,483	81.9
3 to 15	53,705	16.7	27,335	9.2
15 to 50	92	0.0	131	0.0
Greater than 50	0	0.0	0	0.0

Source: Cambridge Systematics, Inc.

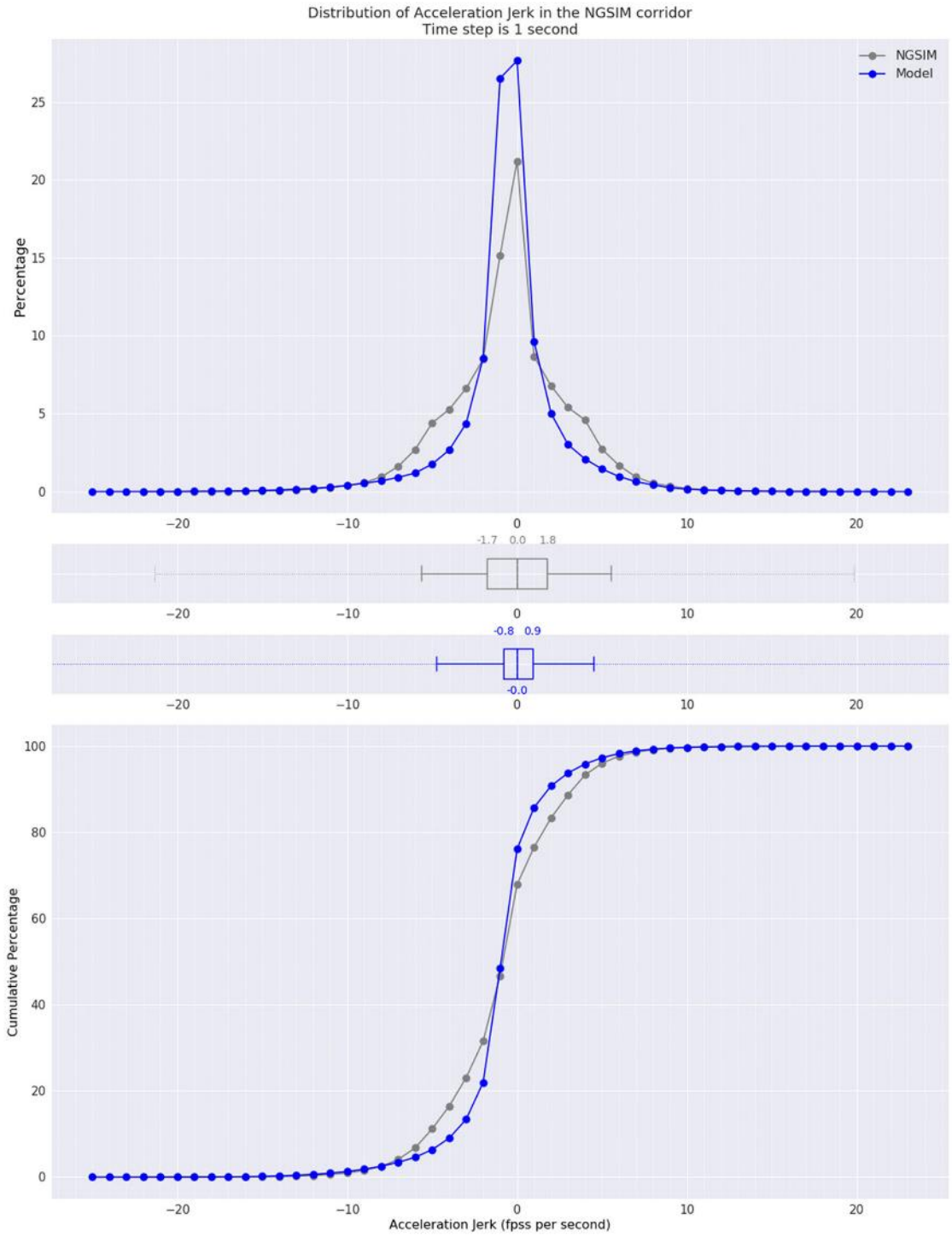


Figure 5-6. Graphs. Distribution of acceleration jerk (model and data).
(Source: Cambridge Systematics, Inc.)

5.3 Acceleration Root Mean Squared

ARMS captures the variability or fluctuations of acceleration. ARMS is associated with experienced levels of driver comfort rather than the physical capabilities of the vehicles involved. With some caveats explained in the validation framework, ARMS can be used to classify a vehicle ride to different levels of driver comfort based on ARMS thresholds presented in a relevant ISO standard. The ARMS calculation involves squaring all acceleration values in a given data set and speed bin, calculating the mean of these squared values, and computing the square root of the mean.

Table 5-4 lists ARMS values calculated by different speed bin. This is because ARMS may vary considerably by speed. Calculations were based on native time step of each dataset; resampling was not required because ARMS is insensitive to time steps less than one second based on project findings. In addition to the ARMS value for each speed category, the number of data points for the same category also are included in Table 5-4 for reference. Values greater than 45 miles per hour have not been included since not all datasets contained a sufficient number of observations.

Note also that the ARMS is dependent upon speed and roadway type. Therefore, care must be taken to ensure that the simulation trajectory dataset is compared to an appropriate reference dataset (e.g., one with similar speed characteristics and from a comparable facility type). Depending on the extent of the simulation model, it may be appropriate to perform this test multiple times for different facility types and speed conditions in the simulation.

Based on Figure 5-7 the following conclusions can be drawn:

- At most speed levels, the modeled data have a significantly higher ARMS compared to the instrumented vehicle data which are the most accurate I-80 data collected.
- Modeled data have a lower value of ARMS at around 40 miles per hour compared to instrumented vehicle data. This is because in the model there are vehicles traveling on the leftmost lane of the freeway based on the speed heat maps presented above. Those vehicles seem to cruise at around 40 miles per hour without significant acceleration fluctuations. This also is evident in Figure 5-3 where the bandwidth of the percentile curves becomes the smallest observed.
- Reconstructed NGSIM data have a lower ARMS than raw NGSIM data at all speed levels possibly due to the reduction of measurement error. However, both datasets have significantly higher ARMS values than the instrumented vehicle data.
- Field data regardless of source, have ARMS values that are either constant or moderately increasing by speed. In contrast, modeled trajectories have an ARMS profile that varies significant by speed. In particular, modeled ARMS is the highest around 20 miles per hour which is the speed maximum that acceleration is achieved.

Table 5-4. Acceleration root mean square values (model and field data).

Speed (mph)	Reconstructed NGSIM (4 to 4:15 p.m.)		NGSIM (5 to 5:30 p.m.)		Instrumented Vehicle on Freeway (3 to 6 p.m.)		Model (5 to 5:30 p.m.)	
	ARMS	n	ARMS	n	ARMS	n	ARMS	n
0-5	0.79	67,733	1.02	455,584	0.45	271,481	1.23	473,613
5- 10	1.01	118,339	1.25	583,204	0.63	310,568	0.84	601,255
10 – 15	0.90	266,193	1.32	560,678	0.66	233,963	1.05	258,667
15 – 20	0.95	285,409	1.37	501,341	0.63	161,972	1.37	70,397
20 – 25	0.94	185,552	1.39	408,505	0.60	109,558	1.61	17,823
25 – 30	1.15	52,241	1.43	274,290	0.57	78,489	1.44	6,998
30 – 35	1.09	28,014	1.50	171,050	0.57	50,246	0.79	9,907
35 – 40	1.14	22,468	1.56	107,616	0.55	30,812	0.47	17,700
40 – 45	1.13	15,321	1.64	79,615	0.58	17,311	0.24	40,355

Source: Cambridge Systematics, Inc.

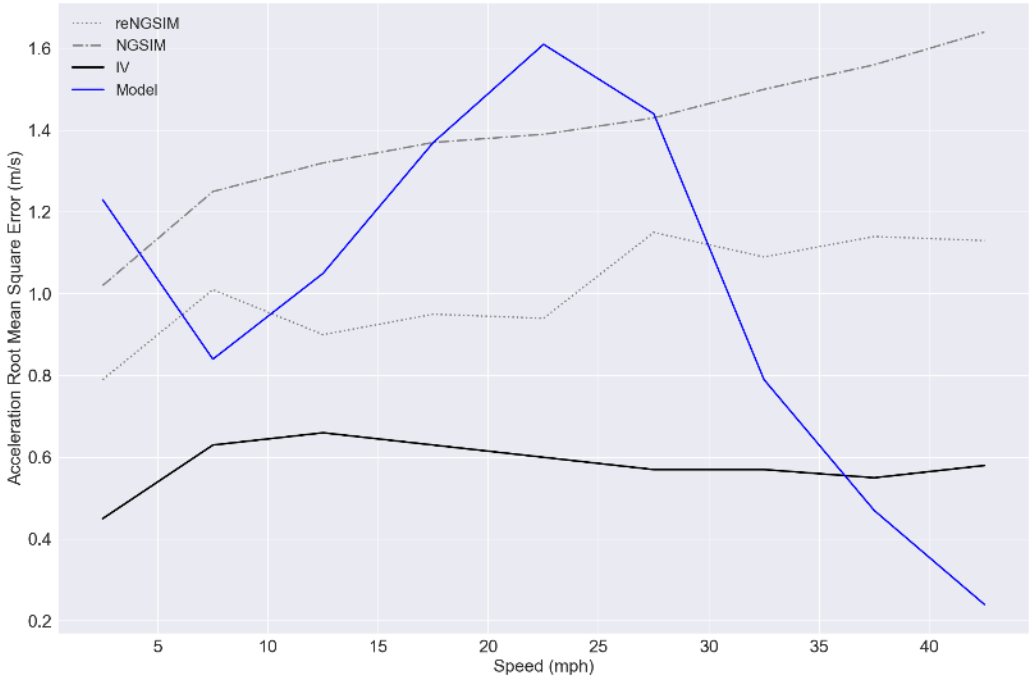


Figure 5-7. Graph. Distribution of acceleration root mean square by speed (model and data).
(Source: Cambridge Systematics, Inc.)

Chapter 6. Traffic Flow Validation Measures and Tests

This chapter focuses on the application of the following validation measures and tests related to traffic flow and its properties:

- Number of lane changes per driver.
- Lane change rate.
- The fundamental diagram relating flow and density.

Additional background and detail about each of these tests is available in the Validation Framework report.

6.1 Number of Lane Changes per Vehicle

Simulation models generally handle discretionary and mandatory lane changes based on multiple heuristic rules, triggers, thresholds, and parameters. A mandatory lane change is a necessary movement to position the vehicle on the appropriate lane for a turn or a highway exit. Discretionary lane changes are influenced by a number of factors such as the speed on the current and adjacent lanes. A high rate of lane changing may be due to higher than usual frequencies of discretionary lane changes.

Figure 6-1 shows the number of lane changes per vehicle in the model and field data. In the field data, the majority of the vehicles do not change lanes. In the model, the majority of the vehicles do at least one lane change. The distribution of the number of lane changes per vehicle is completely different in Table 6-1 which shows the total number of lane changes by lane. In the model no vehicle is leaving lane one because it is the high-occupancy vehicle lane that travels significantly faster than the rest of the lanes. In the NGSIM data the number of lane changes for every other lane but the first one is significantly higher than the number of lane changes in the model.

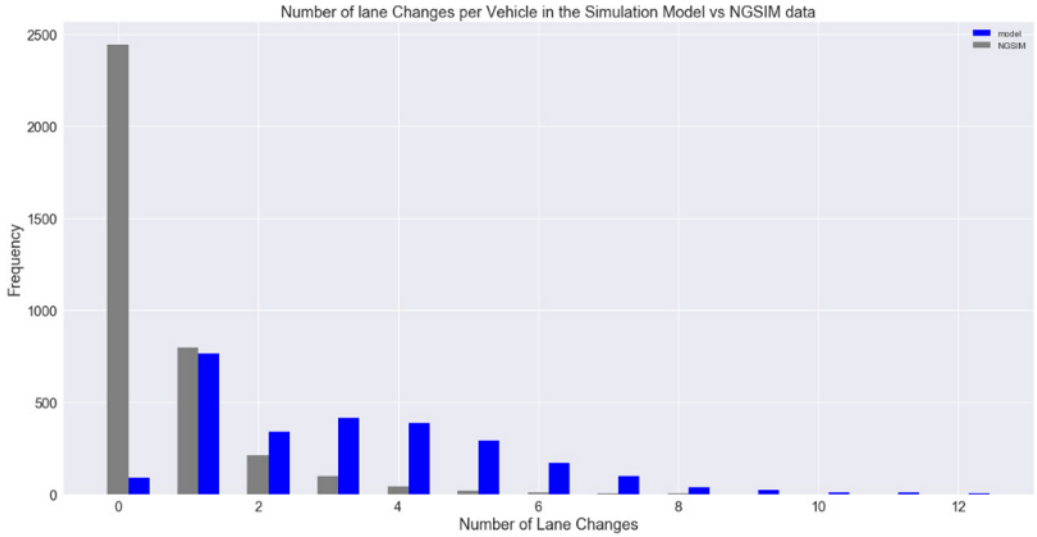


Figure 6-1. Graph. Number of lane changes per vehicle (model and data).
(Source: Cambridge Systematics, Inc.)

Note that the classification of lane changes into either discretionary or mandatory types is often subjective and involves the judgment of the analyst reviewing the data. Note also that the relative frequency of each lane change type is dependent upon roadway type/configuration. Therefore, care must be taken to ensure that the simulation trajectory dataset is compared to an appropriate reference dataset (e.g., one with similar configuration). Depending on the extent of the simulation model, it may be appropriate to perform this test multiple times for different roadway segments in the simulation.

Table 6-1. Number of lane changes per lane (model and data).

	Total Number of Lane Changes (Vehicles Leaving their Current Lane)	
Lane	**Model**	**NGSIM Data**
1 (HOV)	0	57
2	1126	121
3	1730	197
4	1668	275
5	1534	405
6	1261	371
7 (rightmost)	795	446

Source: Cambridge Systematics, Inc.

A statistical analysis of these data can be performed using a Chi Square test to compare the number of lane changes per vehicle between the NGSIM field data set and the reference simulation model data. This test can be applied to compare the overall distributions across all

bins, or to look at individual lane change count bins (e.g., the proportion of vehicles that made exactly 3 lane changes). Using an overall test, the conclusion is that the two distributions are different at a 95 percent level of statistical significance. Figure 6-2 shows the analysis results by individual bin, where dark gray borders indicate a statistically significant difference. As the figure shows, even when considered individually, every bin exhibited a statistically significant difference between the simulated and observed field data.

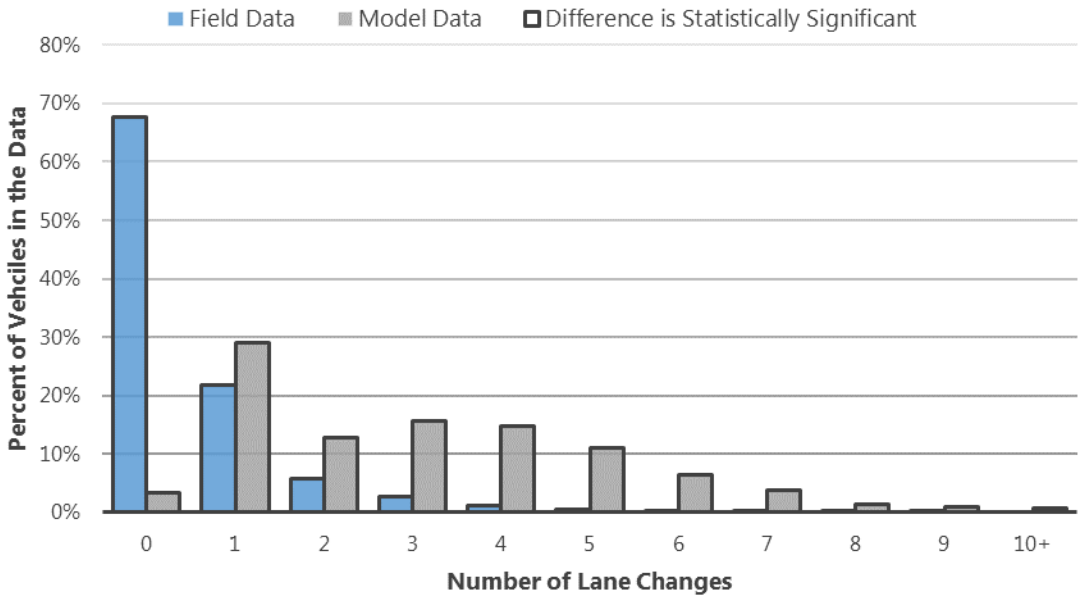


Figure 6-2. Chart. Comparisons of distributions of lane change counts per vehicle.
(Source: Cambridge Systematics, Inc.)

6.2 Lane Change Rate

Lane change rate (LCR) provides insight into the intensity of lane-changing maneuvers on a corridor level, which has effects on speed, capacity, and flow. LCR can be visualized over space and time in a heat map to identify locations and times of intense lane changing. Results from this test are highly site-specific and cannot be generalized because weaving patters are location specific. Therefore, in the absence of site-specific field observations of lane changing behavior (e.g., a data set similar to the NGSIM data), this validation test for simulation models is not feasible. However, even without on-site full trajectory data the LCR heat map can provide significant insight on the spatial and temporal distribution of lane changing that cannot be obtained otherwise.

In Figure 6-3 the LCR heat map for the field data is drawn using a spatial bin of 100 feet and a temporal bin of 60 seconds. The lane change rate is computed as the number of vehicles leaving a given lane per mile and hour. The LCR maximum value depends on the time and space discretization, and for this reason it is recommended that the same spatial and temporal discretization is used for all comparisons. In Figure 6-3 a discrete color map is used to associate LCR levels to colors. The maximum value is 72 thousand lane changes per mile per hour. Each lane in the corridor is represented by a separate rectangle. Distance along the direction of travel is shown on the y-axis while time is shown on the x-axis for all the lanes/rectangles. It is clear

from Figure 6-3 that most lane changing happens on lane 7, which merges into lane 6. The horizontal stripe of high LRC values represents vehicles that merge to lane 6 at the merge point. On the rest of the lanes, lane changing does not seem to have a specific pattern and seems evenly distributed in time and space. The minimum amount of lane changing happens on lane 1 and lane changing intensity increases from left to right in the corridor, as expected. Table 6-1 provides the total number of vehicles leaving any given lane.

Figure 6-4 shows the LCR for the modeled trajectories using the same temporal and spatial discretization. It is clear that the intensity of lane changing is significantly higher in the model compared to the field data even though the maximum value of LCR is similar. In the field data a great amount of lane changing happened at the end lane 7. In the model however, the same amount of high intensity of lane changing that is observed at lane 7 is happening on all lanes but the first. Specifically, the horizontal stripe of high intensity indicates that at about 400 feet from the beginning of the corridor vehicles swap lanes at the maximum rate. It is unclear what triggers this behavior in the model since the speed heat map of Figure 3-3 shows that there is not a significant speed advantage among lanes except lane 1. A minor difference between the field data and the model is that on lane 7, where vehicles don't wait to reach the end of the lane to merge into lane 6. If LCR values are added up by time and space they will result in the lane change numbers provided in Table 6-1. Lane 1 in Figure 6-4 shows zero lane changing activity because no vehicle leaves lane 1 to merge to lane 2. LCR also can be calculated for the vehicles that enter a lane. In this case, lane 1 would have shown a significant amount of vehicles entering.

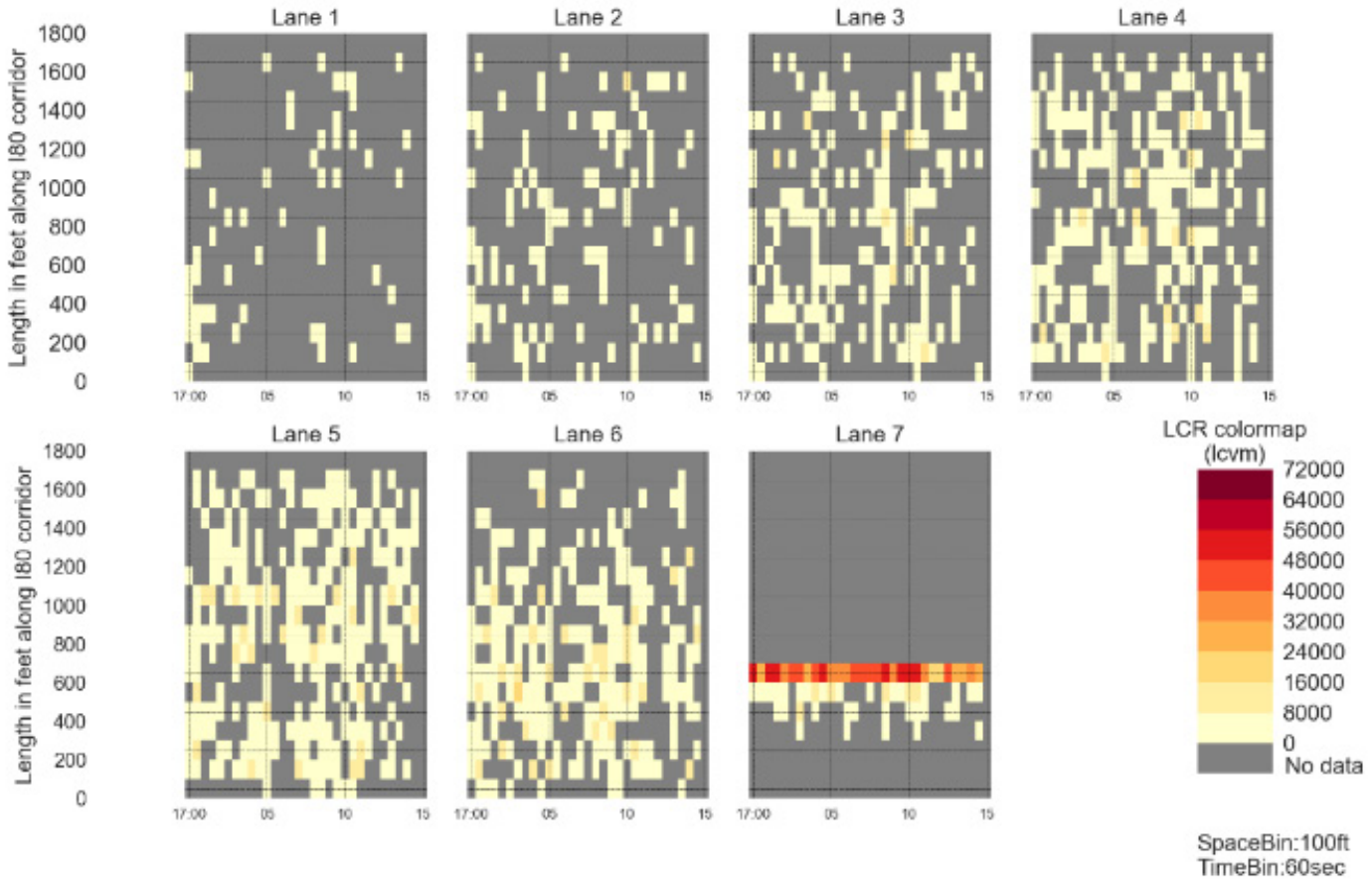


Figure 6-3. Charts. Lane change rate heat map for vehicles leaving their current lane (Next Generation Simulation).
(Source: Cambridge Systematics, Inc.)

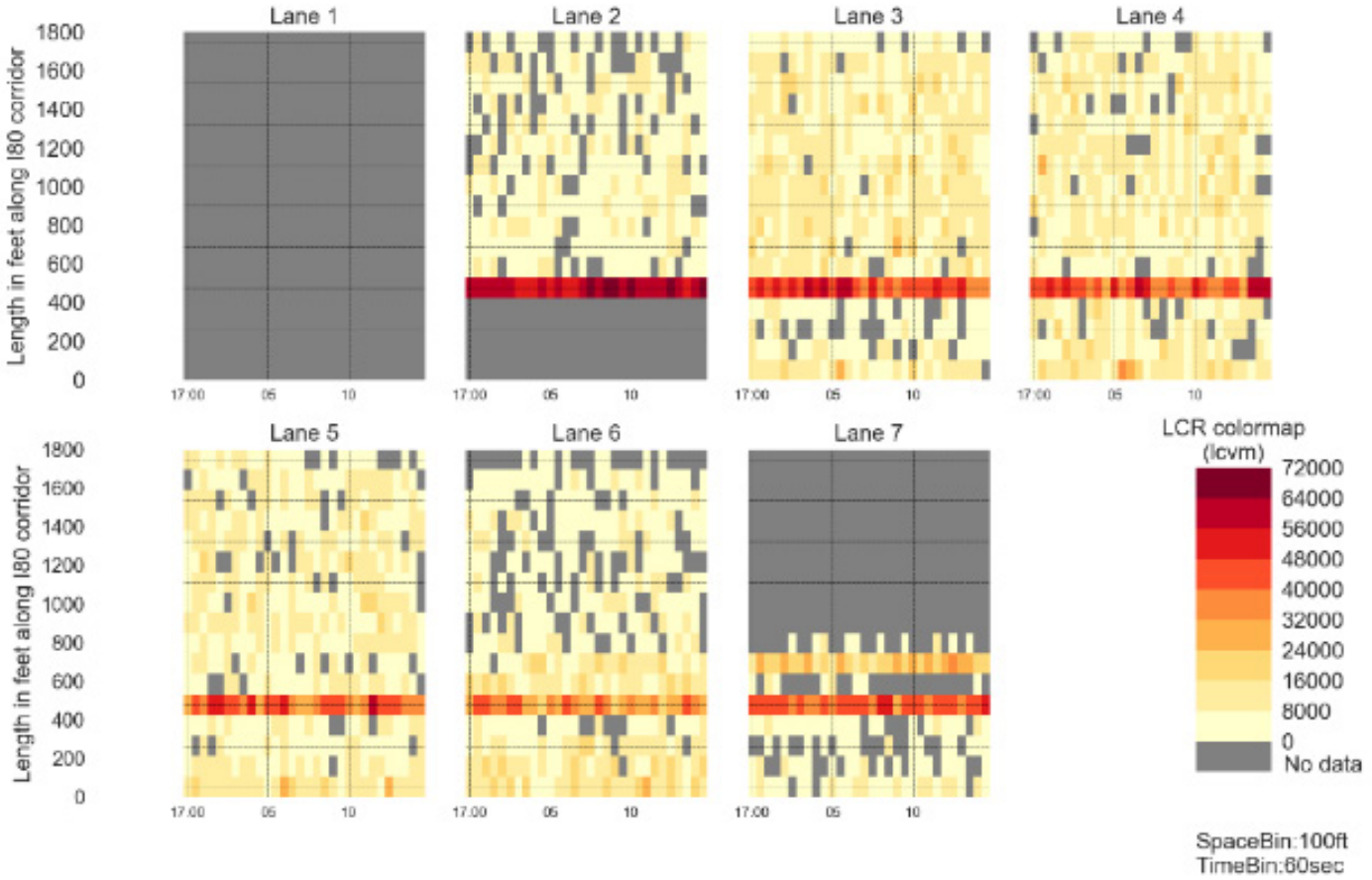


Figure 6-4. Charts. Lane change rate heat map for vehicles leaving their current lane (model).
(Source: Cambridge Systematics, Inc.)

6.3 Fundamental Diagram Comparisons

The Fundamental Diagram (FD) is the primary method to express and document traffic flow properties for a specific section or link. Furthermore, it is the basic and broadly relevant means to convey changes in flow properties such as the impact of weather and connected vehicles. In those cases, the shape of the diagram with respect to capacity and speed-at-capacity varies. In general, the shape of the fundamental diagram is site specific, though it is sometimes possible to generalize based on such factors as roadway grade, traffic characteristics, weather, and other factors. If the fundamental diagram is estimated using traffic measurements from a single spot location over time, the results can be less accurate due to the assumptions that are required to calculate the required quantities from spot measurements.

Comparing the observed and modeled FD is important because it provides insight on how the model behaves at different combinations of flow, speed, and density. This case is particularly important when the model is calibrated to a limited number of base conditions but it is used to predict congestion to a wide range of diverse scenarios in an Active Transportation and Demand Management (ATDM) application. For a given segment, a modeled FD that closely matches the observed FD is a check that the model behaves similar to observations under different flow and density conditions.

In Figure 6-5, the speed versus density relationship of the fundamental diagram is shown for both the model and the field data (NGSIM). This diagram combines all the lanes together into one plot. Figure 6-6 and Figure 6-7 present the FD for each lane. Significant differences exist between lane 1 which is the most uncongested lane in the model and field data and the rest of the lanes which have similar levels of congestion. To obtain the space-mean speed and density measurements a space bin of 100 feet and a time bin of 10 seconds were used.

The top graph in Figure 6-5 shows the field data and the bottom figure the modeled data. A point corresponds to a single space-time bin of 100 feet and 10 seconds. In both the top and bottom graphs, the dispersion of points shows the deviation from the homogeneous conditions (e.g., same speed and gap) that are assumed in the analytical derivation of the FD. To make comparisons easier, the average speed by density level also is plotted as a separate line on top of the individual points. Based on the top and bottom plot in Figure 6-5 the following observations can be made:

- For densities greater than 50 vehicles per mile (vpm), field space-mean speeds are approximately twice as great as modeled speeds. For example, at 50 vpm field space-mean speed is approximately 21 miles per hour and modeled speed is about 11 miles per hour. For densities greater than 50 vpm space mean speeds decrease linearly in both graphs. As a reference, based on the 2000 version of the Highway Capacity Manual, less than 11 passenger cars per hour per lane (pcphpl) correspond to level of service A. Level of service B corresponds to 11 to 18 pcphpl. Traffic flow is classified as level of service C when pcphpl is between 18 and 26. Level of service D and E correspond to ranges of pcphpl between 26 and 35 and 35 and 45. Finally, the F level of service occurs when pcphpl is greater than 45.
- For densities less than 50 vpm there is a very wide dispersion in the NGSIM field data and a smaller number of measurements. For low-density values closer to 10 vpm, the average space-mean speed is significantly lower than the free-flow speed in the corridor. Such a deviation is attributed to the fact the corridor was heavily congested when data were collected and the number of points with low density is not sufficient to draw the curve for the uncongested portion of the FD. The top left graph in Figure 6-6 which breaks down the FD by lane has more information of the shape of the FD for uncongested conditions, even though

this information is still insufficient. This is because the left most lane is the most uncongested of the seven in both the field data and the model.

- For densities less than 50 vpm, modeled data show a higher speed close to 40 miles per hour and significantly less variation than the field data. Based on the speed heatmaps presented in section 2.4, vehicles that travel with 40 miles per hour are mostly located on the left most lane while the rest of the lanes are heavily congested. This also is clear from the lane-based diagrams in Figure 6-7.

Clearly, observations from a larger time period and traffic conditions are needed for the derivation of the fundamental diagram for both the modeled and the observed data. The available data presented in Figure 6-5 and Figure 6-6 correspond mostly to heavily congested conditions and may not be relied upon for the estimation of the uncongested portion of the FD.

Statistical tests for comparing the two distributions are not easy to construct because the points in each of the figures of this section are part of a two-dimensional distribution. Statistical comparisons between two-dimensional distributions are harder to construct and can depend heavily on the underlying normality assumptions. However, even though no statistical tests have been conducted, it can be concluded empirically that the differences are significant. This is mainly because for the same congested density the model has approximately half the space-mean speed observed data have.

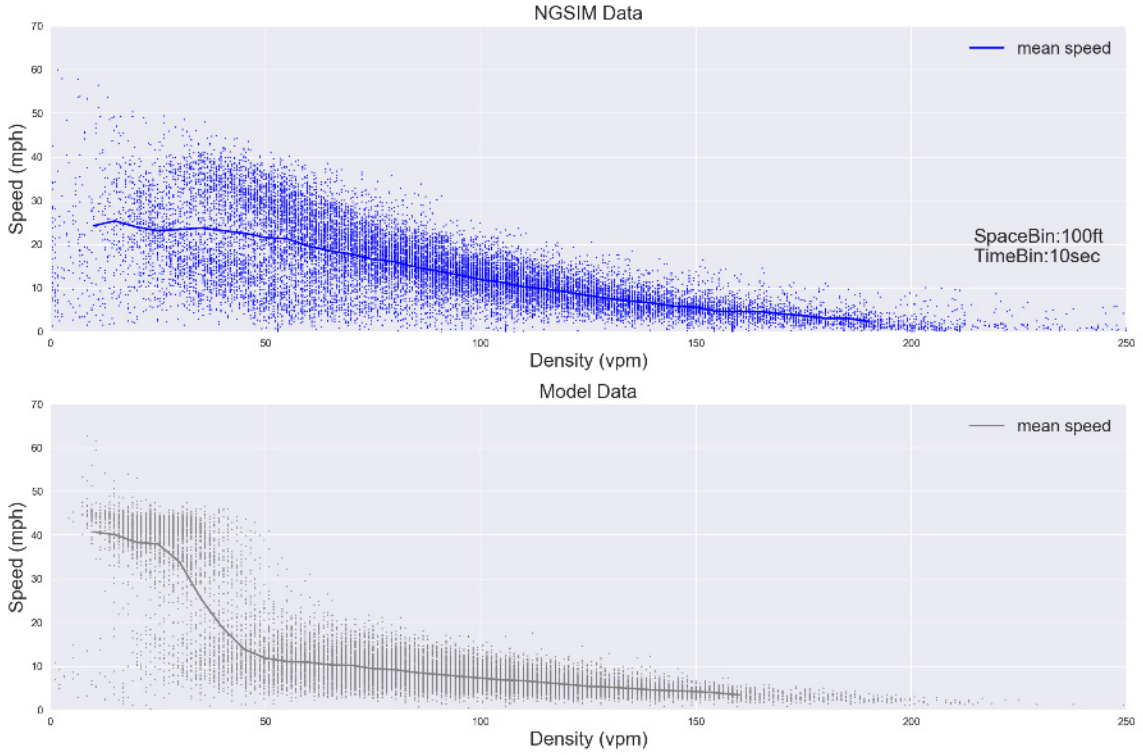


Figure 6-5. Graphs. Fundamental diagram (speed versus density) for all lanes.
(Source: Cambridge Systematics, Inc.)

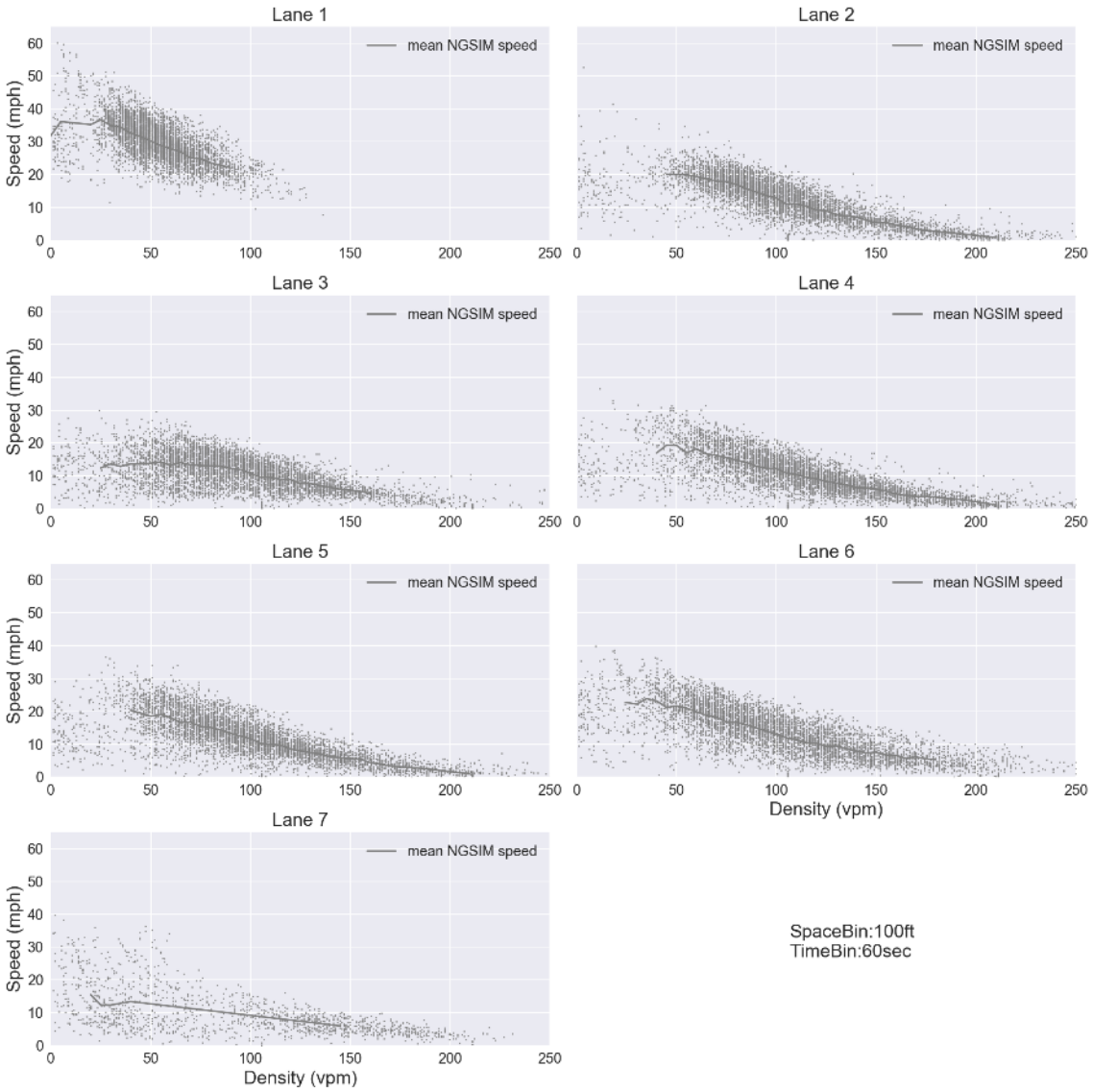


Figure 6-6. Graphs. Fundamental diagrams for each lane (data).
(Source: Cambridge Systematics, Inc.)

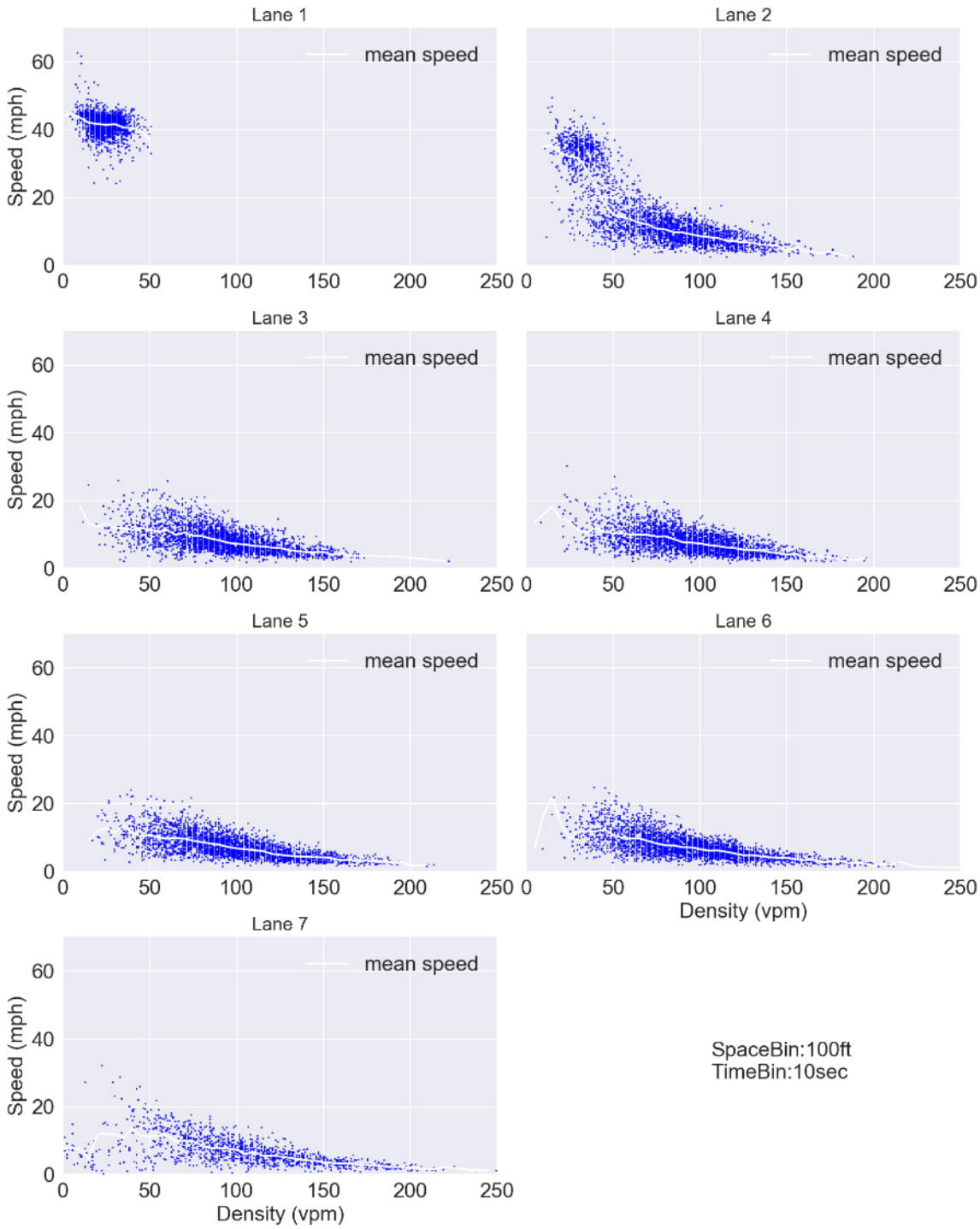


Figure 6-7. Graphs. Fundamental diagrams for each lane (model and data).
(Source: Cambridge Systematics, Inc.)

Chapter 7. Conclusions

A model that is calibrated at the aggregate level and successfully replicates the observed flow or travel time may have a significant number of vehicles behaving unrealistically at the trajectory level. The trajectory-based checks, as it has been demonstrated in this report, can reveal significant deviations from what it is considered acceptable driving behavior. For example, a model that cannot reproduce a realistic speed versus acceleration pattern will likely provide biased results in any emissions related scenario evaluation. Simulating the interactions of autonomous vehicles that are guided by an external logic with the rest of the traffic may not produce valid results if the vehicles that adhere to the car-following rules have unrealistic decelerations. It is recommended that near-future scenarios that involve connected vehicles or advanced driver assistance technologies are not conducted unless the simulation model is validated for realistic driver dynamics.

The trajectory-based validation tests can also be beneficial when simulating a vehicle fleet with the present technological characteristics. For example, excessive lane changing or unrealistic lane changing patterns can be revealed by the related tests in the framework and the lane change rate heatmap visual. In microscopic simulation models target macroscopic measures such as capacity at a weaving section arise from the collective behavior and interactions of the individual drivers. Therefore, the higher the realism in simulated driving dynamics the easier it is for the modeler to replicate macroscopic phenomena and the wider the range of traffic conditions that can be successfully replicated.

The specific tests applied in this report revealed unrealistic behavior in the model for both the longitudinal and the lateral direction. In the lateral direction, at the microscopic level, the aggressiveness or collision risk as captured by the relative distances and speeds in the corresponding validation tests do not reveal significant qualitative differences. Even though statistically the distributions are different, the differences are markedly smaller than the discrepancies revealed in other tests such as the number of safety events. At the macroscopic level, there are significantly more lane changes compared with ether the field or the naturalistic data. Such a high number of lane changes clearly reveals a tendency for modeled drivers to benefit from minor speed differences on adjacent lanes. Even though the lane changes themselves are not significantly more risky or aggressive at the micro level than those in the field data their high number per driver is clearly unrealistic. The intensity and propensity for lane changing may have an impact on the modeled weaving capacity of the junction.

In the longitudinal direction, modeled vehicles have mechanically infeasible decelerations, a finding that also has been reported in the past for some simulation models. It is not clear however if this behavior is the outcome of certain parameter ranges in the car-following model or the byproduct of some other behavior for which the relaxing of deceleration limit acts as a remedy or benefits the macroscopic model performance. Collisions were found in the model but the individual circumstances of their occurrence or their sensitivity to model parameters were not investigated. Clearly, near crashes and crashes happen unrealistically frequently in the model, orders of magnitude more frequently than in the naturalistic data. The speed versus acceleration relationship is especially important for a number of applications including emissions. It was found that this relationship in the model deviates significantly from observations, which show that the 95th percentile of acceleration decreases as speed increases. The variation of modeled

acceleration is significantly higher than in observations, which would constitute a bumpy and uncomfortable ride. At the macroscopic level and in the longitudinal direction, the congested portion of the modeled fundamental diagram is flatter than in the field data. This implies a smaller shockwave speed and smaller speed at the same density level.

Appendix A. Trajectory Plots

In this appendix, vehicle trajectories are plotted using a space-time diagram. Each plot is specific to a lane. Each trajectory point in the data or in the model is a dot. Dots are color coded by speed using the same color map scale that was used in the speed heat maps. The dots that represent trajectory start and end points are plotted with dots that have twice the diameter of intermediate trajectory points.

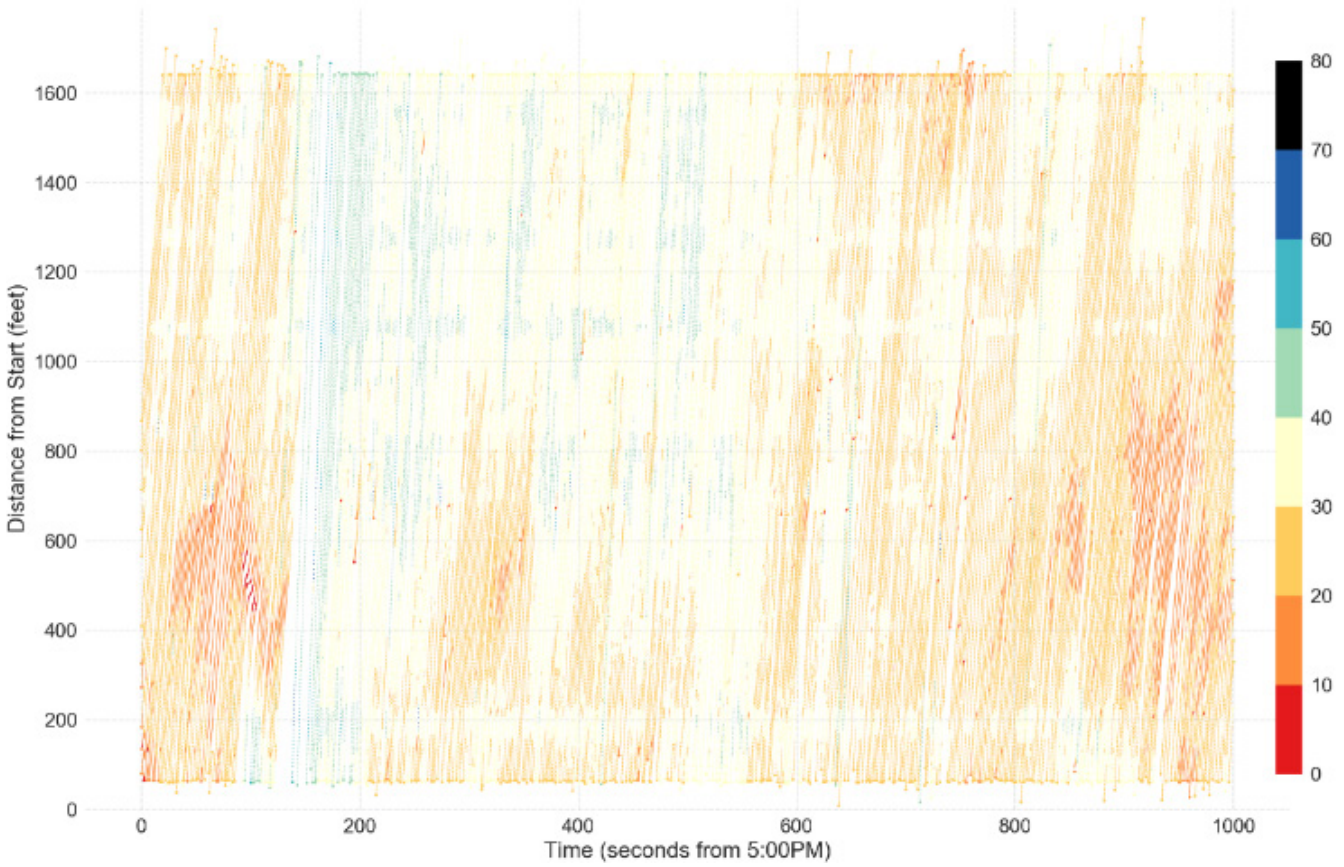


Figure A-1. Graph. Field trajectories for lane 1.

(Source: Cambridge Systematics, Inc.)

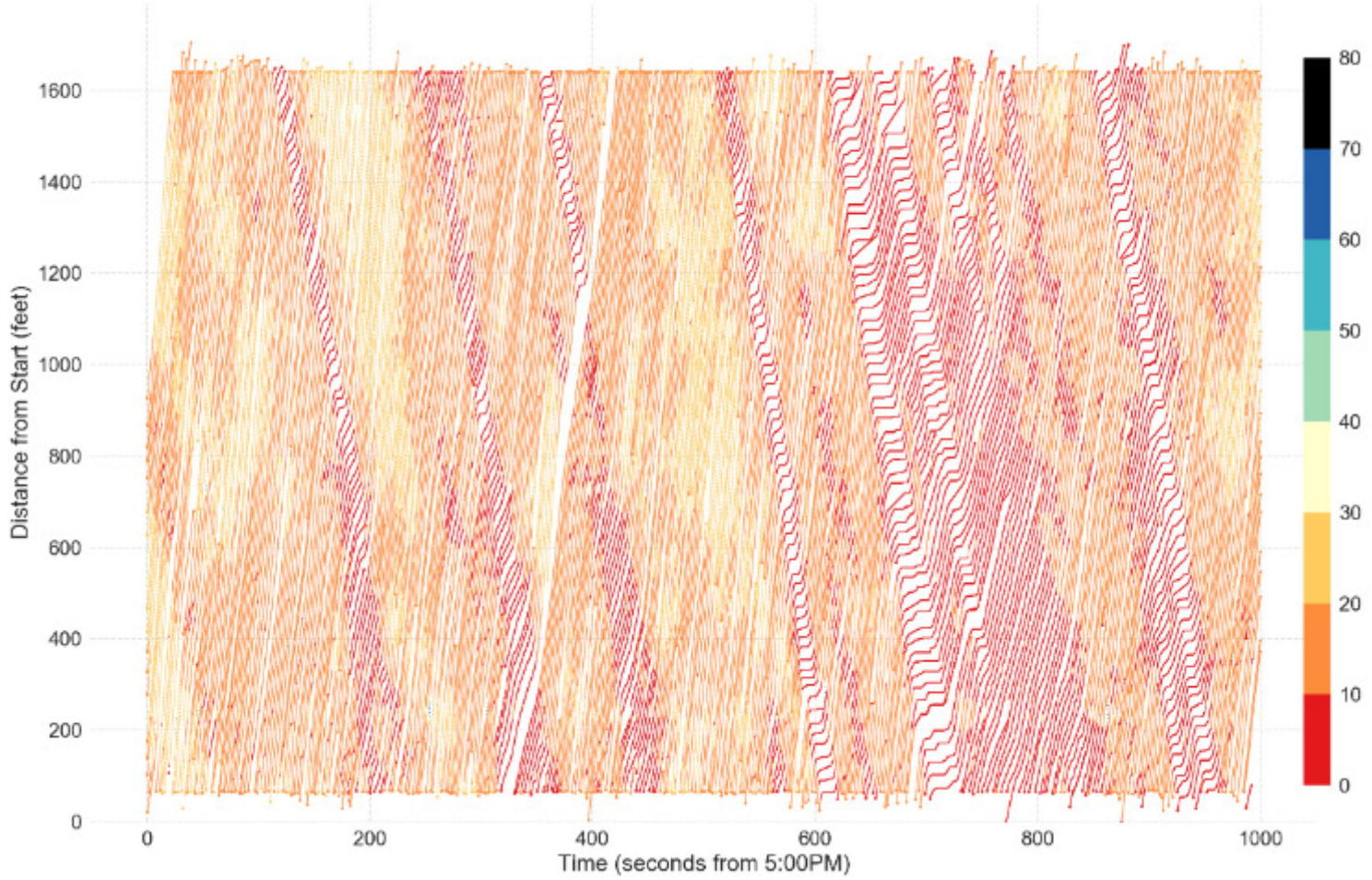


Figure A-2. Graph. Field trajectories for lane 2.
(Source: Cambridge Systematics, Inc.)

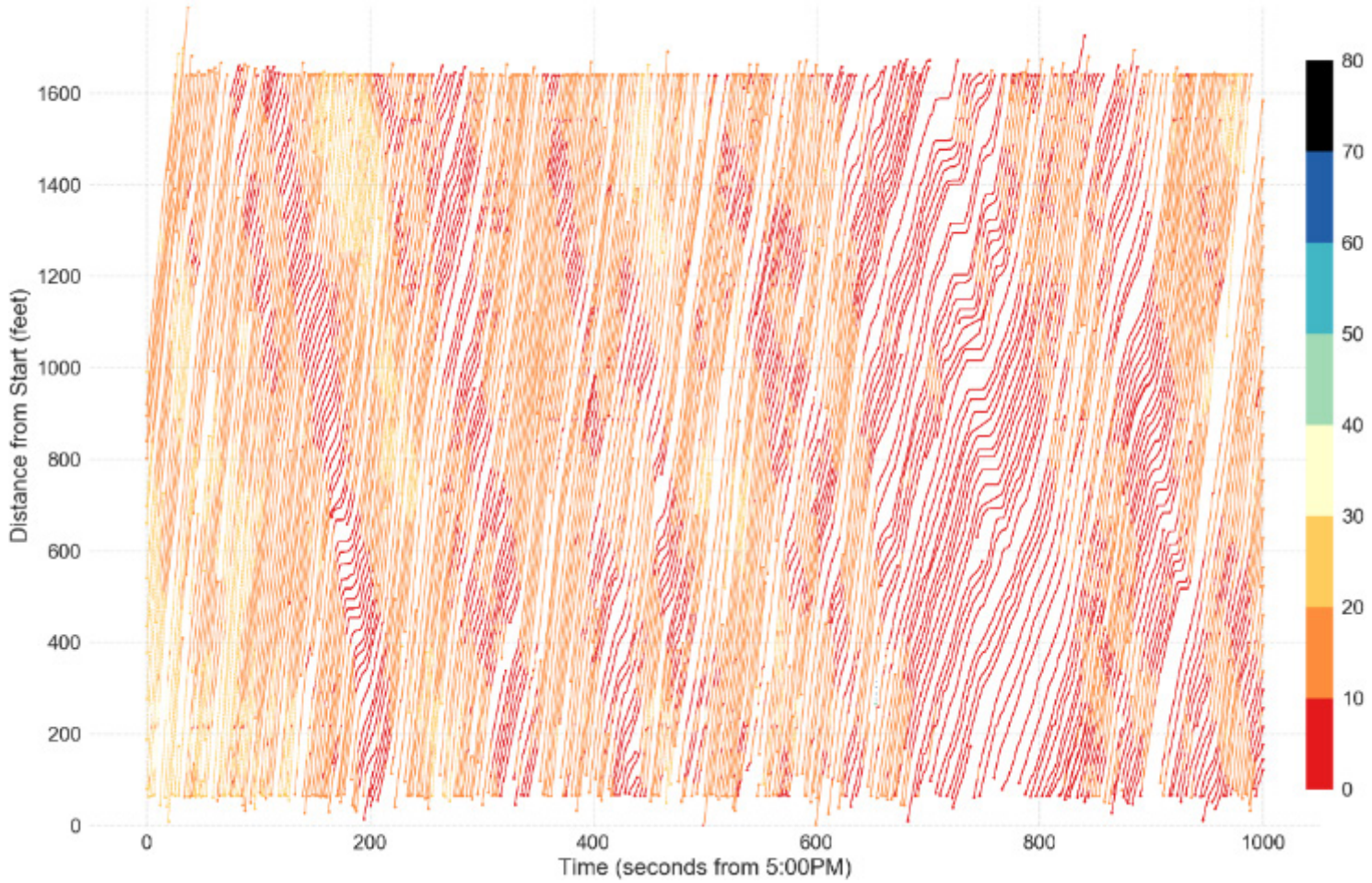


Figure A-3. Graph. Field trajectories for lane 3.
(Source: Cambridge Systematics, Inc.)

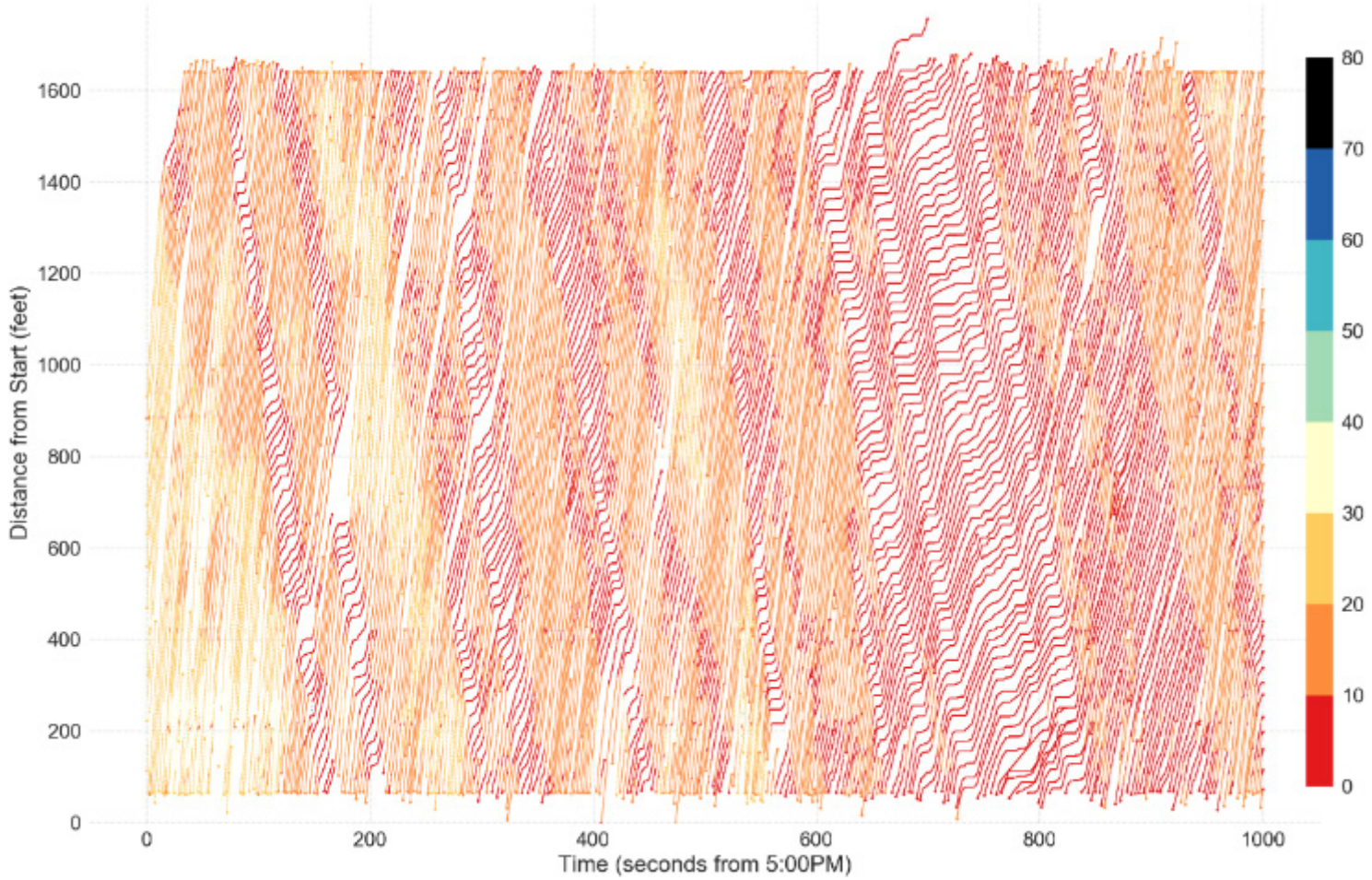


Figure A-4. Graph. Field trajectories for lane 4.
(Source: Cambridge Systematics, Inc.)

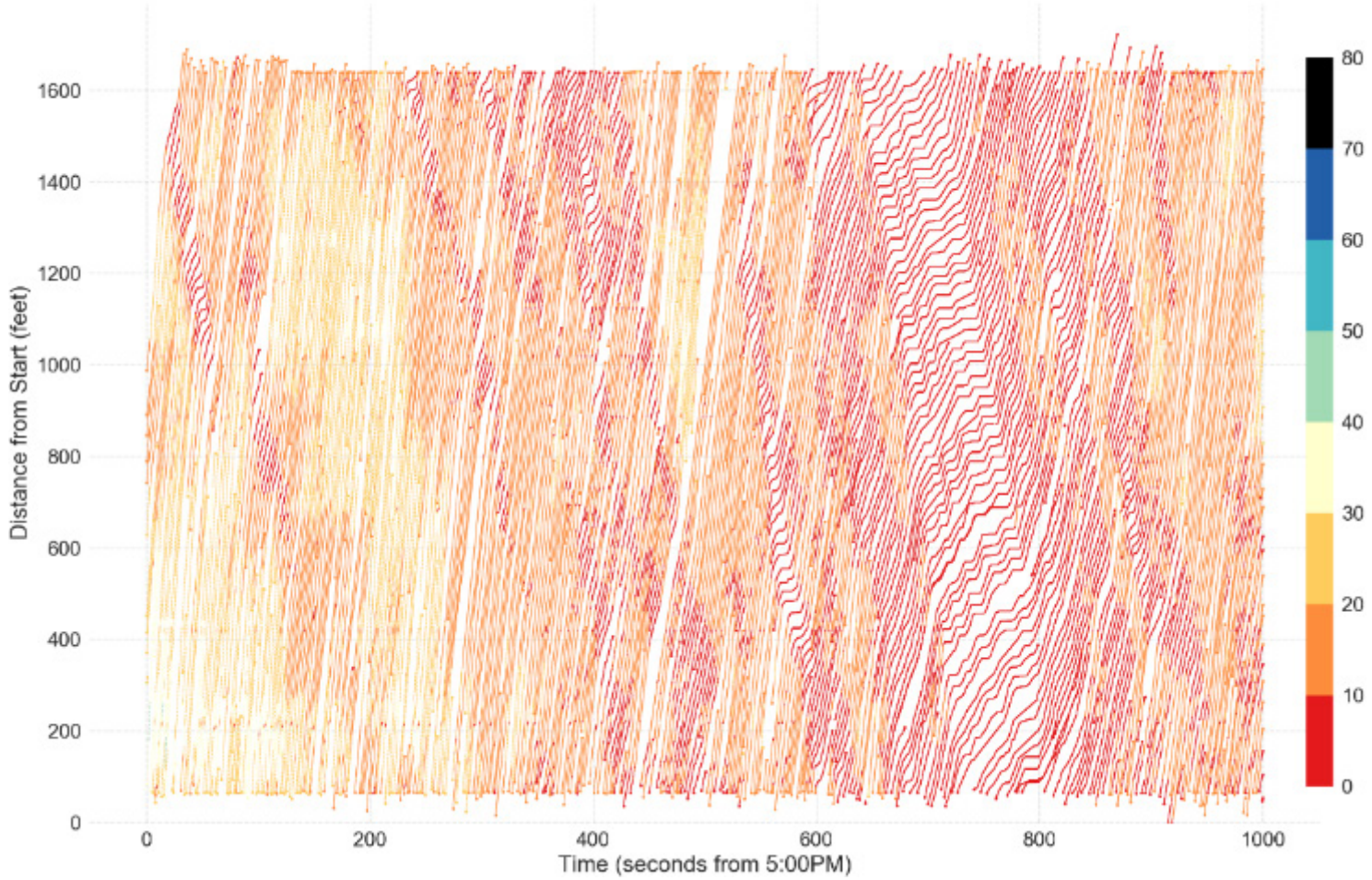


Figure A-5. Graph. Field trajectories for lane 5.
(Source: Cambridge Systematics, Inc.)

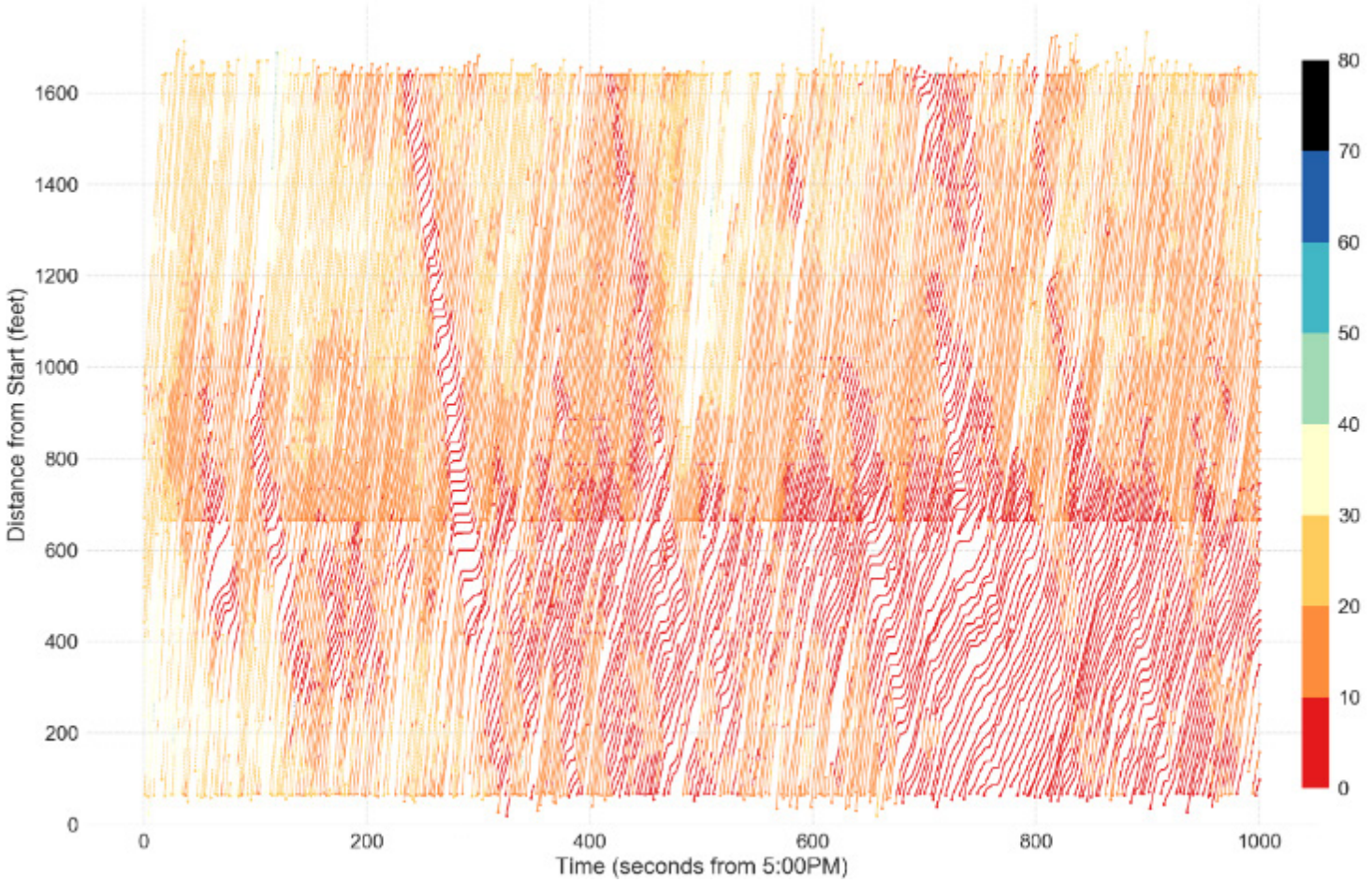


Figure A-6. Graph. Field trajectories for lane 6.
(Source: Cambridge Systematics, Inc.)

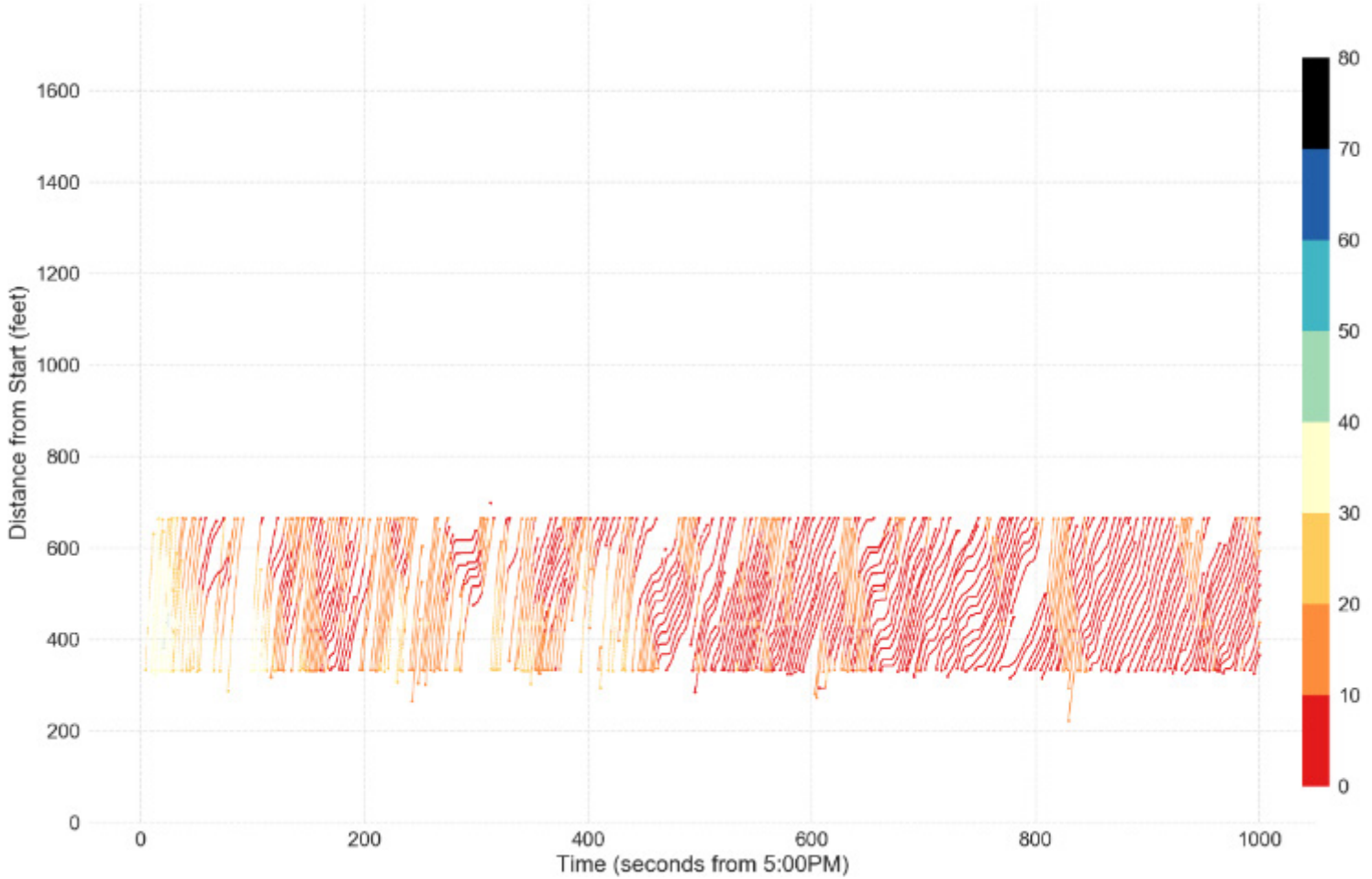


Figure A-7. Graph. Field trajectories for lane 7.
(Source: Cambridge Systematics, Inc.)

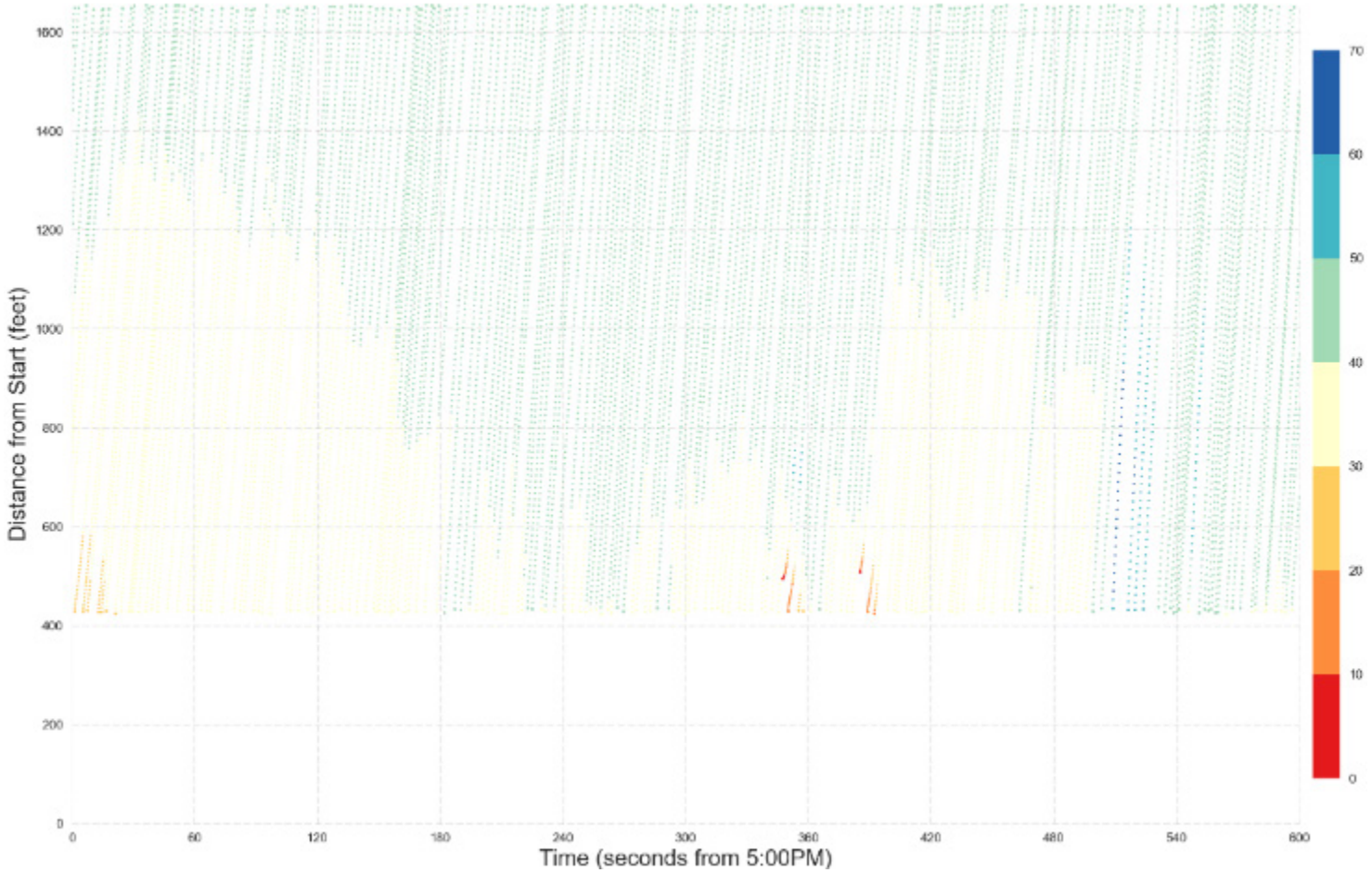


Figure A-8. Graph. Model trajectories for lane 1.
(Source: Cambridge Systematics, Inc.)

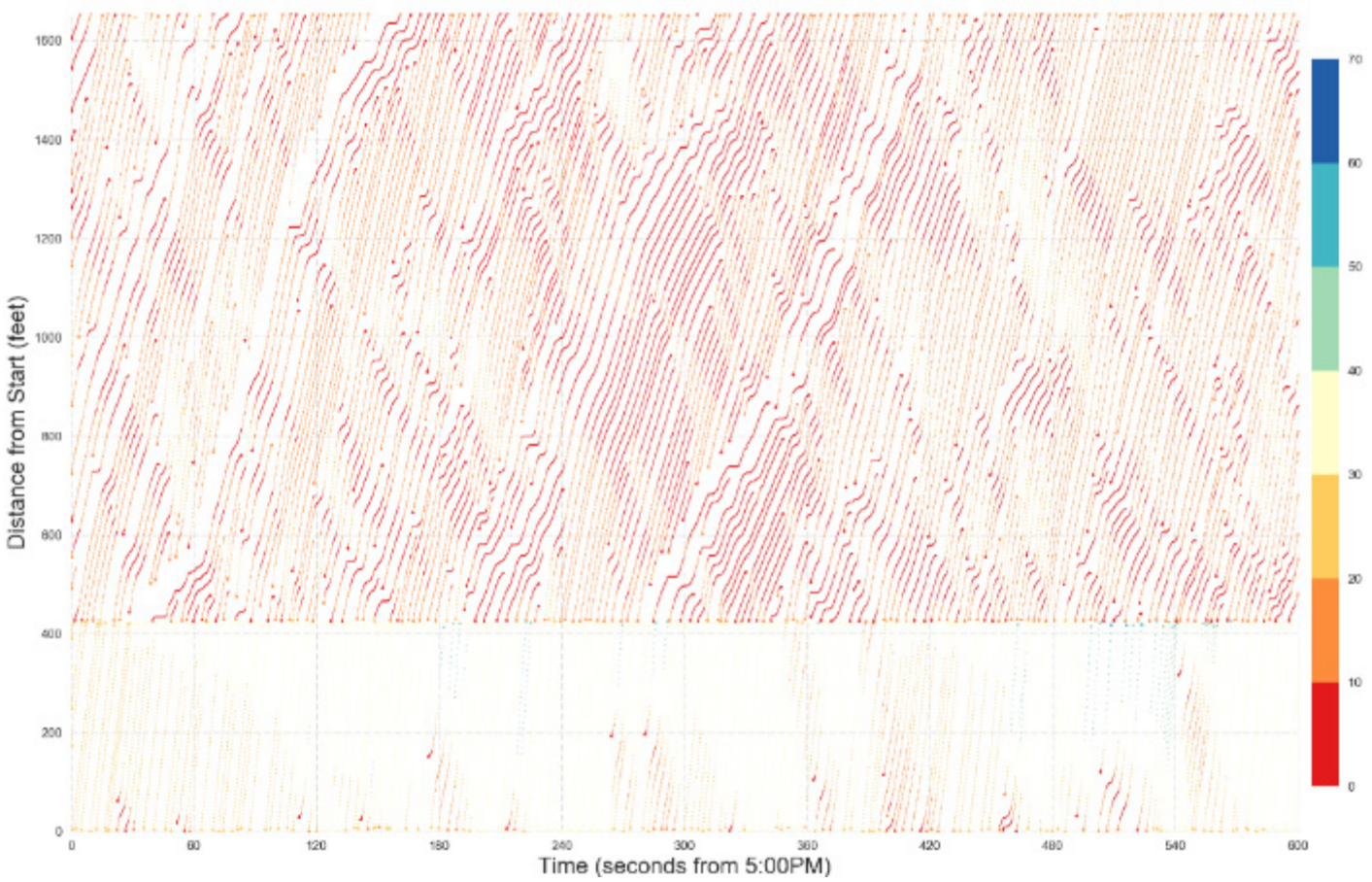


Figure A-9. Graph. Model trajectories for lane 2.
(Source: Cambridge Systematics, Inc.)

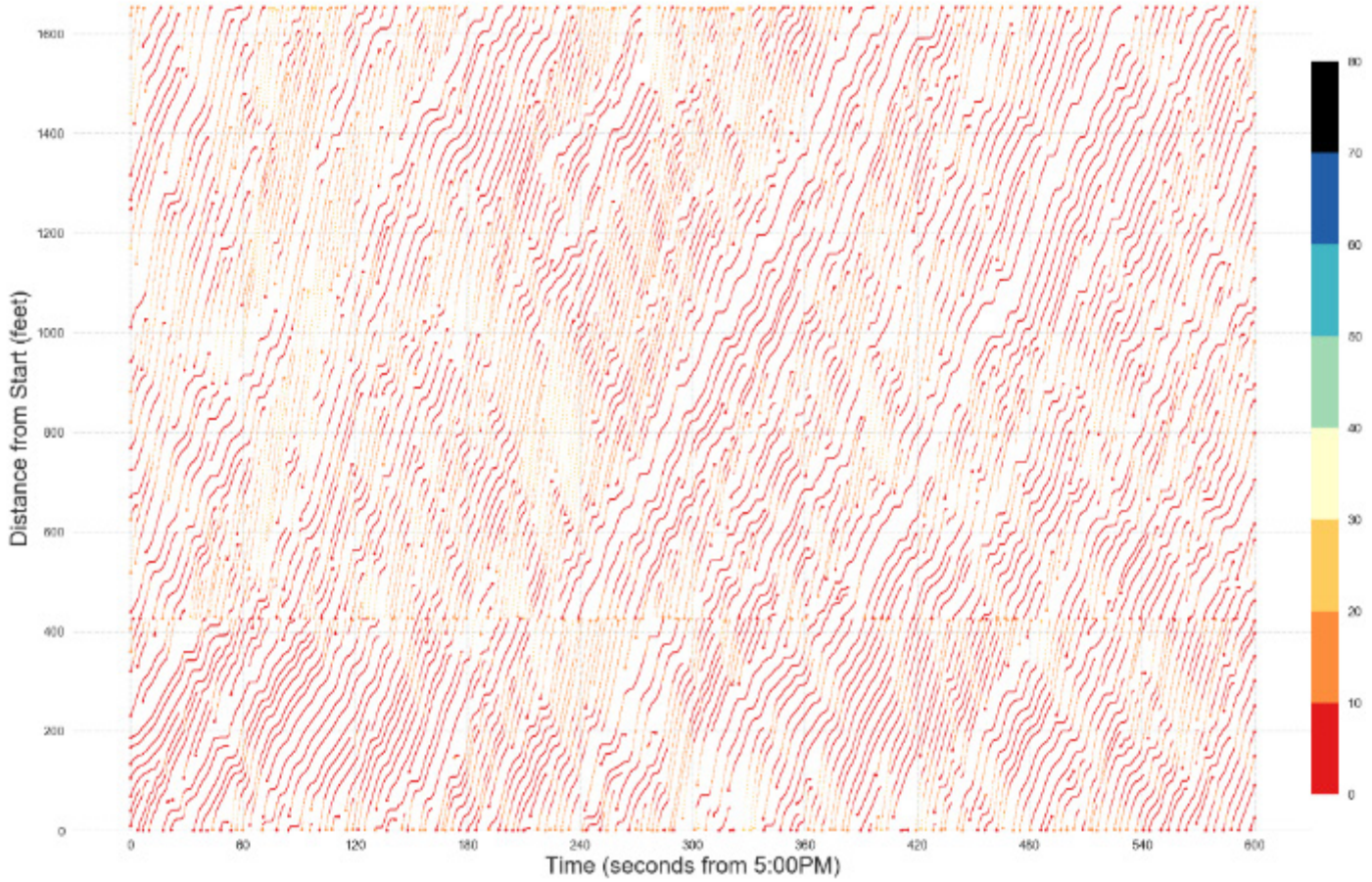


Figure A-10. Graph. Model trajectories for lane 3.
(Source: Cambridge Systematics, Inc.)

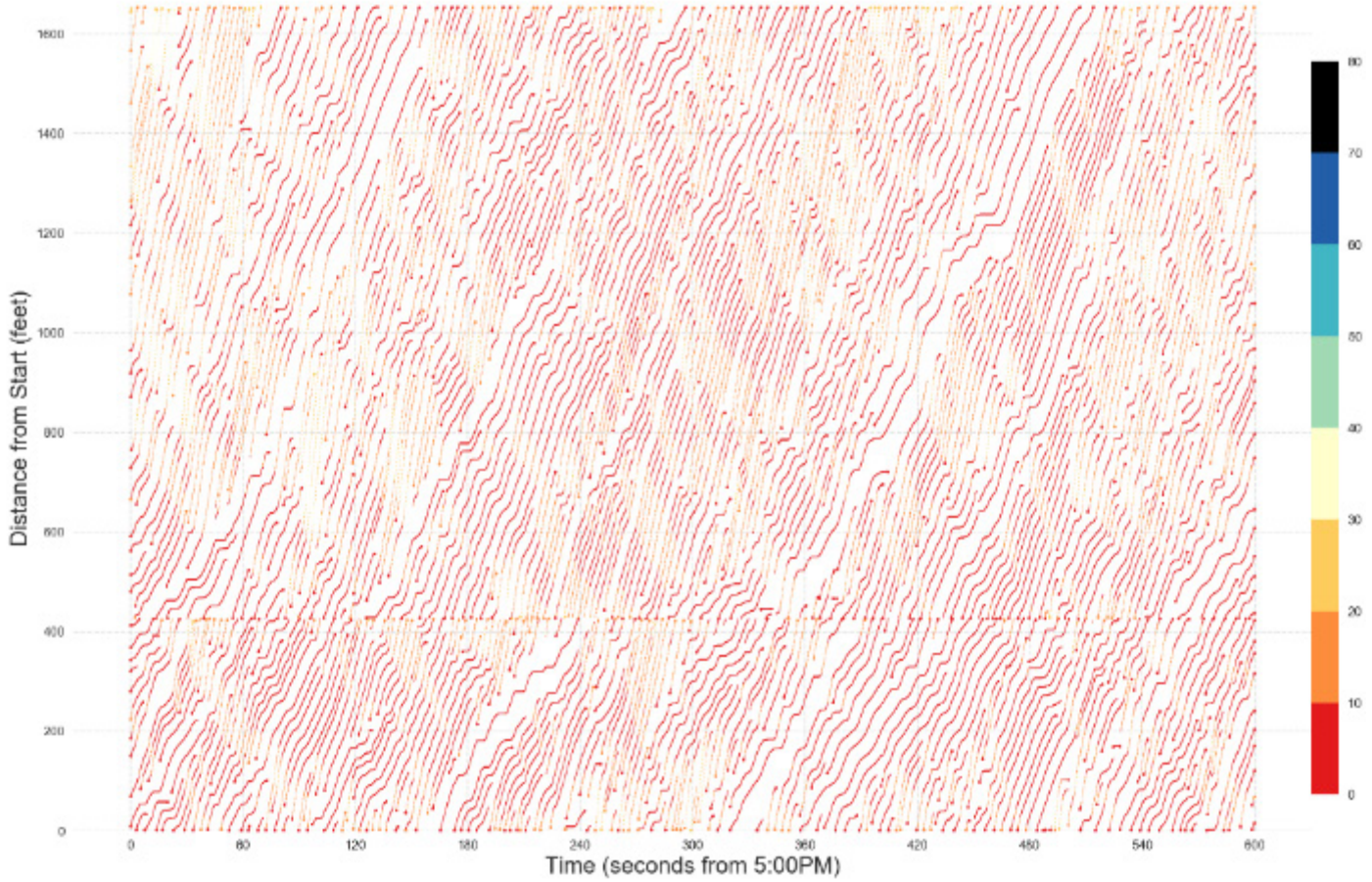


Figure A-11. Graph. Model trajectories for lane 4.
(Source: Cambridge Systematics, Inc.)

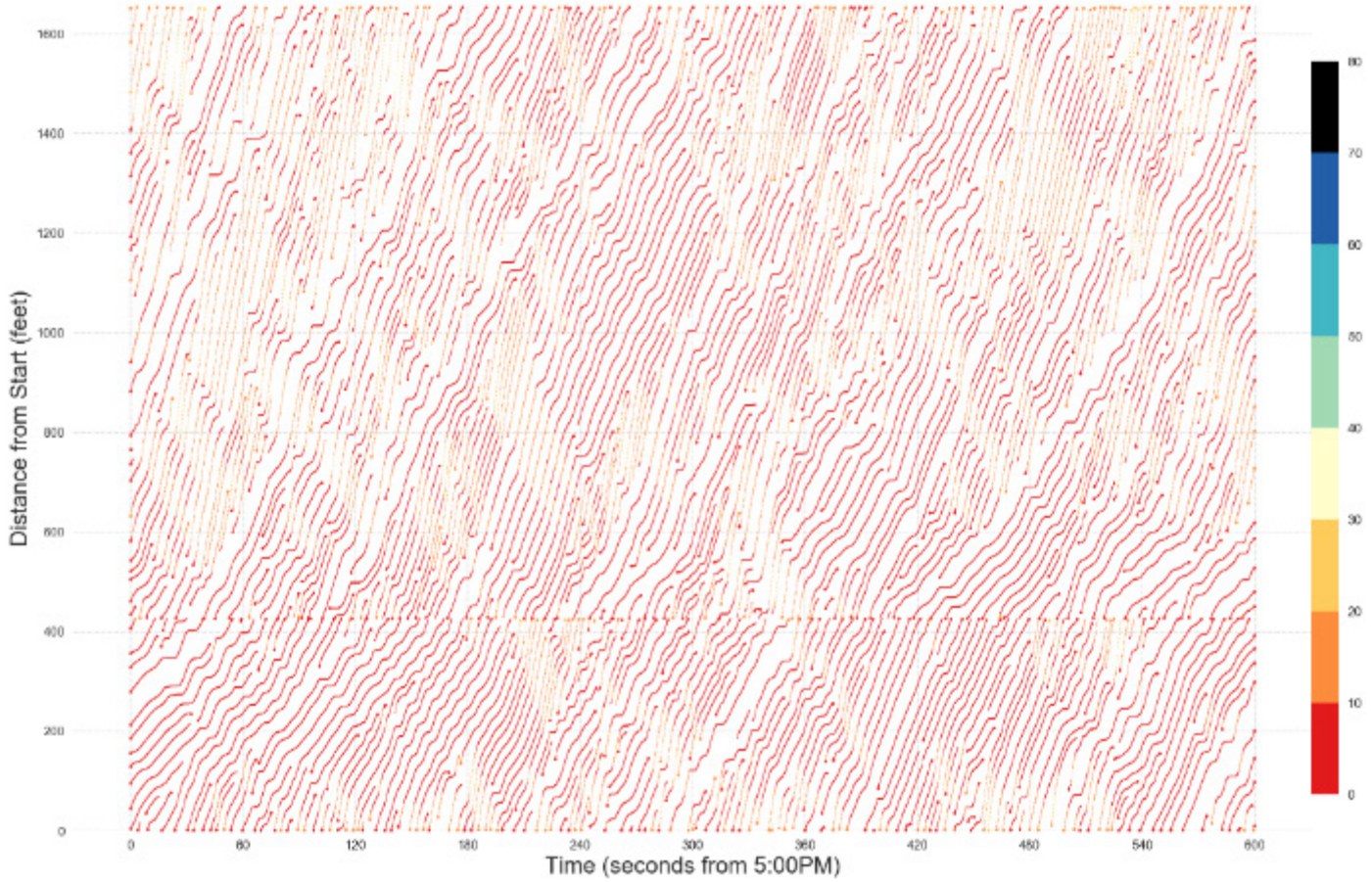


Figure A-12. Graph. Model trajectories for lane 5.
(Source: Cambridge Systematics, Inc.)

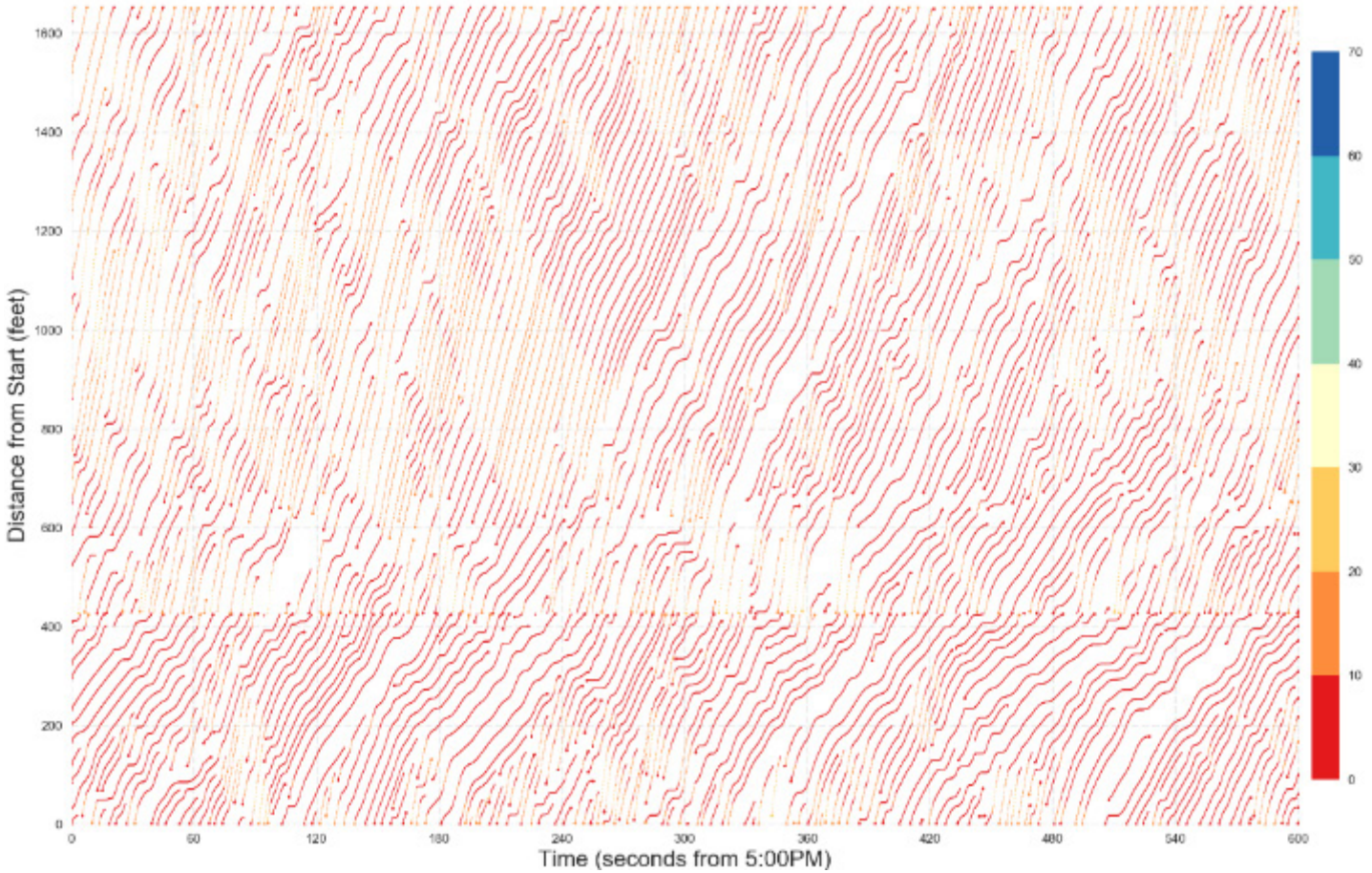


Figure A-13. Graph. Model trajectories for lane 6.
(Source: Cambridge Systematics, Inc.)

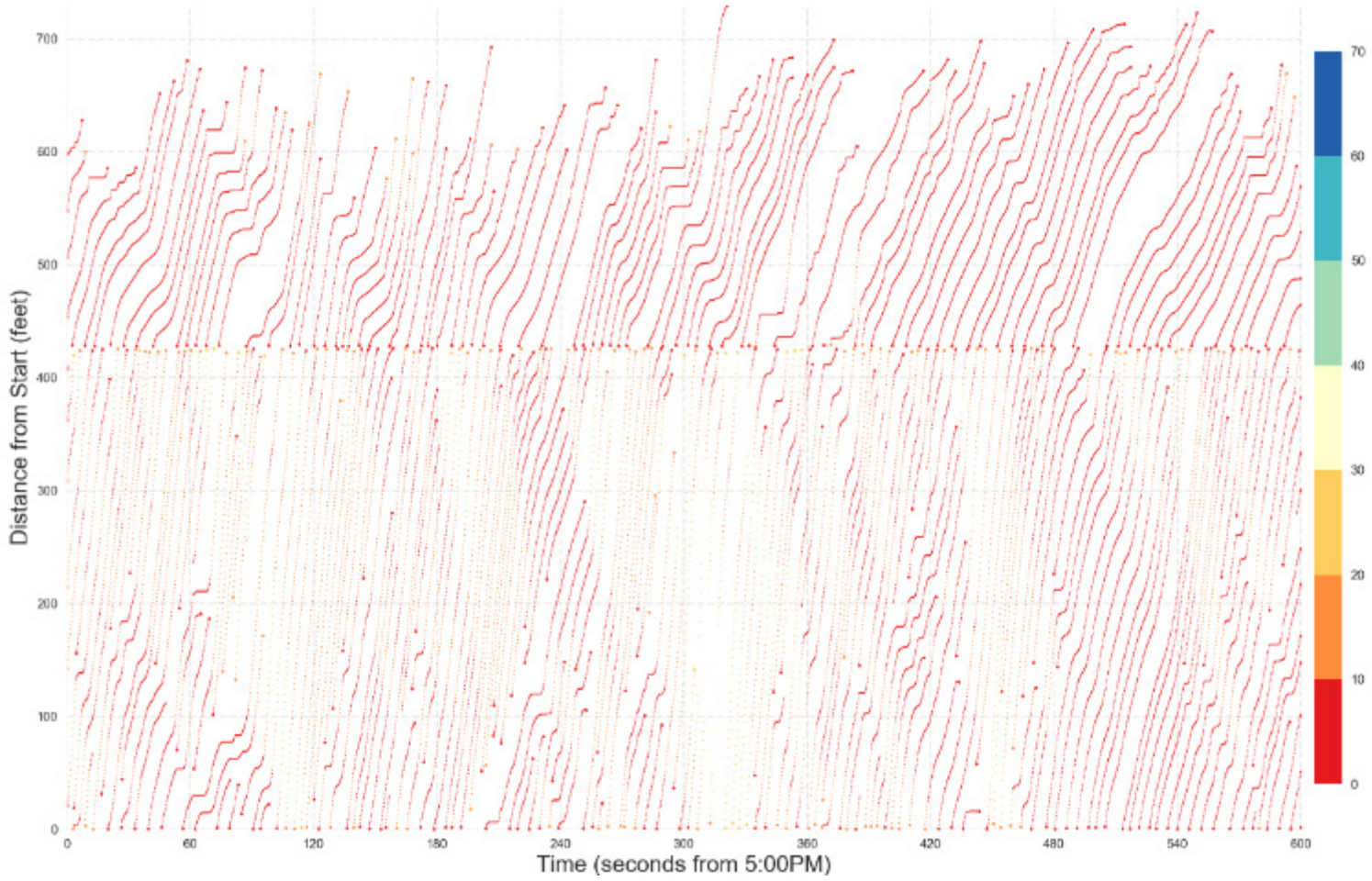


Figure A-14. Graph. Model trajectories for lane 7.

(Source: Cambridge Systematics, Inc.)

Appendix B. Reconstructed Next Generation Simulation

Reconstructed Next Generation Simulation (NGSIM) data for the same corridor but for a different day and time period are shown below. The reconstructed data by V. Punzo maintain consistency between position, speed, and acceleration calculations and have been smoothed using a Kalman Filter. (Montanino, Marcello, and Vincenzo Punzo. "Trajectory data reconstruction and simulation-based validation against macroscopic traffic patterns." *Transportation Research Part B: Methodological* 80 (2015): 82-106.) The percentages in the last column are based on total driving time.

Table B-1. Distribution of time spent at time to collision values between 0 and 14 seconds.

	Frequency	Percentage
0 <= TTC < 1	65	0.1
1 <= TTC < 2	412	0.5
2 <= TTC < 3	1,090	1.3
3 <= TTC < 4	2,001	2.3
4 <= TTC < 5	3,087	3.6
5 <= TTC < 6	4,766	5.6
6 <= TTC < 7	6,325	7.4
7 <= TTC < 8	7,202	8.4
8 <= TTC < 9	8,570	10.0
9 <= TTC < 10	9,592	11.2
10 <= TTC < 11	10,320	12.0
11 <= TTC < 12	10,701	12.5
12 <= TTC < 13	10,560	12.3
13 <= TTC < 14	11,053	12.9

Source: Cambridge Systematics, Inc.

Table B-2. Number of safety events.

Type	Number of Events
Collisions	0
Safety events	475
Warnings	812

Source: Cambridge Systematics, Inc.

Table B-3. Lane change severity distribution.

Category	Frequency	Percentage
1	33	4.1
2	60	7.4
3	317	39.4
4	266	33.1
5	129	16.0

Source: Cambridge Systematics, Inc.

Appendix C. List of Acronyms

AMS	Analysis, Modeling and Simulation
AASHTO	American Association of State Highway Transportation Officials
ABS	Anti-Lock Braking System
ACAS	Advanced Collision Avoidance Systems
ADAS	Advanced Driver Assistance Systems
AJ	Acceleration Jerk
AR	Acceleration Range
ARMS	Acceleration Root Mean Square
ATDM	Active Transportation and Demand Management
CAS	Collision Avoidance Systems
CAV	Connected Automated Vehicles
FCW	Forward Collision Warning
FD	Fundamental Diagram
FHWA	Federal Highway Administration
FMCSA	Federal Motor Carrier Safety Administration
fps	Frame per second
fpss	Frame per second squared
GPS	Global Positioning System
ISO	International Organization for Standardization
IVBSS	Integrated Vehicle-Based Safety Systems Project
K-S	Kolmogorov Smirnov
LC	Lane Change
LCVM	Lane Changes Per Mile
LCR	Lane Change Rate
LCS	Lane Change Severity
LCU	Lane Change Urgency
LIDAR	Light Detection And Ranging
MicroFD	Microscopic Fundamental Diagram

NDS	Naturalistic Driving Study
NGSIM	Next Generation Simulation
NHTSA	National Highway Traffic Safety Administration
NTSB	National Transportation Safety Board
pcphp	passenger cars per hour per lane
POV	Principal Other Vehicle
RMVMT	Rate per Million Vehicle Miles Traveled
SAE	Society of Automotive Engineers
SSAM	Surrogate Safety Assessment Model
SV	Subject Vehicle
TRB	Transportation Research Board
TTC	Time to collision
TTCa	Time to collision (speed plus acceleration based)
TTCs	Time to collision (speed based)
U.S. DOT	U.S. Department of Transportation
vpm	vehicles per mile
VTTI	Virginia Tech Transportation Institute

U.S. Department of Transportation
ITS Joint Program Office-HOIT
1200 New Jersey Avenue, SE
Washington, DC 20590

Toll-Free “Help Line” 866-367-7487
www.its.dot.gov

FHWA-JPO-16-407

U.S. Department of Transportation