Calibration in Quantitative Alternatives Analysis

$$Y = f(X,\beta) + e$$

U.S. Department of Transportation
Federal Highway Administration

NOTICE

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The U.S. Government assumes no liability for the use of the information contained in this document.

The U.S. Government does not endorse products or manufacturers. Trademarks or manufacturers' names appear in this report only because they are considered essential to the objective of the document.

QUALITY ASSURANCE STATEMENT

The Federal Highway Administration (FHWA) provides high-quality information to serve Government, industry, and the public in a manner that promotes public understanding. Standards and policies are used to ensure and maximize the quality, objectivity, utility, and integrity of its information. FHWA periodically reviews quality issues and adjusts its programs and processes to ensure continuous quality improvement.

Unless accompanied by a citation to laws or regulations, the contents of this document do not have the force and effect of law and are not meant to bind the public in any way. This document is intended only to provide clarity regarding existing requirements under the law or agency policies. This document is not legally binding in its own right and will not be relied upon by the Department as a separate basis for affirmative enforcement action or other administrative penalty.

Cover photos: Getty Images

Technical Report Documentation Page

1. Report No. FHWA-HOP-20-057	**2. Government Accession** No.	**3. Recipient's Catalog No.**	
4. Title and Subtitle Calibration in Quantitative Alternatives Analysis		**5. Report Date** November 2020	
		6. Performing Organization Code	
7. Authors Hani Mahmassani, David Hale, Moein Hosseini, Amir Ghiasi		**8. Performing Organization Report No.**	
9. Performing Organization Name and Address Leidos 11251 Roger Bacon Drive Reston, VA 20190		**10. Work Unit No. (TRAIS)**	
		11. Contract or Grant No. Contract No. DTFH61-16-D-00053, T-0006	
12. Sponsoring Agency Name and Address Federal Highway Administration U.S. Department of Transportation 1200 New Jersey Avenue, SE Washington, DC 20590		**13. Type of Report and Period Covered** Technical Report	
		14. Sponsoring Agency Code HOP	
15. Supplementary Notes The Contracting Officer's Technical Representative was John Halkias			
16. Abstract In an era of emerging vehicle automation technologies and advanced traffic management strategies, traffic simulation has become an indispensable tool for giving agencies the confidence they need for adoption and implementation. Likewise the importance of calibration cannot be overstated, because there is no other way to ensure reliability of the simulation results. Current practice calls for analysts to calibrate their analytical tools to a base (or existing) condition, and then use those tools to predict performance of a future condition. However, many times these future conditions incorporate improvements that are significantly different than the base condition modeled when the analysis tool was calibrated. This can inhibit the accuracy of the calibrated model. Development of new calibration methods could allow the tools to be calibrated to data that are reflective of what the future condition will be. The Calibration in Quantitative Alternatives Analysis Primer proposes and describes the following five major components of the framework: scenarios, robustness, parameter libraries, local density, and the role of vehicle trajectories. Although developed with future conditions in mind, application of the framework would also lead to improved calibration of existing conditions. The Primer includes chapters on the framework underpinnings, case studies, and step-by-step instructions for different analysis types. To some extent the proposed framework could be applied with existing software, but future development of intermediate tools is recommended, for improved efficiency and practicality. Follow-on work will faciliatate development of more detailed recommendations, parameters, model forms, and tools. The ultimate result will be more accurate traffic analyses, leading to improved trust in analysis tools, and improved transportation decision-making overall.			
17. Key Words Calibration, scenario, robustness, parameter library, local density, trajectory		**18. Distribution Statement** No restrictions.	
19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No of Pages** 140	**22. Price** N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized.

SI* (MODERN METRIC) CONVERSION FACTORS

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 1,000 L should 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

*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)

SI* (MODERN METRIC) CONVERSION FACTORS (continued)

APPROXIMATE CONVERSIONS TO 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)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES (CONTINUED)

LIST OF FIGURES (CONTINUED)

LIST OF FIGURES (CONTINUED)

LIST OF TABLES

LIST OF ACRONYMS

ARTEMiS	Analysis of Road Traffic and Evaluation by Micro-Simulation
AV	automated vehicle
CORSIM	CORridor SIMulation
CV	connected vehicle
DBF	directed brute force
DLC	discretionary lane changing
DOT	department of transportation
DYNASMART	DYnamic Network Assignment-Simulation Model for Advanced Road Telematics
FHWA	Federal Highway Administration
FM	forced merging
GA	genetic algorithm
GEH	Geoffrey E. Havers (statistical measure)
GPS	Global Positioning System
GRE	global relative error
IDM	intelligent driver model
LMRS	Lane-Changing Model with Relaxation and Synchronization
MANE	mean absolute normalized error
MITSIMLab	MIT Intelligent Transportation Systems sIMulation Laboratory
MLC	mandatory lane changing
MOBIL	Minimizing Overall Braking Induced by Lane changes
MOE	measure(s) of effectiveness
MULTITUDE	Methods and tools for supporting the Use caLibration and validaTIon of Traffic simUlation moDEls
NGSIM	Next Generation Simulation
O-D	origin destination
pc/hr/ln	passenger cars per hour per lane

LIST OF ACRONYMS (CONTINUED)

RMSE	root mean square error
SPSA	simultaneous perturbation stochastic approximation
TAT	Traffic Analysis Tools
TrEPS	Traffic Estimation and Prediction System
V2I	vehicle-to-infrastructure
V2V	vehicle-to-vehicle
VANET	vehicular ad-hoc networks
veh/km/lane	vehicles per kilometer per lane
VOR	value of reliability
VOT	value of time
WAF	weather adjustment factor

EXECUTIVE SUMMARY

Models defined in simulation tools typically entail mathematical relations and algorithmic procedures that take certain variables as input to produce the output variables of interest. The relations and procedures are typically specified up to a set of parameters whose values are estimated for the context at hand. The process of estimating the model parameters, and otherwise determining the values of parameters that govern various aspects of the model to best represent the given area and context is referred to as calibration.

Current practice typically assumes that parameter values obtained using observations of current conditions will remain applicable under different future conditions. However, the values of parameters may be subjected to changes under future conditions. This diminishes the capability of the simulation tools to produce realistic predictions of the impact of various policies and interventions under future conditions. The objective of this study is to present a calibration framework that helps traffic models to produce meaningful and reliable predictions of the system performance under future conditions and contemplated policy/operational interventions. The scope of this study is to provide information for transportation professionals on how to calibrate analytical/simulation models and tools to data that are reflective of the target future conditions.

Chapter 1 defines the problem, its objective, and its challenges. The general approach that some previous studies have utilized to overcome the challenges and develop models that can capture the behavior of simulation agents under future conditions are discussed. Chapter 2 provides an in-depth analysis of previous studies on the calibration of traffic simulation models. Different studies on calibration such as the MULTITUDE (Methods and tools for supporting the Use caLibration and validaTIon of Traffic simUlation moDEls) project are discussed. The chapter includes information on popular fitness functions used for the purpose of calibration. Furthermore, some of the previous studies are discussed that attempted to address the challenge of producing reliable calibrated models. These studies were grouped in order to shape the main components of the proposed calibration framework including scenario-based calibration, consideration of parameter correlation, and use of vehicle trajectories.

To develop a calibration framework that is mindful of future conditions, an assessment is made regarding limitations of the traffic analysis tools and the data they depend on. The tools require many different layers of input data. Some categories of input data are fundamentally more precise than others. Other data may only have high precision when collected or obtained properly. On top of this, the forecasting method(s) being used adds another potential obstacle to the precision that is necessary for calibration. Moreover, the accuracy, precision, and availability of some data are rapidly changing over time. Chapter 3 contains an assessment of the data required for calibrating transportation simulation models and the limitation associated with the data. The role of trajectories, as a major source of understanding the behavior of drivers (and other actors in the transportation network), in the proposed framework is elaborated. In order to account for parameter correlation, a data structure is introduced to classify the model parameters into three different categories based on the level of uncertainty associated with the parameters. Besides the parameter categorization, each parameter should go through a preprocessing and postprocessing phase.

In chapter 4, based on the review of previous studies and the assessment performed on the required data, a gap analysis is conducted to identify a set of Traffic Analysis Tool calibration needs for transportation improvement evaluations and describe the impacts of these gaps. The set of calibration needs is intended to provide a "snapshot" of potential updates that may be required for the current calibration methods. The libraries of parameters needed to support the development of calibrated models are introduced in this chapter.

Chapter 5 articulates the proposed Traffic Analysis Tool calibration methodology/framework; which will be sensitive to, and reflective of, future conditions, no matter how different it is from the base condition. The proposed methodology/framework will be designed to achieve better analysis validity for a wider range of improvement alternatives. Using the components introduced in previous chapters, a structure for the library of parameters and the method of selecting values from the libraries to generate simulation agents that capture parameter correlations are discussed. Two analysis methods, namely the scenario-based analysis and the robustness-based analysis, are suggested for the purpose of analyzing the simulation inputs and outputs to achieve a calibrated model.

Three different case studies are presented in chapter 6. The case studies emphasize the role of the main component of the framework. Furthermore, they exhibit different types of analysis (scenario-based and robustness-based) associated with the framework. Each case study was analyzed based on the step-by-step approach provided in chapter 7. The information essentially can help transportation professionals to develop robust models that present a realistic image of the study area under future conditions and various policy interventions.

The information provided in this document is intended to identify methodologies users may consider for calibrating traffic analytical/simulation models and tools to analyze data that are reflective of the target future conditions. This document is not intended to convey any endorsement by FHWA of recommended practices for specific applications, and is not intended to override or augment existing FHWA guidance for the calibration, validation, and reasonableness checking of travel and land use forecasting for project development and NEPA processes, which can be found at https://www.environment.fhwa.dot.gov/nepa/Travel_LandUse/travel_landUse_rpt.aspx.

Additionally, FHWA and the Center for Innovative Finance Support have developed a number of spreadsheet-driven modeling and sketch planning tools that may be used by transportation professionals and policy makers to estimate the effects of road pricing strategies on revenue generation, travel behavior, economic activity, and the environment. Descriptions of these tools and links to user guides and the interactive spreadsheets are available at https://www.fhwa.dot.gov/ipd/tolling_and_pricing/tools.

CHAPTER 1. INTRODUCTION

Note: Unless accompanied by a citation to statute or regulations, the practices, methodologies, and specifications discussed below are not required under Federal law or regulations.

BACKGROUND

Analytical tools and simulation models are the primary tools used to support evaluation, design and planning of operational measures and policies aimed at improving mobility and service quality offered by our transportation systems. Models vary in level of detail, geographic scope and granularity, temporal dynamics, behavioral richness and so on, usually tailored to the needs of a particular study or application. However, in all cases it is expected that the models provide a reasonable and realistic representation of reality that is appropriate for the application at hand. At their core, models typically entail mathematical relations and algorithmic procedures that take certain variables as input to produce the output variables of interest. The relations and procedures are typically specified up to a set of ***parameters*** whose values are estimated for the context at hand. The process of estimating the model parameters, and otherwise determining the values of parameters that govern various aspects of the model so as to best represent the given area and context is referred to as ***calibration***. Calibration methods vary in rigor and complexity depending on the nature of the models and the available data. Calibration inherently requires observation of the existing system conditions, as the analyst seeks to estimate the model parameter values so as to best match the available observations.

Current practice typically assumes that parameter values obtained using observations of current conditions will remain applicable under different future conditions, which may thus be modeled or simulated by providing input values reflecting these future conditions. Whether they describe mathematical relations between physical quantities (e.g., fundamental diagrams at a macroscopic level) or characteristics of individual behavior (e.g., driver risk aversion or traveler preferences at a microscopic level), parameters values may be different under these future conditions than they are under the conditions for which the model was calibrated. This is especially true when new technologies are introduced (new modes, new fuels or propulsion technologies, connected and/or automated vehicles), or major changes are implemented in the operation of the transportation system. To the extent that models are developed to evaluate contemplated interventions and measures to be implemented in the future, rather than merely replicating present and past conditions, it is essential that the models have the capability to produce realistic predictions of the impact of these interventions under potential future conditions. In addition to being able to represent the changes through the variables incorporated in the model formulation/specification, it is highly desirable that the parameter values used in prediction be reflective of the future conditions under evaluation. Unfortunately, typical calibration methods require actual observations of the conditions of interest, and hence are generally possible only for conditions that may be observed. To the extent that the future conditions have not yet occurred, it is obviously challenging to perform a similar calibration exercise for future conditions. Addressing this challenge requires changes not only in calibration procedures and practices, but also in model formulation and development. The main objective is for traffic models to produce meaningful and reliable predictions of system performance under future conditions and contemplated policy/operational interventions.

Accordingly, different calibration methods need to be developed so that the tools are calibrated to produce results that are reflective of what the future condition will be. The scope of this study is to provide information for transportation professionals on how to calibrate analytical/simulation models and tools to data that are reflective of the target future conditions. The developed methodology/framework will allow agencies and transportation professionals to develop and calibrate analytical/simulation tools that are valid for a wide range of different future conditions.

OBJECTIVE

Traffic analysis tools (TATs) are designed to assist transportation professionals in evaluating the transportation improvements that best address the transportation needs of their jurisdiction. TATs can help practitioners improve the decision-making process, evaluate and prioritize improvement alternatives, and improve project design and operations. Over the past 15 years, the Federal Highway Administration (FHWA) TAT Program has developed the TAT Toolbox, a compendium of TAT informational documents (see https://ops.fhwa.dot.gov/trafficanalysistools/index.htm). The TAT Toolbox has helped to establish consistency in practice for traffic analysts across the nation.

Calibration is a key step in the application of TATs to a project or study. Calibration is the adjustment of model parameters to improve the model's ability to reproduce local driver behavior and traffic performance characteristics. The importance of calibration cannot be overemphasized because no single model can be expected to be equally accurate for all possible traffic conditions.

Current practice calls for analysts to calibrate their analytical tools to a base (or existing) condition and then use those tools to predict performance of a future condition. This practice is consistent with information currently in the TAT Toolbox. However, many times these future conditions incorporate improvements that are significantly different than the base condition modeled when the analysis tool was calibrated. This can inhibit accuracy of the performance outcomes.

Different calibration methods are developed so that the tools are calibrated to data that are reflective of what the future condition will be. The scope of this report is to provide information for transportation professionals on how to calibrate analytical tools to data that are reflective of the conditions that the tool is being used to predict future performance. Users should be provided with an understanding of how to amass, understand, prepare, and analyze a wide range of future conditions that impact performance. The methodology/framework that will be presented in this report will allow agencies and transportation professionals to calibrate their analytical tools to enable validity no matter how different the future condition modeled is from the base condition.

The objective of this report is to describe a methodology/framework allowing traffic analysis tools calibration to account for not only base conditions, but the future conditions being modeled. This will enable validity of the analysis tool for improvement alternatives being considered. Such a methodology/framework will enable more accurate traffic analyses, which will in turn lead to improved trust in analysis tools and improved transportation decision-making overall.

CHALLENGES AND SOLUTIONS

Addressing the above objective entails several challenges that are unique to this effort. The challenges arise from the very nature of the problem itself, which entails determining model parameter values for conditions that do not yet exist, and hence my not be observable, either directly or indirectly. Below are selected challenges and proposed mitigation approaches.

Table 1. Challenges and solutions for calibration to account for future conditions.

Anticipated Challenge	Solution/Mitigation Approach
Uncertainty in future values of the input variables is greater than uncertainty in the model parameters.	Develop a scenario-based approach to address model development and calibration challenges, and ensure that model parameters would be applicable no matter the future conditions. The idea is to limit the error in conditional forecasts for particular scenarios.
Difficult to establish benchmark against which to evaluate performance of the models' predictive capabilities.	Hindsight (after the fact) would not be a fair benchmark in this case. The research team will draw on its methodological expertise to devise a systematic approach for comparing the performance of different predictors and calibration methods.
There is no holy grail—cannot observe future conditions before they occur!	The research team's approach has identified a variety of ways for using models in a way to inform evaluation for future conditions.

CALIBRATION NEEDS

Calibration methods vary in rigor and complexity depending on the nature of the models and the available data. Statistical or econometric estimation is the most common method for finding parameter values that provide a best fit to the available data. Other optimization approaches with varying degrees of formalism may also be used to make the model outputs best match available data on the system of interest. Conceptually, in simple terms, the model output Y can be formulated as a function of input variables X, and a set of parameters β that characterize that relation, i.e.,:

The process of estimating the model parameters, and otherwise determining the values of parameters that govern various aspects of the model so as to best represent the given area and context is referred to as calibration.

$$Y = f(X,\beta) + \varepsilon$$

Figure 1. Formula. Relationship between inputs, outputs, and model parameters.

Where ε is an error term that is the discrepancy between the actual value of the output Y and the one calculated using the function $f(.)$ The calibration problem, in its simple form, is to find the values of β, given joint observations of the input and output variables (X,Y). Calibration, inherently, requires knowledge and observation/measurement of actual conditions in order to be meaningful. Application of the model for prediction generally entails the input of new values of X (say at some future date), and calculation of the corresponding Y using the function with calibrated parameter values. What is typically assumed is that the relation established for "current" or "past" conditions will continue to hold in the future, including the values of the model parameters β. The only elements assumed to change when the model is applied to forecast future conditions consist of the values of X. These are typically forecast independently, or determined as the result of a particular intervention.

In what follows, future conditions, with or without intervention, are referred to as a ***scenario***. A scenario is defined by set of operational conditions (reflecting external events, such as weather, demand surges, etc.), interventions (infrastructure changes, control actions, dynamic system management schemes, etc.), as well as characteristics of the general activity system (land use, activity locations) and associated technologies (e.g., connected and autonomous vehicles, Internet of Things, smart cities). The features and characteristics of a given scenario are represented through the values of X in the above simple model.

In general, the longer the horizon over which a scenario is defined, i.e., the further out into the future one is trying to forecast, the greater the uncertainty associated with the values of X. This is especially true not only for the activity system variables and the associated technologies, but also the social fabric and associated preferences/lifestyles/ norms. Traffic planners generally deal with the uncertainty in the future values of X by defining alternative future scenarios corresponding to different rates of population and economic growth, alternative assumptions on urban core densification vs. sprawl, and so on. However, in modeling the impacts of these changes for a given scenario, the values of model parameters β are assumed to remain constant. There are of course many reasons to assume that these parameters would change due to the same forces that are influencing the values of X, but also cases where these parameters may remain stable and still applicable to the new scenario conditions.

Addressing this challenge involves a significant shift in the mindset of traffic modelers—from model estimation and calibration aimed at replicating existing and past conditions, to greater emphasis on prediction quality and accuracy.

The main problem addressed in this document is how to develop and calibrate models so that the output they produce for a future scenario is meaningful, and provides a realistic and accurate depiction of future conditions defined by the scenario of interest. Note that model development and calibration are intrinsically related from the standpoint of the model's ability to predict future scenario conditions. It is important that the model be responsive to the scenario features and associated interventions, i.e., that it allows the analyst to represent these features in the model—this is an issue of model specification. As a pre-requisite, the specification should be responsive to future scenarios, and include the appropriate X variables in the model.

Given the specification, the main challenge addressed is how to ensure that the parameter values used in applying the model to predict traffic system performance under the future scenario of interest are adequate, i.e., that the model is "calibrated for future conditions." Of course, since these future conditions have not yet occurred at the time the model is used, it is not possible to directly observe actual conditions for such scenarios as a basis for model calibration. Addressing this challenge requires a significant shift in the mindset of traffic modelers—from model estimation and calibration aimed at replicating existing and past conditions, to greater emphasis on prediction quality and accuracy.

CHAPTER 2. REVIEW OF CALIBRATION METHODOLOGIES

Note: Unless accompanied by a citation to statute or regulations, the practices, methodologies, and specifications discussed below are not required under Federal law or regulations.

States are calibrating their traffic analysis tools to existing conditions, and then using those same calibration settings with future projected demand volumes to predict future performance. This practice is reflected in the state-specific modeling and simulation informational manuals published by various states. Moreover, there is little apparent variance in the calibration approaches when analyzing emerging technologies (e.g., advanced transportation demand management, integrated corridor management, connected and automated vehicles). As discussed in the first chapter, future conditions incorporate improvements that are significantly different than the base condition modeled when the analysis tool was calibrated. This can inhibit accuracy of the performance outcomes.

In recent years, there has been a growing interest in developing frameworks for calibration procedures that could be applied to different spatiotemporal settings. Such developments require a thorough understanding of the fundamental concepts underlying model calibration, as well as related validation principles. Calibration and validation procedures have tended to vary widely in practice, reflecting variation in available data and resources, as well as in the purpose of the motivating applications. In addition, there appears to be little consensus in the published literature and consultant informational documents regarding concepts, methods and best practice, including appropriate measures of effectiveness and calibration metrics. Many academic efforts have been devoted to modifying the optimization-based methods used for performing the calibration. As a result, the literature still lacks a systematic approach for calibration (So, et al. 2016)

Wegmann and Everett (2012) mentioned that the even the meaning of "calibration" and "validation" have changed over time, especially where funding limitations preclude the possibility of conducting travel surveys (for planning models) and/or taking needed measurements (for operational models). Under such limitations, default values are used for the model parameters in lieu of calibration using local data. Furthermore, lack of an independent database results in the calibration and validation processes becoming a single task. They provided information for calibrating each step of the transportation four-step model. For each step, they provided a list of measures along with candidate values and thresholds that could be used to evaluate the model.

As another attempt in this area, the MULTITUDE (Methods and tools for supporting the Use caLibration and validaTIon of Traffic simUlation moDEls) project, funded by the European Commission, has mainly focused on calibration methodologies and tried to bridge the gap between academic research and industry practice. According to a survey conducted in 2011 by Brackstone et al. (2012) among practitioners in the transportation industry, 19 percent of the respondents performed no calibration on the models that they developed and used. Moreover, among all those who performed the calibration process, only 55 percent followed sample steps provided in an informational manual instead of personal experience. Another finding is that calibration and validation have received more attention in the travel demand modeling area than the traffic flow simulation area.

The MULTITUDE document discusses several issues related to performing traffic simulation projects, including calibration:

- Definition of calibration.
- Structuring a simulation calibration activity.
- Calibration methodologies.
- Sensitivity analysis and the way to perform it.
- "Fall-back strategies" when calibration data are not available, and data transferability over various calibrations.
- Effect of model parameters considered in the calibration.

There have been numerous studies on the calibration of traffic simulation models. These studies differ from one another in different aspects including choice of algorithm, metrics used to compare the field-collected data and simulation model results, objective function utilized throughout the calibration process, parameters selected to calibrate, and choice of software.

Calibration of microscopic traffic behavior model parameters poses special challenges related to the difficulty of obtaining closed form expressions of the objective functions typically maximized or minimized in commonly used estimation techniques, such as generalized least squares or maximum likelihood. Thus estimation typically requires use of simulation or other numerical techniques to evaluate the objective function in the context of the estimation process. Furthermore, the lack of closed form expressions often precludes establishing rigorous mathematical properties that ensure convergence and uniqueness of the solution. For this reason, a variety of search techniques such as metaheuristics have been used in connection with this problem. Several studies in the literature have focused on one or the other algorithm for this purpose. Among these, genetic algorithm (GA) and variants (Cheu, et al. 1998, Kim, Kim and Rilett 2005, Ma and Abdulhai 2002, Park and Qi 2005, Chiappone, et al. 2016, Menneni, Sun and Vortisch 2008) and simultaneous perturbation stochastic approximation (SPSA) (Paz, et al. 2015, Balakrishna, et al. 2007, Lee and Ozbay 2009) are widely used in the literature. Some studies also performed a comparison of the performance of different algorithms. For instance, Ma et al. (2007) performed a comparison of the SPSA, GA, and a trial-and-error iterative adjustment algorithm over a portion of the SR-99 corridor in Sacramento, California. The comparison showed faster convergence of the SPSA algorithm; however, the other two algorithms reached a more stable set of calibrated parameters. Unfortunately, it is difficult to generalize from one set of comparisons as the results of such techniques often depend on the characteristics of the specific data set and model forms used.

Due to the high dimensionality of the problem, multiple local optima might exist for the problem. Accordingly, since the algorithms used in the literature are heuristic methods, a global optimum solution could not be guaranteed. Therefore, the calibrated parameter set found for the same problem might be different when a different heuristic algorithm and a different starting parameter vector are used.

The common set of metrics used in the calibration procedure are based on comparing simulated values (with the estimated parameters) to the observed values of traffic volume/flow, speed, capacity, density/occupancy, travel time, and/or origin-destination (O-D) flow. All these quantities could be derived if there is enough information regarding the observed and simulated vehicle trajectories.

The following formulas are some of the fitness functions used as the objective function for calibrating the simulation models (Yu and Fan 2017).

$$MANE = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M} w_j \frac{\left|m^j_{obs,i} - m^j_{sim,i}\right|}{m^j_{obs,i}}$$

$$GRE = \sum_{j=1}^{M} w_j \frac{\sum_{i=1}^{N}\left|m^j_{obs,i} - m^j_{sim,i}\right|}{\sum_{i=1}^{N} m^j_{obs,i}}$$

$$PMAE = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M} w_j\left(\left|m^j_{obs,i} - m^j_{sim,i}\right|\right)$$

$$PMRE = \sum_{j=1}^{M} w_j \sqrt{\frac{1}{N}\sum_{i=1}^{N} w_j \left(\frac{m^j_{obs,i} - m^j_{sim,i}}{m^j_{obs,i}} \times 100\right)^2}$$

$$GEH = \sqrt{\frac{2\left(vol_{obs,i} - vol_{sim,i}\right)^2}{vol_{obs,i} + vol_{sim,i}}}$$

Figure 2. Formula. Fitness functions used as the objective function in calibration.

In figure 2 *MANE* is the mean absolute normalized error; *GRE* is the global relative error; *PMAE* is the point mean absolute error; *PMRE* is the point mean relative error; and *GEH* is a statistical measure named after Geoffrey E. Havers. Within the formulas, *N* is the total number of observations; *M* is the total number of measurements considered for comparing the simulation results and the observations; $m^j_{obs,i}(m^j_{sim,i})$ denotes the ith sample value of the observed (simulated) metric j; $vol_{obs,i}(vol_{sim,i})$ represents the ith sample observed (simulated) traffic volume. Last but not least, w_j is the weight assigned to each metric. Usually an equal weight is selected for each metric unless there is an inconsistency between the units of the components of the fitness function, such as the ones in the PMAE formula, or there is a situation that the researcher values one metric more than another. For instance, Paz et al. (2015) used a PMRE formulation for their objective function with different weights assigned to the link vehicle counts and link speeds.

So et al. (2016) presented a step-by-step approach for building, calibrating, and validating a microscopic model. In the calibration phase of their methodology, they first adjust the traffic flow descriptors (traffic counts, desired speed decisions, and reduced speed areas) and car-following parameters (maximum deceleration and acceleration rates, and minimum headway). Then the process goes through a calibration loop where the volume, speed, and travel time are checked sequentially based on their respective criteria. The modification of the model parameters are performed until the coefficient of determination for the turning movement traffic counts (R^2) is more than 0.8 at all intersections. As the second step in the calibration process, the speed distributions on all links and the reduced speed areas in the simulation model are adjusted. Finally, the travel time is verified so that the R^2 is more than 0.8 on all network sections. Since these three sequential steps might have a contradicting effect, a final calibration is performed considering all three goals simultaneously. Unfortunately, many values of the parameters to be calibrated could satisfy such aggregate criteria.

They tested the model on a large network consisting of 6 arterials and 160 intersections. This is one of the limited examples that a microsimulation calibration was performed on a relatively large network. One of the disadvantages of modeling such a large network is the high computational time required to develop and calibrate the model. Although they reached a significant accuracy in the calibration process, the validation process was only partially successful. They speculated that one of the reasons might be seasonal differences in the traffic demand distribution.

An innovation in the model calibration area is a software developed by Hale (2015) that has automated the calibration process to some extent. The software, with a published patent titled "System and Method for Automated Model Calibration, Sensitivity Analysis, and Optimization," intends to make calibration a faster, cheaper, and easier procedure that requires less engineering expertise. Based on the software algorithm, the user selects the parameters and outputs for calibration as well as the level of calibration for each parameter. The software then performs the calibration and displays the field data metric, the simulated metrics, and the recommended values for each parameter. The software-assisted calibration architecture pertains to a simulation-based optimization category of simulation models. It contains a predefined narrow set of trial values for the input parameters. Similar to assigning a weight for each metric in other studies, the software is capable to prioritize the metrics evaluated in the objective function and the input parameters. Besides calibration, the software can be used to perform sensitivity analysis and solve a network optimization problem. Although the calibration process could conceivably be streamlined by deploying such software, the method has never been licensed to a commercial tool.

Hale et al. (2015) compared two optimization methods using the calibration software, namely directed brute force (DBF) and the aforementioned SPSA. They used CORSIM to model the I-95 corridor near Jacksonville, FL. The study period was divided into twelve 15-minute intervals. After calibrating the model using both optimization methods, they found that SPSA optimized the system more quickly. However, the effectiveness of SPSA relies on the suitability of the internal parameters and starting point. Moreover, the algorithm performs better when trial values are used instead of

optimum (theta) values. Even with all these considerations, the SPSA algorithm usually provides local optimum solutions. On the other hand, the DBF algorithm would find the global optimum solution if the solution is within the user-defined search space. Furthermore, DBF is an appropriate algorithm for conducting sensitivity analysis. The authors concluded that a combination of the algorithms could be used to solve a calibration-related optimization problem. More specifically, the SPSA could be used to quickly find a local optimum solution then this solution could be fed as starting point to the DBF algorithm to locate the global optimum solution.

In summary, there have been many efforts in calibrating simulation models; these efforts share common elements in terms of the base model, the optimization algorithm, formulation of the objective function, etc., but differ in the specific choices made for these elements. One of the main limitations in existing studies is that the calibration process is based on available data, which mostly represent the normal prevailing condition of the transportation system. Some studies identified and specified additional scenarios developed on the basis of historical data. Offline calibration performed based on these datasets may not be robust enough to capture the future behavior of the system, especially when major changes are anticipated. Transferability of these models to future conditions of a system that might be subjected to fundamental changes in both demand and supply sides could be questioned. Furthermore, models that are calibrated online could only perform in a reactive manner. This reduces the effectiveness of the strategies suggested by the system when an anomaly occurs in the system that is not realized in the historical data under consideration.

Traffic is a complex system, and models developed to capture and predict its behavior often have many parameters that are estimated as part of the calibration process. Frequently, these parameters exhibit correlations across individuals, as they are intended to capture human behavior. Most of the previous studies performed the calibration procedure without considering the correlations that might exist between various parameters of a model. Ignoring these effects could result in solutions that are inconsistent with reality. Furthermore, as discussed later, considering these correlations, one could simplify the calibration process by reducing the dimensions of the pertinent optimization problem.

Different measures of effectiveness (MOE) have been selected in various studies presented in the literature. These are typically measured at fixed locations, and presented as aggregates over space and time. Calibration of microscopic behavioral models has often entailed specialized data collection with instrumented vehicles, or image processing of videos taken from fixed locations with limited visual field. An alternative that is gaining popularity is the use of vehicle trajectories (e.g., via Global Positioning System (GPS) tracking), which provide a comprehensive record of individual vehicle movement through a network. As such, trajectories could be used as an input from which various metrics/descriptors at various levels of detail could be extracted. Trajectories could be directly related to the demand parameters since they provide information regarding the origin and destination of a trip, though they do not convey information on other behavioral aspects associated with travel.

SCENARIO-BASED CALIBRATION

One of the approaches that would help calibration of transportation simulation models for anticipated/unexpected future conditions is a scenario-based simulation. There are many conditions that could impact a transportation system and induce variability in the model outputs. For example, major causes of variability in travel time could be traffic incidents, work zones, weather, special events, traffic control devices, demand fluctuations, and insufficient base capacity (J. Kim, H. Mahmassani and P. Vovsha, et al. 2013). Figure 4 shows a schematic of possible scenarios that might impact operation of a transportation network. It is essential to consider these sources of variability in the calibration process to develop robust models that could capture the relevant behavior of a transportation system.

Hou et al. (2013) developed a systematic procedure for calibration of a weather-sensitive Traffic Estimation and Prediction System (TrEPS). TrEPS is a tool that not only is able to capture the behavior of the system under different weather conditions, but can also implement weather-responsive strategies in the traffic management system. In the absence of relevant data, one can use the library of parameter values derived from TrEPS to construct some inputs of a dynamic network simulation model. This weather-sensitive system was integrated into a dynamic traffic assignment simulation framework, DYNASMART-P (DYnamic Network Assignment-Simulation Model for Advanced Road Telematics). The weather effect was formulated as a coefficient:

$$WAF_i = \frac{f_i^{weather\ event}}{f_i^{normal}}$$

Figure 3. Formula. Weather effect coefficient in DYNASMART-P.

Where WAF_i is the weather adjustment factor for parameter i, and $f_i^{weather\ event}$ and f_i^{normal} represent the value of parameter i under a certain weather condition and under normal condition, respectively. After deriving the factor corresponding to each parameter, the factor is related to weather-related variables through a regression analysis. The coefficients of these regression models and the weather adjustment factors constitute a library that could be used in similar studies.

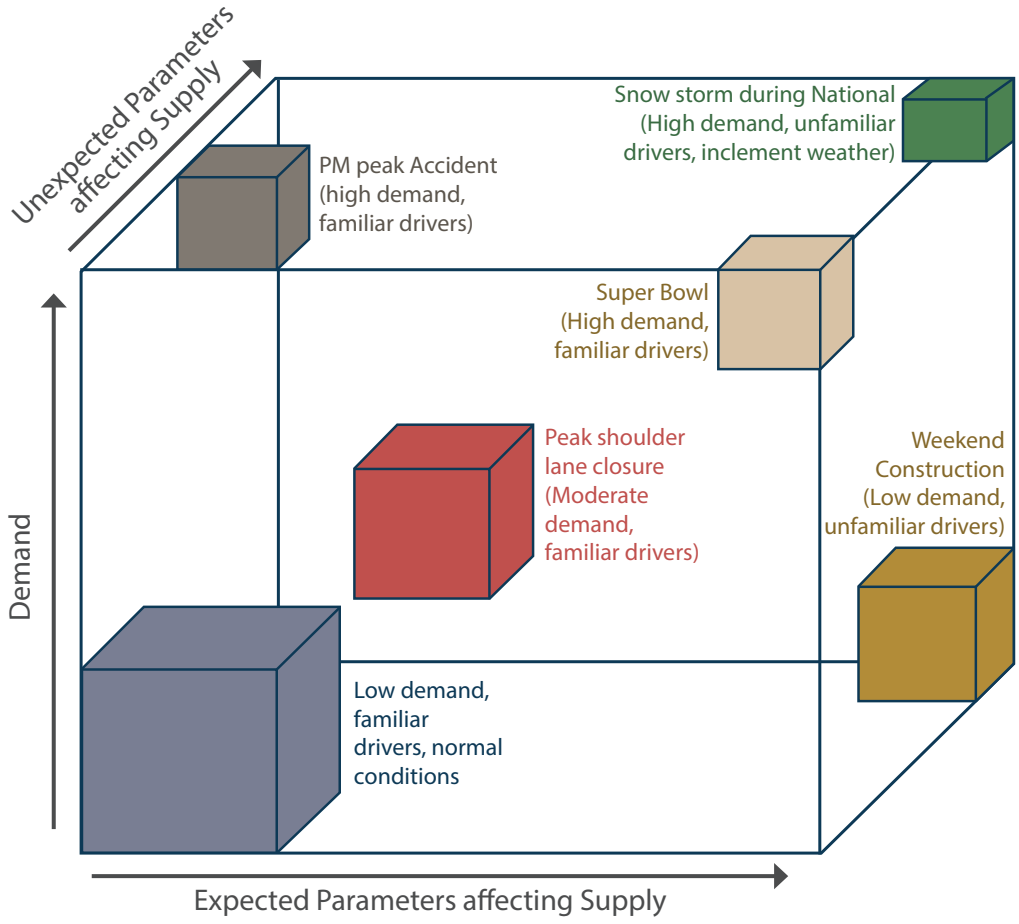


Figure 4. Diagram. Schematic representation of different conditions that impact a transportation network.

The methodology was applied to the data collected for four metropolitan areas over several years to test the robustness of the methodology in calibrating a dual-regime modified Greenshields speed-density model under different weather conditions. Models developed for the case studies were calibrated through an iterative procedure that minimizes the root mean square error (RMSE) of speed. The results show that incorporating the weather adjustment factor in the model leads to a more realistic representation of the traffic conditions. In addition, a significant effect of visibility and precipitation intensity on some parameters of the fundamental diagram such as free flow speed and maximum flow rate was exhibited.

Kim et al. (2013) proposed a conceptual framework that uses commonly used traffic simulation models to capture the probabilistic nature of travel time. The framework is comprised of a scenario manager, a traffic simulation model, and a trajectory processor. The exogenous and endogenous sources of travel time variations are controlled respectively in the scenario manager component and in the traffic simulation model component. The last component of the framework processes the trajectories created by the simulator to measure the travel time for each scenario. This component also makes this study relevant to the trajectory-based calibration topic which is further discussed in the last section. The results of this framework could be translated into reliability measures for the urban network.

The methodology starts with determining the distribution of the demand and the equilibrium between demand and supply. Then, the scenario is specified for the distribution of scenario-specific parameters. The scenario is selected from the distribution of scenarios using a combination of the Monte Carlo and the mix-and-match approaches. After running the simulation model, the trajectories are developed. Then, the travel time distribution could be derived using the trajectories. Moreover, other reliability measures such as standard deviation of travel time, different percentile values of travel time, buffer index, misery index, planning time index, planning index, travel time index, and the probability of on-time arrival could be calculated using the vehicle trajectories. In order to calibrate the model for each scenario, an inner loop controls the demand-side scenario parameters and an outer loop adjusts the supply-demand equilibrium. The framework was tested on the data collected from 27.5 mi of the Long Island Expressway during the morning peak hours within a winter season. The results showed that the model parameters were consistent with the assumed scenario after the calibration procedure was performed.

Besides all the known sources of variability in transportation systems that might be reflected in historical data, there are conditions that have not been realized yet but are expected to be seen in the future. To improve the accuracy of the models in forecasting the network behavior under these situations, there is a need to generate these scenarios and develop reasonable probabilistic distributions for them.

Scenario-based approaches offer an important direction for making traffic models ready for a variety of future conditions.

CORRELATED PARAMETERS

A possible source of error in previous calibration-related studies is that they did not consider the correlation between the parameters being estimated. Some studies investigated the effect of these correlations on the final results and proposed some methods to tackle this issue.

Kim and Mahmassani (2011) studied the effect of ignoring correlations among parameters when a car-following model is calibrated. The authors selected three car-following models including the Gipps model, the Helly linear model, and the intelligent driver model (IDM). Six metrics were chosen to compare the effect of parameter correlation on different objective functions defined for the calibration procedure: network exit time, total travel time, mean of average spacing, the standard deviation of average spacing, mean of the coefficient of variation of spacing, and standard deviation of coefficient of variation of spacing. They evaluated the correlation effect using the Next Generation Simulation (NGSIM) trajectory data. First, the correlations between parameters of the dataset were studied using the factor analysis approach. To investigate the parameter correlation effect, they developed two sets of simulation models. One by sampling from the marginal distribution of the model parameters (ignoring the correlation effects) and the other by sampling from the joint distribution of the model parameters (preserving the correlation effects). Then, the models were calibrated by solving the optimization problem using the downhill simplex method. The distributions of the results were compared to the relevant distribution developed based on the NGSIM data using the Kolmogorov–Smirnov test. The comparison showed significant discrepancies between the result of the calibrated models that correctly considered parameter correlations and the calibrated models that ignored parameter correlations. As expected the results of the models that preserved the correlation effect were closer to the field data.

Prakash et al. (2018) suggested a generic calibration methodology to perform dynamic traffic assignment in real time. As mentioned earlier, one of the limitations associated with the state-of-the-art approaches in calibration is that they are spatially limited to a small network or they are applied to a limited set of parameters, incorporating both demand-side and supply-side parameters. The authors addressed these limitations by systematically reducing the dimension of the problem. They proposed a state-space formulation and solved the problem using the Constrained Extended Kalman Filter approach. In order to reduce the problem dimension, they performed a principal component analysis. The following formula shows the state-space formulation used for the generic online calibration problem:

$$\Delta\pi_h = \sum_{i=h-q}^{h-1} F_i^h \, \Delta\pi_i + \eta_h$$

$$\Delta M_h = S\left(\Delta\pi_h + \pi_h^H, \dots, \Delta\pi_{h-p} + \pi_{h-p}^H\right) - S\left(\pi_h^H, \dots, \pi_{h-p}^H\right) + \zeta_h$$

$$\Delta\pi_i = \pi_i - \pi_i^H$$

$$\Delta M_i = M_i - M_i^H$$

Figure 5. Formula. State-space formulation for generic online calibration.

Where π_i (M_i) represents the model parameters (measurements) in the time interval i. On the other hand, π_i^H (M_i^H) contains the historical values of the parameters (measurements) in time interval i; η_h (ζ_h) is a vector of random errors in the transition (measurement) equation; F_i^h is a matrix that relates the parameter estimates of interval i to the estimates of interval h; S(.) denotes the simulation model; q is the degree of the autoregressive process; and p denotes the maximum number of previous time interval's parameters which influence the current interval's parameters.

The authors listed three main assumptions of the methodology as follows: using time-invariant principal component directions, the linearity of transition equation, and performing a sequential online calibration that neglects the effect of estimated parameters on the previous time intervals. However, these simplifying assumptions do not sacrifice the accuracy of the model significantly.

They applied the methodology to a large Expressway network in Singapore that constitutes 936 nodes, 1,157 links, 3,906 segments, and 4,121 O-D pairs. The time required for estimating the calibrated parameters of the model for such a large network based on this methodology was less than 5 minutes. Followed by the model calibration, they performed a sensitivity analysis to study the effect of the degree of dimensionality reduction on the accuracy of the model.

As mentioned previously, there are several ways to consider the correlation that might exist between parameters that are considered in the calibration procedure. Recognizing the correlation could help reduce the size of the problem and develop more reliable models. With the advent of new technologies and fundamental transitions in travel behavior, it is important to study the correlation between parameters that would be introduced to the transportation system and already existing parameters to improve the performance of simulation models.

TRAJECTORY-BASED CALIBRATION

Vehicle trajectories are becoming an increasingly available form of observation of individual traffic and traveler movement in networks, largely due to the prevalence of GPS-enabled cell phones and navigation systems in vehicles. Trajectories constitute records of vehicle presence at various points along a network.

Information that can be extracted from trajectory data includes (H. S. Mahmassani 2016):

- Time; i.e., position of this moment on the timescale.
- Position of the vehicle in space.
- Trip origins and destinations.
- Direction of the vehicle‘s movement.
- Speed of the movement.
- Dynamics of the speed (acceleration/deceleration).
- Accumulated travel time and distance.
- Individual path and temporal characteristics.

In addition, from groups of trajectories, the following could be obtained:

- Distribution of speed/travel time (for reliability analysis).
- Probe vehicle density.
- Inferred traffic volume.

Trajectories provide a more complete and compact description of system state than fixed sensors, as they capture all aspects of individual actions (most complete record of actual behavior), with no loss of ability to characterize systems at any desired level of spatial and temporal aggregation or disaggregation. Furthermore, they retain the ability to extract stochastic properties of both individual behaviors and performance metrics (Kim and Mahmassani 2015).

Trajectories are also a common output of particle-based (microscopic and mesoscopic) simulation tools is one of the model outputs that could be used to generate other system metrics at a different level of aggregation. The level of aggregation could vary from reflecting the performance of the system in a small segment of the road to summarizing the performance of the entire network in a single number (Saberi, et al. 2014). As a result, they enable better model formulation/specification at all levels of resolution, effectively unifying model calibration and network performance analysis.

Trajectories have played a key role in the calibration of microscopic models of car following (Hamdar, et al. 2008) and lane changing (Talebpour, Mahmassani and Hamdar, 2015). Ossen and Hoogendoorn (2008) mentioned that calibration of car-following models could be impacted by methodological issues, practical issues related to the use of field data, and measurement errors. The authors investigated the effect of measurement error on calibration results. To grasp a better understanding of this effect, they introduced several types of measurement errors to a synthetic

dataset. They simulated the trajectory of the follower based on the given trajectory of the leader and a selected set of parameters using the Gipps model and the Tampere models. As the first step, they used the Bosch trajectory data collected in Germany as the leading car trajectory and then produced a clean synthetic dataset for the following car based on each car-following model and random sets of model parameters with uniform distribution. In the second step, the errors were incorporated in the leading car and following car datasets considering a variety of statistical distributions. Several statistical distributions were used to accommodate the effect of error magnitude, systematic errors, and nonsymmetrical errors. The calibration was performed in the next step by minimizing two separate objective functions: one based on vehicle speed and the other based on headway. The problems were solved using the simplex method. They found that in addition to the significant biases caused by measurement errors, the optimization problem could generate results that contradict the car-following behavior. Another problem caused by these errors is the reduction in sensitivity of the objective function and reliability of the results. The authors suggested smoothing the data by the moving average method to alleviate some these issues.

As a result, the vehicle trajectories as mentioned earlier are very useful in calibration of simulation models. However, due to the large effect of small errors introduced in the models that generate vehicle trajectories, care should be exercised in such analyses.

CONCLUSIONS

States are calibrating their traffic analysis tools to existing conditions, and then using those same calibration settings with future projected demand volumes to predict future performance. Unfortunately, because future conditions incorporate improvements that are significantly different than the base condition modeled, this can inhibit accuracy of the performance outcomes.

Significant research into calibration of traffic analysis tools has taken place in recent years. Thanks to these studies, we have learned more about the potential effectiveness of certain heuristic methods (e.g., genetic algorithm) and fitness functions (e.g., RMSE). However, these studies have not addressed the challenge of developing robust models for future conditions.

Other research avenues appear more promising for informing an overall methodology or framework for modeling future conditions. One of these is a ***scenario-based simulation***. To improve the accuracy of models in forecasting network behavior under various situations, there is a need to generate scenarios and develop reasonable probabilistic distributions for them. Secondly, a possible source of error in previous studies is that they did not consider ***parameter correlation***. There are several ways to consider the correlation that might exist between parameters considered in the calibration procedure. Recognizing this correlation could help reduce the size of the problem, and develop more reliable models. Lastly, the use of ***vehicle trajectories*** is gaining popularity. Trajectories provide a more complete description of system state. They retain the ability to extract stochastic properties of both individual behaviors and performance metrics. Trajectories could be directly related to the demand parameters, though they do not convey information on other behavioral aspects associated with travel.

CHAPTER 3. DATA ASSESSMENT

A fundamental element of traffic analysis tool (TAT) application is the supporting data that enables tool calibration, validation, and validity of results for evaluating the transportation improvements. Calibrating Traffic Analysis Tools to conditions that are ultimately evaluated involves data from improvements that may not be deployed yet. This challenge is magnified for evaluating emerging technologies and strategies that have not been mainstreamed into our transportation networks (e.g., advanced transportation demand management, integrated corridor management, and connected and automated vehicles). Since these emerging technologies and strategies are not yet widespread, this chapter aims to assess the practical issues and impacts relating to data, both existing and future. This chapter identifies and documents what data sources exist, or are likely to be available in the future, for the transportation improvement being evaluated. This will also include data sources that may be promising, or that may require transformation, in meeting objectives where no suitable data is anticipated to be available in the near term.

Note: Unless accompanied by a citation to statute or regulations, the practices, methodologies, and specifications discussed below are not required under Federal law or regulations.

DATA COLLECTION TECHNIQUES

The objective of calibration a traffic simulation model is to estimate the parameter values that establish the structural relations, amongst the relevant variables, that govern the behavior of the system. Validation further verifies the ability of the model to replicate and predict traffic conditions given specific inputs such as the traffic demand, operational control strategies and interventions, and external environmental factors. Errors in the collected data could influence the simulation in two different ways. First, errors in the raw data influence the outputs derived from the simulation tool. Second, estimated error-prone variables compared to simulation results could impact the optimal set of parameter values estimated as a result of the calibration process. Therefore, the model input should be carefully checked and pre-processed prior to the calibration of model parameters.

Daamen et al. (2014) classified the data used in traffic simulation studies into six groups:

- Local detector data.
- Section data (vehicle re-identification).
- Vehicle-based trajectory data.
- Video-based trajectory data.
- Behavior and driving simulation.
- Stated and revealed preferences.

The first two data groups primarily reveal macroscopic properties of the system. The last two typically address individual behavior, at the microscopic level; driving simulation data could synthetically reflect driver's behavior at a microscopic level. On the other hand, the trajectory data (vehicle-based and video-based) could be used for both microscopic and macroscopic studies.

However, the main limitations associated with the trajectory data are the limited coverage in terms of the number of vehicles equipped with tracking devices (for the vehicle-based trajectory data) and the limited spatial coverage (for the video-based trajectory data). Emerging technologies in the vehicle connectivity area both in the vehicle-to-vehicle and vehicle-to-infrastructure communication could address these limitations. Since the trajectory data is more informative than the other types of data, the calibration of simulation models that deal with car-following, lane changing behaviors, and other microscopic behaviors should be more emphasized.

DATA PROCESSING AND ENHANCEMENT

The data collected to serve as an input for the simulation tool should be filtered, aggregated, and corrected before the calibration and validation processes. As a first step, data should be processed, and data fusion techniques should be used (when dealing with multiple data sources) to generate an estimate of the traffic state of the system. Typical techniques in this stage are digital filters, least square estimation, expectation maximization methods, Kalman filters, and particle filters. In the second step, features and patterns that would be used to compare the results of the simulation model for calibration purposes should be derived from the preprocessed data. Classification and inference methods are the common techniques utilized in this step. Examples of these techniques are statistical pattern recognition, Dempster-Shafer, Bayesian methods, neural networks, correlation measures, and fuzzy set theory. Last but not least, the key input of the simulation tool should be specified (Daamen, Buisson and Hoogendoorn 2014).

Two types of error are usually seen in raw trajectory data: random error and systematic error. The noise in trajectory data is magnified when the position information is converted to speed and acceleration through a differentiation process. This is an example of random errors that can happen in trajectory data. On the other hand, the cumulative distance traveled by a vehicle is usually overestimated. The bias is a non-negative, non-decreasing function of the number of observations. The magnitude of this error increases as the speed of the vehicle decreases. This is an example of a systematic error in the trajectories. The space error in trajectory also becomes important when the interaction of vehicles is considered such as in car-following or lane changing models. Platoon consistency is the term defined in this area that verifies the consistency of the following vehicles and the traveled space. This error could be resolved by projecting coordinates on the road lane alignment (Punzo, Borzacchiello, and Ciuffo 2011).

As mentioned earlier, the calibration of the collected data is an import step toward reaching reliable results from the simulation tool. The way of performing this calibration also influences the accuracy of the resulting preprocessed data. For example, the vehicle trajectory positions (1) could be smoothed first then differentiated to get velocities and accelerations; (2) could be differentiated first and then all three types of data could be smoothed simultaneously; (3) a sequential process could be undertaken that smoothens the positions then derives velocities, then smoothens velocities and derives the accelerations afterwards. The preservation of internal consistency should be considered in any of these calibration processes. To be more specific, the integration of speed and acceleration should return the measured travel space (Daamen, Buisson and Hoogendoorn 2014).

One of the main limitations in assessing the accuracy of calibrating trajectory techniques (averaging, smoothing, Kalman filtering, Kalman smoothing) is that there is no established method that can quantify the level of accuracy. This is an important issue even with a large dataset such as the NGSIM data (Daamen, Buisson and Hoogendoorn 2014).

Macroscopic traffic data sensors such as loop detectors, radar, or infrared sensors, and cameras are prone to both random errors (noise) and structural errors (bias). Since data assimilation techniques are ineffective in resolving structural errors, other techniques such as speed bias correction algorithms should be used to address this issue (Daamen, Buisson and Hoogendoorn 2014).

For instance, one of the most common errors is to use time mean speed instead of space mean speed when deriving other traffic state parameters. The relationship between these two parameters could be written as:

$$v_t = v_s + \frac{\sigma_s^2}{v_s}$$

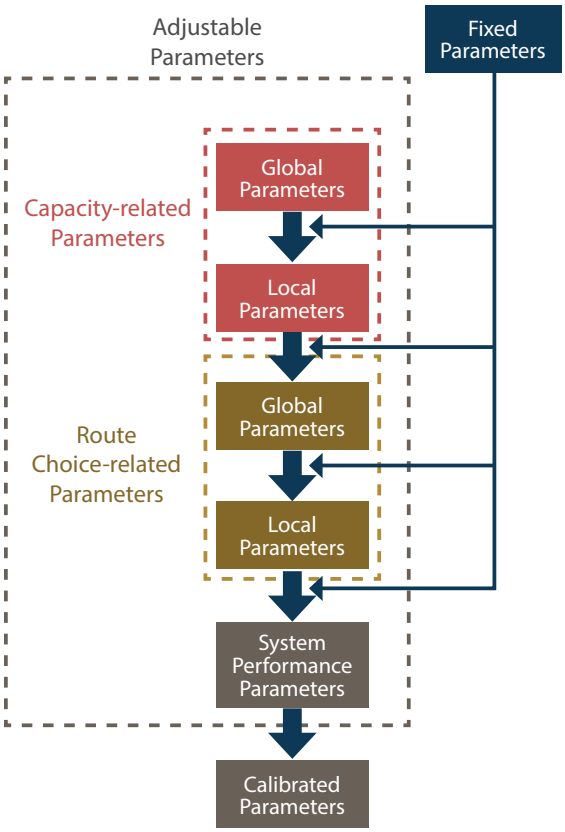
Figure 6. Formula. Relationship between time mean speed and space mean speed.

Where v_t is the time mean speed, v_s is the space mean speed, and σ_s is the variance of space speed. The variance is an unknown quantity in this formula, and the contribution of the second term could reach up to 25 percent in congested conditions. As a result, directly using the time mean speed to derive other traffic flow characteristics, such as density and flow, results in systematic biases (Daamen, Buisson and Hoogendoorn 2014).

ROLE OF VEHICLE TRAJECTORIES IN THE TAT CALIBRATION PROCEDURE

The traffic analysis toolbox (Dowling, Skabardonis and Alexiadis 2004) suggests a step-by-step calibration procedure. Figure 7 shows the procedure schematically. As could be seen in the figure, the parameters that serve as inputs for the simulation tool are divided into two groups of adjustable and fixed parameters based on a trade-off between the computational effort required for the calibration and the degree of freedom for fitting the calibrated model to the local conditions. The adjustable parameters are further subdivided into parameters that influence the system capacity and parameters that influence the route choice. These two groups of parameters could be further subdivided into global parameters and local parameters. As a sequential process, first, the global capacity-related parameters are calibrated by analyzing the queue discharging points at the network bottlenecks. Average following headway, driver reaction time, critical lane changing gap, minimum separation in the stop-and-go situation, startup lost time, queue discharge headway, and gap acceptance for unprotected left turns are examples of parameters that directly affect the capacity. Then, the model is fine-tuned by adjusting local parameters such as the presence of on-street parking, driveways, or narrow shoulders. The next step involves adjusting the global and local parameters that directly influence the route choice until the link volumes are best fitted to the collected data. Examples of global route choice-related parameters are route weightings for travel time and travel cost, drivers' familiarity with different each route, and the error in the drivers'

perception of the cost and time for each route. The last step in the calibration process is to modify some link free-flow speeds and link capacities. The modification should be limited in order to avoid significant deviations of the model behavior from the behavior derived from the capacity-related and route choice-related calibration steps. The aim of the system performance calibration is to achieve certain goals when model performance estimates are compared to field measurements. Examples of these calibration targets are the statistical thresholds for traffic flows, travel times, and the visual audits for link speeds and bottlenecks which are used by Wisconsin Department of Transportation (DOT) for its Milwaukee freeway system simulation model.



Source: FHWA, 2019.

Figure 7. Flowchart. The traffic analysis tools calibration procedure.

Sufficient information about vehicle trajectories could be very helpful in this sequential calibration process. The vehicle trajectories could not only specify the bottlenecks of the system but also could evaluate the capacity of the system at the queue discharging points. Both vehicle-based and video-based trajectory data provide useful information for the capacity-related model calibration stage. For the route choice-related parameters, the behavior of the drivers could be studied through the data collected by a vehicle-based trajectory data collection system.

LOCAL DENSITY THROUGH VEHICLE TRAJECTORIES

The vehicle trajectories could be processed to derive a variety of system performance metrics. An important system performance measure which is used in many traffic information systems is the traffic density. The commonly used vehicle density estimation that is designed based on infrastructure-based traffic information systems suffers from low reliability, limited spatial coverage, high deployment cost, and high maintenance cost. Another approach for estimating the road density is to use information derived from vehicle trajectories. This approach becomes more reliable as more vehicle become equipped with sensors and wireless technologies, and the connectivity level increases in vehicular ad-hoc networks (VANET).

The following formula shows the relationship between density and space mean speed in Pipes proposed a car-following model.

$$u = \lambda\left(\frac{1}{k} - \frac{1}{k_{jam}}\right)$$

Figure 8. Formula. Relationship between density and space mean speed.

Where λ is the vehicle interaction sensitivity, and k_{jam} is the jam density.

In the two-fluid theory, space mean speed could be calculated by the following formula:

$$u = u_{max}(1 - f_s)^{\eta+1}$$

Figure 9. Formula. Space mean speed in the two-fluid theory.

Where u_{max} is the allowed maximum speed, fs is the fraction of stopped vehicles, and ηs an indicator for service quality of vehicular networks.

Combining the above equations would result in a mathematical relationship between road density and the fraction of stopped vehicles as follows:

$$\frac{k}{k_{jam}} = \left[\frac{u_{max}k_{jam}(1 - f_s)^{\eta+1}}{\lambda} + 1\right]^{-1}$$

Figure 10. Formula. Relationship between road density and fraction of stopped vehicles.

Artimy (2007) used the ratio of stopping time to trip time for probe vehicles as an estimator for the fraction of stopped vehicles. In another study, Luo et al. (2017) proposed a velocity and acceleration aware density estimation algorithm to estimate the neighboring density of a vehicle with known trajectory using the vehicle acceleration and velocity:

$$D_s = \begin{cases} \dfrac{|a|}{\alpha u_i + (1-\alpha)u_{i-1}}, & 1 \le i \le n, u_i \ne 0 \\ 1, & u_i = 0 \end{cases}$$

Figure 11. Formula. Density as a function of velocity and acceleration.

Where *i* represents a time step, and *a* is equal to one when the vehicle is accelerating and zero when it is decelerating.

In order to convert the real-time local density to the road segment density, averaging techniques could be utilized, or the segment length should be short enough to avoid significant variations in the estimated local density values. One of the main limitations stated by both studies is the inaccuracy of the local density estimate in free-flow traffic conditions.

Simpler techniques also exist to estimate the density of a road segment using the local density around vehicles that can collect information of their surrounding area. One approach is to use the inverse of the average inter-vehicular spacing. The other approach is to find the density of the vehicles in the area that could be visualized by the vehicle sensors. The analysis would be performed on a case study to infer the road segment density from local density values.

DATA REQUIREMENTS

One of the main data requirements, which serves both on the input side as a basis for calibration of the simulation model parameters, and on the output side to perform validation, is the vehicle trajectory data. Mahmassani et al. (2014) enumerate trajectory information requirements as follow:

- Provide at least longitude and latitude (or X-Y) coordinates with associated timestamps.
- Capture a wide range of operational statuses such as recurring and nonrecurring congestion.
- Cover a wide range of road facilities from freeways to arterial roads and intersections.
- Include sufficient sampling and time-series to support statistically meaningful analysis.
- Provide the capability of fusing other ancillary data with the trajectory data.

Satisfactory knowledge over the trajectories, the independent behavior of simulation agents, and the interactions between agents in the simulation environment could support the development of libraries of agents with attributes associated with them representing the agent behavior in the simulation environment. The heterogeneity among members of an agent class could be defined via a probabilistic distribution as an attribute for the agent class or the agent could be further divided into subagent classes.

Beside the endogenous effects that could be derived from the underlying interaction between vehicle trajectories, another important source of information used in the simulation is related to exogenous sources of variability, such as policy interventions, traffic incidents, construction zones, special events, and weather conditions. These datasets are usually provided by transportation authorities that are in charge of the management, control, or monitoring of transportation systems or by third parties such as meteorological organizations. Table 2 shows the data used to characterize exogenous sources of variability. Libraries of scenarios corresponding to different operational conditions could be created based on relevant historical data.

Table 2. Typical data elements for exogenous sources of variation in system performance measures.

Event Type	Data Elements
Incident	• Type (e.g., collision, disabled vehicle). • Location. • Date, time of occurrence, and time of clearance. • Number of lanes/shoulder and length of roadway affected. • Severity in case of collision (e.g., damage only, injuries, fatalities). • Weather conditions. • Traffic data in the area of impact before and during the incident (e.g., traffic flows; speed, delay, travel time measurements; queues; and other performance measures or observations, if available).
Work zone	• Work zone activity (e.g., maintenance, construction) that caused lane/road closure, and any other indication of work zone intensity. • Location and area/length of roadway impact (e.g., milepost), number of lanes closed. • Date, time, and duration. • Lane closure changes and/or other restrictions during the work zone activity. • Weather conditions. • Special traffic control/management measures, including locations of advanced warning, speed reductions. • Traffic data upstream and through the area of impact, before and during the work zone (e.g., traffic flows and percentage of heavy vehicles; speed, delay, travel time measurements; queues; and other performance measures or observations, if available). • Incidents in work zone area of impact.
Special event	• Type (e.g., major sporting event, official visit/event, parade) and name or description. • Location and area of impact (if known/available). • Date, time, and duration. • Event attendance and demand generation/attraction characteristics (e.g., estimates of out-of-town crowds, special additional demand). • Approach route(s) and travel mode(s) if known. • Road network closures or restrictions (e.g., lane or complete road closures, special vehicle restrictions) and other travel mode changes (e.g., increased bus transit service). • Special traffic control/management measures (e.g., revised signal timing plans). • Traffic data in the area of impact before, during, and after the event (e.g., traffic flows; speed, delay, travel time measurements; queues; and other performance measures or observations, if available).
Weather	• Weather station identification or name (e.g., KLGA[1] for the automated surface observing system station at LaGuardia Airport, New York). • Station description (if available). • Latitude and longitude of the station. • Date, time of weather record (desirable data collection interval: 5 minutes). • Visibility (miles). • Precipitation type (e.g., rain, snow). • Precipitation intensity (inches per hour, liquid equivalent rate for snow). • Other weather parameters (temperature, humidity, precipitation amount during previous 1 hour, if available).

Source: Mahmassani, Kim, et al., 2014.

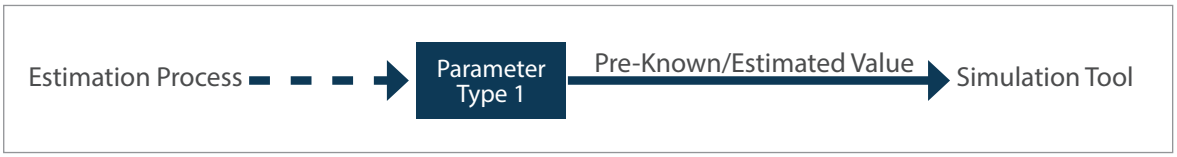
[1] KLGA is the four-letter International Civil Aviation Organization designation for at LaGuardia International Airport and its associated weather station.

DATA STRUCTURE

As mentioned in the review and assessment of calibration methodologies, the correlation between input parameters is an important factor that should be considered. The correlation among the parameters could be captured through a joint probabilistic distribution assigned to the parameters. In general, the type of parameters that would be used in this project could be classified into three groups as follows:

- Type 1: Parameters with the least level of uncertainty.
- Type 2: Parameters with some level of uncertainty.
- Type 3: Parameters with deep uncertainty.

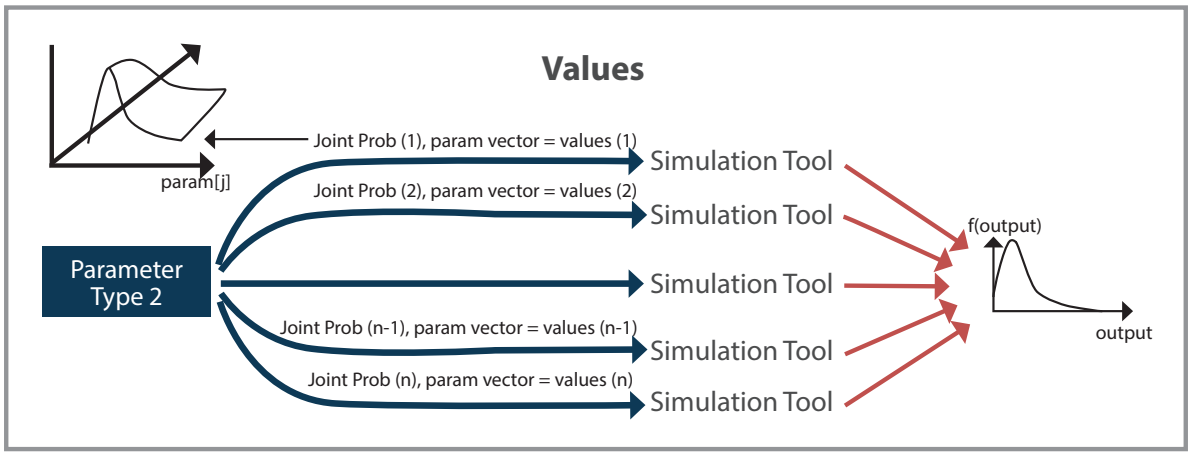
The first type of parameters pertains to most of the parameters available in the base year that capture the behavior of different agents in the normal prevailing condition of the system in the absence of new technologies. The heterogeneity among the individuals in the same agent class is preserved through an assigned probabilistic distribution or a sub-classification of the agent. The least uncertainty level exists in the parameter values and the interaction effects. Historical data and predefined libraries of agents could be used to determine or estimate the values of these parameters (figure 12). The data still should be preprocessed to remove random errors and systematic biases.



Source: FHWA, 2019.

Figure 12. Illustration. Parameters with the least level of uncertainty [type 1].

The second type of parameters imposes more uncertainty on the analysis. The uncertainty could be related to the value of the parameters such as the unknown composition of risk-averse and risk-taking individuals among the drivers when the cognitive aspects are incorporated in the car-following models (Hamdar, et al. 2008). The uncertainty could also be attributed to the correlation between parameters. For example, there are uncertainties involved when the effect of exogenous sources that cause variations in the system performance measures. Examples are mentioned in table 1. In this case, a library of scenarios is available and through sampling techniques, such a Monte Carlo sampling, the values, and associated probabilistic distributions would be selected. The sampled sets would be entered in the simulation tool and the outputs would be processed and combined to develop probabilistic distributions of system performance measures (figure 13).



Source: FHWA, 2019.

Figure 13. Illustration. Parameters with some level of uncertainty [type 2].

The third type of parameters encompasses the highest level of uncertainty. This case is relevant to emerging technologies in the transportation system such as autonomous vehicles and innovative control strategies in connected environments. In this case, the parameter values, the correlation effects, and the joint probabilistic distributions are unknown. A process similar to the scenario-based value development for parameters type 2 should be performed for parameters type 3. However, due to the high level of uncertainty in the values and joint distributions, a robust sensitivity approach is utilized. Several criteria from decision theory have been used in the literature to perform robust sensitivity analysis when a large number of scenarios should be analyzed, namely, the maximax, the weighted average, the minimax regret, and the limited degree of confidence (Gao and Bryan 2016, Gao, Bryan and Nolan, et al. 2016).

The maximax decision criterion (S^{max}) first finds the maximum possible payoff (gain or loss) for each scenario. For example, the maximum travel time for a specific path or the least capacity for a specific link. Then, it selects the scenario that has the highest maximum payoff as the selected scenario. The weighted average decision criterion (S^{avg}) is utilized when multiple performance measures are considered simultaneously, and the sensitivity indicators are combined with appropriate weights to calculate an overall sensitivity indicator that could be used to specify the selected scenarios. The minimax regret decision criterion (S^{reg}) is a commonly used criterion in decision theory. Regret is defined as follows:

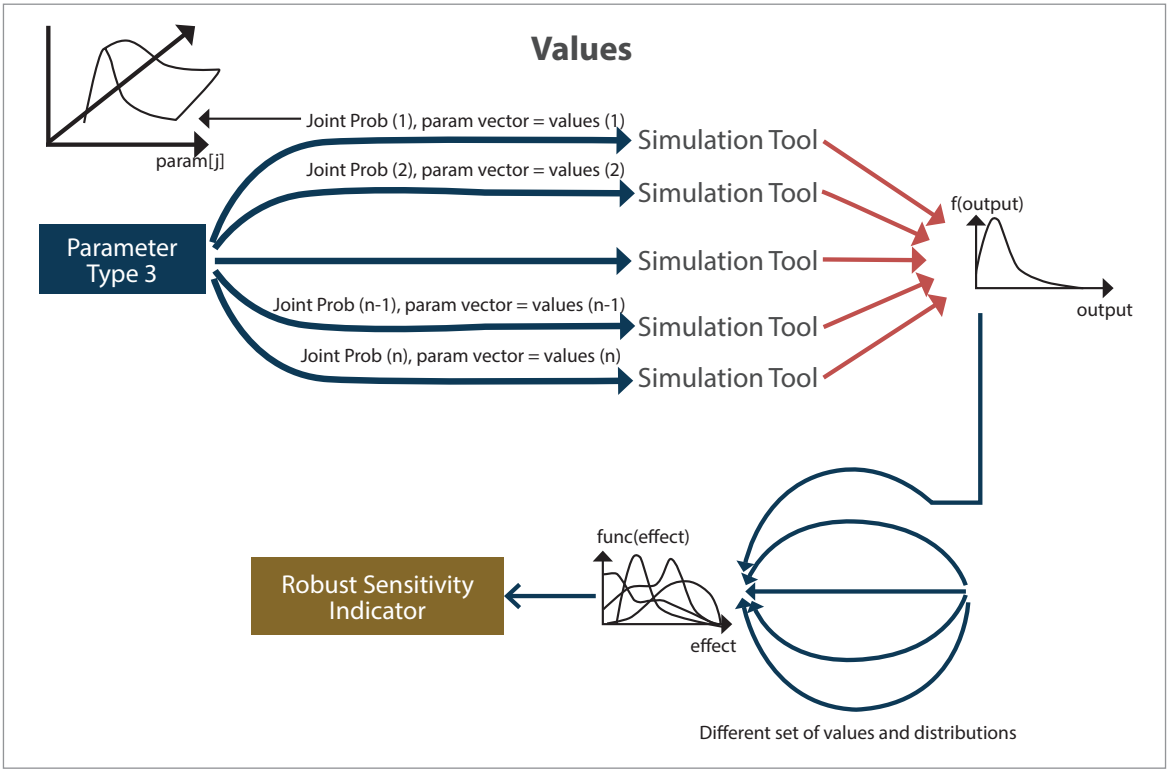
$$\min_a \max_b(r_{a,b}) = \min_a \max_b(|S_a - S_b|)$$

Figure 14. Formula. Regret formulation.

Where r_{ab} is the regret when the scenario "a" occurs for type 3 parameters and "b" pertains the type 1 parameter values and the type 2 parameter scenario-based values. S_b is the system performance measure when there are no new technologies in the system. This is developed by using preset values for type 1 parameters and sampled scenarios from the scenario library for type 2 parameters. S_a is the system performance measure when parameters related to the new technology (parameter type 3). The minimax regret, first, derives the highest system performance gap that is

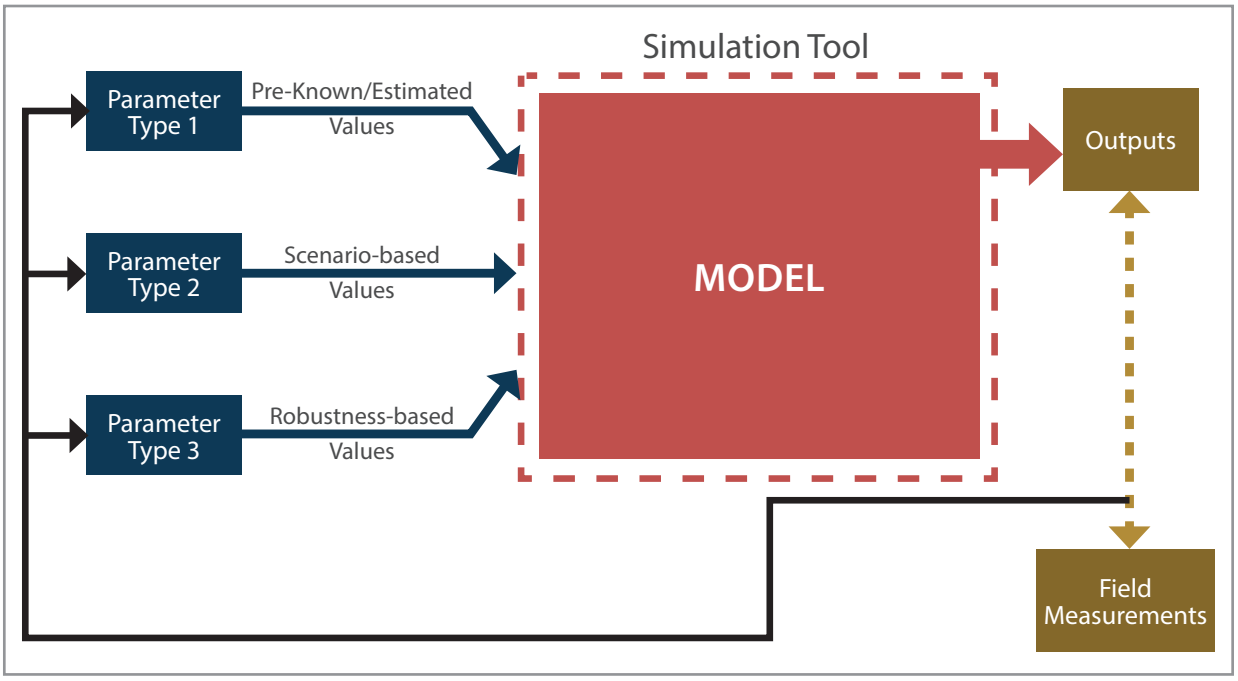
achieved between the system influenced by selected values and probabilistic distributions of type 3 parameters and the system influenced by any feasible combination of type 1 and type 2 parameters. Then, among all the feasible scenarios of type 3 parameters, the one which possesses the lowest gap is selected as the type 3 parameter set that complies with the minimax regret criterion.

The limited degree of confidence decision criterion (S^{ldc}) is a criterion where the indicator value is a weighted average of the weighted average indicator value and the minimax regret indicator value. This criterion implies the decision-maker's degree-of-confidence in the probability distributions of the assumed scenarios. The degree-of-confidence is quantified using an appropriate weight. Figure 15 shows the robustness-based value generation procedure which is similar to the scenario-based value generation except that the robustness-based value generation includes a larger number of scenarios. Furthermore, evaluation of the robust sensitivity indicator is the key component that determines the selected scenario (values and probabilistic distributions for type 3 parameters). The overall framework, shown in figure 16, represents the simultaneity of entering different types of data into the traffic simulation tool.



Source: FHWA, 2019.

Figure 15. Illustration. Parameters with the deep uncertainty [type 3].



Source: FHWA, 2019.

Figure 16. Flowchart. Overall framework.

CONCLUSIONS

Vehicle trajectories provide the most complete record of a vehicle's behavior over space and time, and afford the analyst the most flexibility and richness in characterizing virtually all aspects of traveler behavior and traffic system performance. Trajectories provide a unifying framework for conducting model calibration and subsequent validation of large-scale traffic operational models.

As noted, vehicle trajectories are still not regularly available to traffic analysts, though the situation is rapidly changing with greater willingness by system integrators and data vendors to share this information (albeit with all kinds of limitations on use). However, greater deployment of connected vehicle systems promises to dramatically increase the availability and accuracy of this type of data. From the standpoint of this project, in addition to possible sources of actual trajectory data to support the calibration framework development and application, a useful strategy would be to use simulated trajectories, i.e., the output of simulation models, as a way to develop and test the main calibration concepts and frameworks under development. An important advantage lies in the ability to then obtain measurements on the complete population of drivers in the network, allowing us the ability to test different partial observation scenarios (sampling schemes) on the accuracy of the results.

CHAPTER 4. CALIBRATION NEEDS

Given the assessments of calibration methodologies and data presented in the previous chapters, this chapter presents a gap analysis. The analysis will identify a set of traffic analysis tool calibration needs for transportation improvement evaluations, and describe the impacts of these gaps. The set of calibration needs is intended to provide a "snapshot" of the potential updates that may be beneficial for the current calibration methods.

Note: Unless accompanied by a citation to statute or regulations, the practices, methodologies, and specifications discussed below are not required under Federal law or regulations.

GAP ANALYSIS AND GAP IMPACTS

Current calibration practices mostly use statistical, econometric, or optimization techniques to find parameter values that make model outputs best match the available data. Calibration, inherently, requires knowledge and observation/measurement of actual conditions. When comparing the model outputs to the data representing the actual conditions of the system, it is typically assumed that the relations established at the calibration stage would remain valid in the future under new system conditions. Therefore, the model specifications and parameters are assumed to remain unchanged. The only changes considered for future conditions are in the values of the variables that serve as input for the simulation model. The values of variables are typically forecast independently or determined as the result of an intervention. These assumptions limit the effectiveness of the simulation tool in capturing the system behavior under the future conditions and desired interventions.

Whenever there are uncertainties or heterogeneity associated with model parameters considered in the calibration process, the typical process is to draw samples from the set of parameter values using the marginal probability distributions of the parameters. As the sampled parameters are inserted into the simulation tool, unrealistic behavior could appear in the system output. One of the reasons for the unrealistic behavior is the correlation existing among the parameters which could not be captured when the samples are taken from the marginal probability distributions.

IMPACTS OF THE GAPS

The limiting assumption regarding the consistency of the parameter values in the calibration stage and those under the future conditions of the system has two major effects. First, the calibrated model is unable to capture the system performance under future conditions and policy interventions. The other impact is that a more reliable understanding of the transportation system could be achieved only if a full re-calibration process is performed. Therefore, the computationally expensive calibration should be repeated more frequently as the previously calibrated model become ineffective when applied to a new area and context.

The main influence of ignoring the correlations among parameters incorporated in the simulation model is that unrealistic behavior could evolve in the simulated environment; hence the reliability of the model would be decreased.

Based on the impact of the gaps in the current calibration practices, the main challenge is how to ensure that the model specifications and parameter values used in applying the model to predict traffic system performance under the future scenario of interest are adequate; i.e., that the model is "calibrated for future conditions." Of course, since these future conditions have not yet occurred at the time the model is used, it is not possible to directly observe actual conditions for such scenarios as a basis for model calibration. Addressing this challenge involves a significant shift in the mindset of traffic modelers: from model estimation and calibration aimed at replicating existing and past conditions to greater emphasis on prediction quality and accuracy. A new calibration framework should be developed that intrinsically encompasses this mindset shift. The necessary elements of the proposed framework are discussed in the following section.

CALIBRATION NEEDS

The main problem addressed in this study is how to develop and calibrate models so that the output they produce is meaningful and provides a realistic and accurate depiction of future conditions. In order to address this problem, there are several elements that should be incorporated in the calibration tools.

First, libraries of models and parameters associated with each model should be added to the calibration framework. These libraries contain different values for the parameters as well as an estimated probability distribution for the set of values. The library items are constructed based on historical data and forecast values for future conditions not been realized yet. Table 3 provides an example of three well-known car-following models with their parameters. The traffic analysis tool should have the capability of updating the library of models and parameters as new possible sets of values for parameters and new behavioral models are introduced.

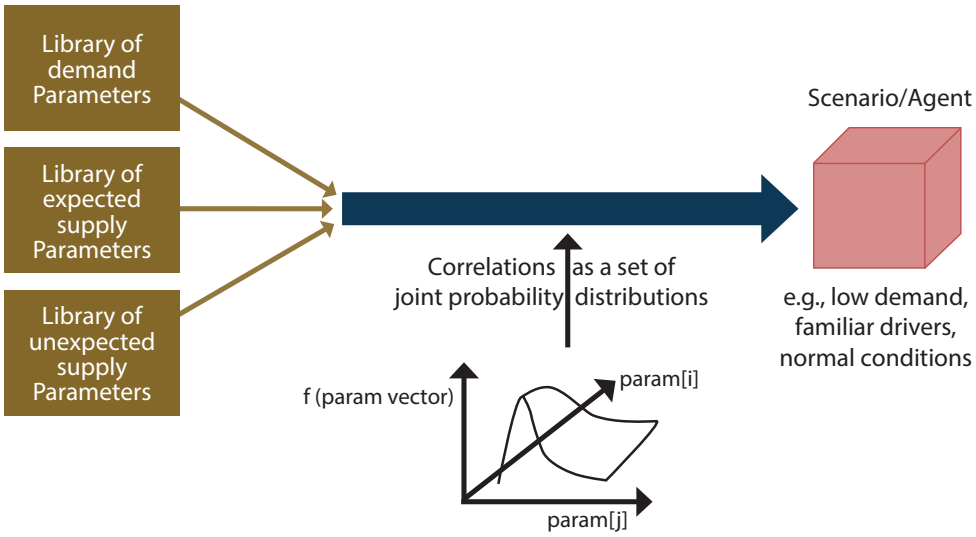
Table 3. Parameters of three car-following models.

Model	Parameter	Description
Gipps	T	Reaction time.
	a	Maximum acceleration.
	d	Maximum desirable deceleration (<0).
	V^*	Desired speed.
	s^*	Minimum net stopped distance from the leader.
	α	Sensitivity factor.
Helly	T	Reaction time.
	C_1	Constant for the relative speed.
	C_2	Constant for the spacing.
	d^*	Desired net stopped distance from the leader.
	γ	Constant for the speed in the desired following distance.
IDM	a	Maximum acceleration.
	b	Desired deceleration.
	V_0	Desired speed.
	s_0	Minimum net stopped distance from the leader.
	T_{HW}	Desired safety time headway.

IDM = intelligent driver model.

Second, the correlations among elements of the libraries should be integrated into the calibration tool. The correlations could be represented as joint probability distributions for vectors of parameters. Throughout the calibration process, selecting the model parameters from the joint probabilities instead of the marginal probabilities of each individual parameter could reflect more realistic behavior of the simulation agents.

As shown in figure 17, combining the libraries of parameters related to different components of the transportation system (demand-related, expected supply-related, and unexpected supply-related components) and selecting from the library values based on the joint probabilities of the selected parameters defines a single scenario/agent block. A scenario is defined by set of operational conditions (reflecting external events, such as weather, demand surges, etc.), interventions (infrastructure changes, control actions, dynamic system management schemes, etc.), as well as characteristics of the general activity system (land use, activity locations) and associated technologies (e.g. connected and autonomous vehicles, Internet of Things, smart cities). On the other hand, an agent is an object in the simulation environment that follow a set of rules defined by various behavioral models. The behavioral characteristics of an agent are defined by the parameters of the agent's behavioral models.

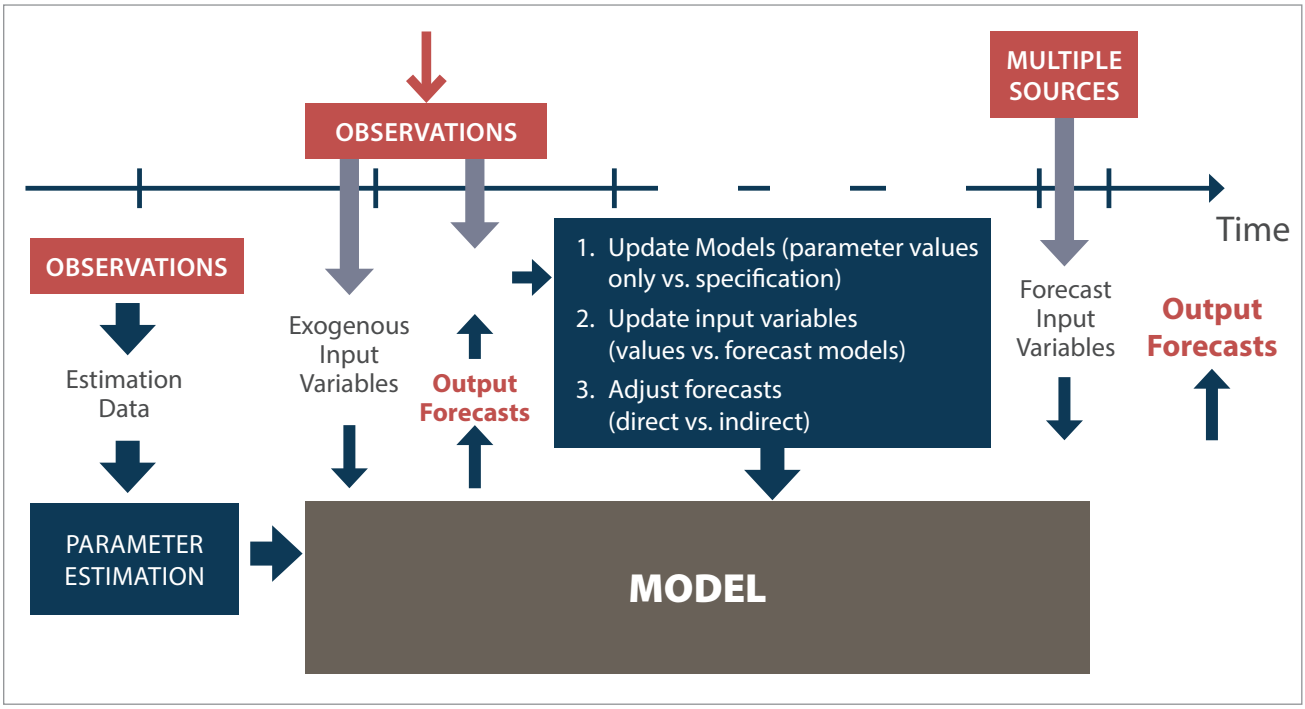


Source: FHWA, 2019

Figure 17. Illustration. Process of developing a scenario/agent.

The third useful element for the calibration of traffic analysis tools is trajectory-based data. As mentioned earlier, one of the main aspects of the new calibration procedure is a significant shift from model estimation and calibration based on existing and past conditions, to greater emphasis on prediction quality and accuracy. To perform this fundamental change, a higher quality and more accurate input data set should also be used. The trajectories of agents (vehicles, pedestrians, bicyclists, etc.) could provide such a level of accuracy through continuous monitoring of the agent behavior over time and space. The simulation tools should become capable to receive trajectories of agents as an input data. As small errors in processing trajectory-based data could result in significant errors in the model outputs, appropriate preprocessing techniques and error handling methods should be included in the algorithms of the simulation tool.

Figure 18 shows the proposed framework for the calibration process. As the first step, after refining the input data and performing an estimation process on the observed data, the appropriate set of parameters are selected/estimated. Simultaneously, the external conditions of the system are incorporated as exogenous input variables. In the proposed framework, the estimated parameters and external conditions are fed to the simulation model in terms of libraries of models, parameters, and related joint probabilities. These library components and probabilities would be used to develop various scenarios and agents as inputs for the simulation model. Using a set of scenarios and agents, the model will be calibrated based on the observations and forecast input variables collected from other sources such as census data.



Source: FHWA, 2019.

Figure 18. Flowchart. Proposed calibration framework.

In general, the longer the horizon over which a scenario is defined, i.e., the further out into the future one is trying to forecast, the greater the uncertainty associated with the values of input variables. This is especially true not only for the activity system variables and the associated technologies but also the social fabric and associated preferences/lifestyles/norms. Furthermore, as the horizon becomes longer, there is a need for a more diversified set of scenarios and agents in the calibration process.

A major difference between the conventional calibration process and the proposed calibration process is the manner that the model is updated over time. The conventional approach involves a full re-calibration process when the model is applied to a new context. On the other hand, in the proposed framework, this computationally intensive process is avoided to some extent by updating the scenarios and libraries. Adaptive updating processes such as Bayesian techniques could be used to update the library parameters and their associated joint probability distributions.

As illustrated in figure 19, probability distributions of the model outputs are developed based on the probability distributions assigned to the scenarios. The outputs of the scenarios are combined using the probability assigned to each scenario in order to generate the final probability distribution of the output. If a comprehensive list of scenarios and agents are considered in the calibration process, as new observations are collected over time, the probabilities assigned to the agents and scenarios could be revised. Consequently, the probability distributions of the outputs could be adjusted accordingly. Throughout this updating process, there would be no need to perform a full re-calibration process and the updates would be directly applied to the inputs and outputs. In the proposed framework, a re-calibration process could become necessary if the conditions of the new observations are not covered in the set of scenarios/agents or if a new behavioral model is introduced to the transportation system.

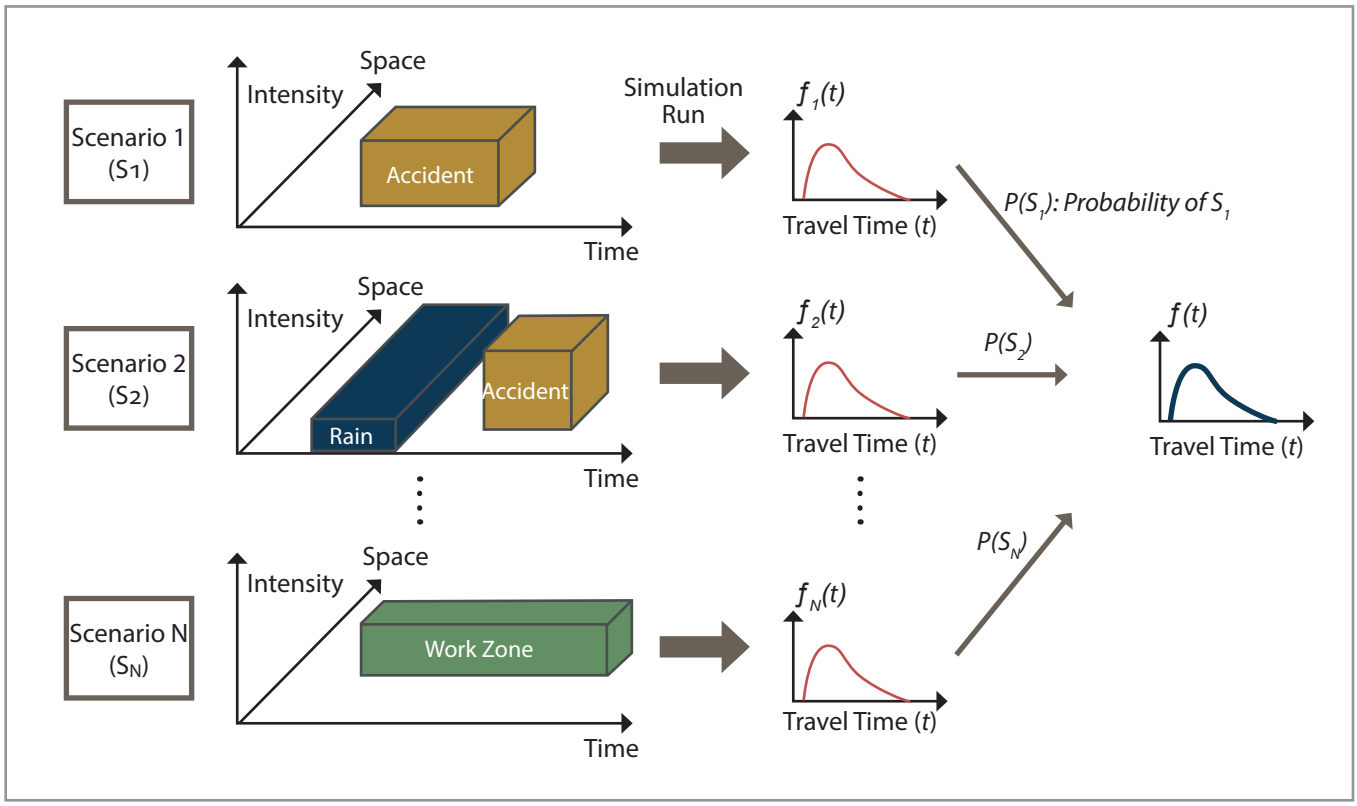


Figure 19. Illustration. Constructing model output (travel time) distribution based on scenario-specific simulation outputs.

CALIBRATION NEEDS LIBRARIES

The first element introduced in the above framework is the libraries of models and parameters. Libraries could be classified into libraries of microscopic, mesoscopic, macroscopic, strategic models. The libraries listed in this section constitutes well-known models which are thoroughly studied in the literature. The final traffic analysis calibration tool should be able to accommodate new models introduced as new technologies are introduced (new modes, new fuels or propulsion technologies, connected and/or automated vehicles), or major changes are implemented in the operation of the transportation system. As the macroscopic and mesoscopic relations reflect the underlying microscopic behaviors of the drivers in a particular setting, it is important to consider the correlations that exist between the macroscopic, mesoscopic, and microscopic parameters. Therefore, besides the correlations that exist among the parameters of a specific library, there are correlations across libraries that are defined for various analysis levels (micro, meso, and macro).

CALIBRATION NEEDS FOR OPERATIONAL MODELS: MICROSCOPIC

Microscopic models of driver behavior in traffic capture the maneuvers of individual drivers as they interact in traffic, including car following, acceleration choice, lane changing, merge decisions, gap acceptance, and queue discharge headways. The model parameters in this case typically capture psychological aspects of driver behavior, including risk attitudes. When using a calibrated model for prediction under new contemplated scenarios, the main factors assumed to change are the situational variables describing respective vehicle positions and speeds. However, parameters governing these behaviors are known to change across drivers and vehicle characteristics (heterogeneity), which in turn vary across locations and over time. Similarly, these are affected by operational conditions such as rain, snow, and visibility.

To deal with this variability, libraries of parameter values may be developed reflecting behaviors under different conditions. Furthermore, it is expected that new vehicle technologies, especially connected and/or automated vehicles will have a major impact on microscopic behaviors. Therefore, libraries related to the microscopic behavior of vehicles should be enhanced as new models and new parameter values are introduced. Car-following models are among the most important models characterizing the microscopic behavior of vehicles. The Newell model is one of simplest car-following models in which the follower imitates the leader trajectory with a lag in time and space. Therefore, it is also known as a trajectory translation model (Punzo and Simonelli 2005).

$$x_n(t+T) = x_{n-1}(t) - d_n$$

Figure 20. Formula. Newell car-following model (trajectory translation model).

The Gipps model is a safety distance model that considers the safe speed in two different regimes: the free-flow regime and the car-following regime (Kim and Mahmassani 2011).

$$v_{a,n}(t+T) = v_n(t) + 2.5aT\left(1 - \frac{v_n(t)}{V^*}\right)\sqrt{0.025 + \frac{v_n(t)}{V^*}}$$

$$v_{b,n}(t+T) = dT + \sqrt{d^2T^2 - d\left[2\left\{x_{n-1}(t) - (l_{n-1} + s^*) - x_n(t) - v_n(t)T - \frac{\left(v_{n-1}(t)\right)^2}{\alpha d}\right\}\right]}$$

$$v_n(t+T) = min\{v_{a,n}(t+T), v_{b,n}(t+T)\}$$

Figure 21. Formula. Gipps car-following model.

The Helly model is a linear model that relates the acceleration of the follower to the speed difference and spacing between the leader and follower considering the reaction time as the time step (Kim and Mahmassani 2011).

$$a_n(t) = C_1[v_{n-1}(t-T) - v_n(t-T)] + C_2[x_{n-1}(t-T) - x_n(t-T) - D_n(t)]$$

$$D_n(t) = (d^* + l_{n-1}) + \gamma v_n(t-T)$$

Figure 22. Formula. Helly car-following model.

The intelligent driver model (IDM) is a continuous response model which considers the follower's acceleration to be a continuous function of the leader's speed, and spacing and speed differences between the two vehicles. A major difference between this model and the other models introduced in this section is that the IDM does not consider any reaction time in the formulation (Kim and Mahmassani 2011).

$$a_n(t) = a\left[1 - \left(\frac{v_n(t)}{V_0}\right)^4 - \left[\frac{s_0 + \left(v_n(t)T_{HW} + \frac{v_n(t)[v_n(t) - v_{n-1}(t)]}{2\sqrt{ab}}\right)}{x_{n-1}(t) - l_{n-1}(t) - x_n(t)}\right]^2\right]$$

Figure 23. Formula. The Intelligent Driver Model.

Last but not least, a stimulus-response acceleration model implemented in the MITSIMLab (MIT Intelligent Transportation Systems sIMulation Laboratory) microsimulator. It combines sensitivity and the effect of a stimulus to produce the relevant response (Punzo and Simonelli 2005).

$$a_n(t) = \alpha^i \frac{\left(v_n(t-T)\right)^{\beta^i}}{\left(x_{n-1}(t-T) - x_n(t-T)\right)^{\gamma^i}} |v_{n-1}(t-T) - v_n(t-T)|^{\lambda^i}$$

Figure 24. Formula. The stimulus-response acceleration model in MITSIMLab.

Where x_n *(t+T)* and $x_{n-1}(t)$ represent the follower and leader positions, respectively; v_n *(t+T)* and $v_{n-1}(t)$ represent the follower and leader speeds, respectively; Similarly, a_n *(t+T)* and $a_{n-1}(t)$ are the follower and leader accelerations, respectively; T is the reaction time; d_n is the desired intervehicle spacing; a is the maximum acceleration; d is the maximum desirable deceleration; V^* represents the desired speed; l_n is the physical length of vehicle; s^* is the minimum net stopped distance from the leader; α is the sensitivity factor; d^* is the desired net stopped distance from the leader; C_1, C_2,and γ are constant parameters related to the relative speed, spacing, and speed in the desired following distance, respectively; b is the desired deceleration; V_o is the desired speed; s_o is the minimum net stopped distance from the leader; T_{HW} symbolizes the desired safety time headway; and α^i, β^i, γ^i, and λ^i are eight parameters related to the car-following regime (with i specifying whether the follower is in acceleration or deceleration mode).

Another important set of microscopic models in the traffic flow is the lane-changing and merging decision models. Rahman et al. (2013) classified the lane-changing models into four groups: rule-based models, artificial intelligence models, incentive-based models, and discrete-choice-based models. The Gipps model, the model implemented in CORSIM (CORridor SIMulation), the ARTEMiS (Analysis of Road Traffic and Evaluation by Micro-Simulation) model, the cellular automata model, and the game theory-based models are examples of the rule-based models. Examples of artificial intelligence models are the models developed based on the fuzzy logic or based on an artificial neural network structure. MOBIL (Minimizing Overall Braking Induced by Lane changes) , as an incentive-based model, is a lane-changing model that combines the attractiveness of a given lane defined by a utility function and the risk associated with the lane-changing maneuver, which is compared to a safety criterion. There are various parameters in the definition of the utility function describing the attractiveness. Furthermore, other parameters could be involved in the combination of the incentive and risk measures. LMRS (Lane-Changing Model with Relaxation and Synchronization) is another example in the incentive-based model category. This model performs a trade-off between the route, speed, and keep-right incentives. The following equation shows one way of combining these incentives to measure the driver's desire to change lane:

$$d^{ij} = d_r^{ij} + \theta_v^{ij}\left(d_s^{ij} + d_b^{ij}\right)$$

Figure 25. Formula. The Lane-Changing Model with Relaxation and Synchronization.

Where d^{ij} is the combined desire to change lane from i to j; d_r^{ij} is the desire to follow a route; d_s^{ij} is the desire to gain speed; $d_b^{ij}$is the desire to keep right; and θ_v^{ij} is voluntary (discretionary) incentive which is a parameter that needs to be determined and adjusted based on the available context.

The fourth category of lane-changing models is the discrete-choice-based models. Ahmed (1999) proposed a dynamic discrete choice model for the three different types of lane-changing maneuvers, namely, the mandatory lane changing (MLC), discretionary lane changing (DLC), and the forced merging (FM). He formulated the behavior using a probabilistic formulation as follows:

$$Pr_t(LC|v_n) = \frac{1}{1+\exp(-X_n^{LC}(t).\beta^{LC} - \alpha^{LC}.v_n)}$$
$$LC = MLC, DLC, FM$$

Figure 26. Formula. A probabilistic model for lane changing.

Where $Pr_t(LC|v_n)$ is the probability of executing MLC, DLC, or FM for driver n at time t; X_n^{LC} is a vector of explanatory variables affecting decision to lane changes; $\beta^{LC}$is the corresponding vector of parameters; v_n is the driver-specific random term; and α^{LC} is the parameter associated with the random term. Jia et al. (2011) estimated and validated the parameters of involved in this formulation using traffic data collected from a typical urban expressway merging section in Guangzhou. They used the lead gap, the lag gap, the relative speed of the lead vehicle to the subject vehicle, the relative speed of the subject vehicle to the lag vehicle, the subject speed, the remaining distance, and driver characteristics as the explanatory variables in the merging probability model.

Gap acceptance models, similar to the lane-changing behavior, possess a rule-based or a probabilistic framework. Mahmassani and Sheffi (1981) proposed a gap acceptance model using a standard cumulative normal distribution mathematical formula as follow:

$$Pr(accept\ gap\ i|\bar{T}_{cr}, \beta, \delta, i, t_i) = \phi\left(\frac{t_i - \bar{T}_{cr} - \beta(i-1)^{\delta}}{\sigma}\right)$$

Figure 27. Formula. A gap acceptance model based on standard cumulative normal distribution.

Where $\phi(.)$ denotes the standard cumulative normal curve; i is the sequential number of the gap; t_i is the gap duration; $\bar{T}_{cr}$ is the average critical gap when facing the first gap; β and δ are parameters that need to be estimated and subsequently calibrated based on the context that the model is being utilized; and σ is the variance across drivers and gaps.

Taylor and Mahmassani (1998) proposed a gap acceptance model with a structure similar to the above formulation. They used the model to characterize the gap acceptance behavior of bicyclists and motorists with a probit mathematical formulation but with different estimated parameters.

$$Pr\begin{pmatrix} vehicle/operator\ unit\ v \\ accepts\ the\ g^{th}\ gap \\ while\ attempting\ maneuver\ m \\ in\ an\ intersection\ environment\ i \end{pmatrix} = \phi\left(\frac{d_{gvmi} - (\beta_0 + \beta_{gap}X_{gap} + \beta_{int}X_{int} + \beta_{man}X_{man} + \beta_{vo}X_{vo} + \beta_{ia}X_{ia})}{\sigma}\right)$$

Figure 28. Formula. A probit gap acceptance model for bicyclists and motorists.

Where $\phi(.)$ denotes the standard cumulative normal curve; d_{gvmi} represents the actual gap duration; β_o is the mean critical gap when all other attributes are zero; X_{gap} is a vector of attributes characterizing the gap and surrounding events; X_{int} is a vector of attributes characterizing the intersection environment; X_{man} is a vector of attributes characterizing the acting vehicle maneuver; is a vector X_{vo} of attributes characterizing the vehicle/operator unit; X_{ia} is a vector of interactions between attributes from different groups; β_{gap}, β_{int}, β_{man}, β_{vo}, and β_{ia} is the total variance across vehicle/operator units, gaps, maneuvers, and intersection types.

Finally, the queue discharge models, that explain the queue discharge behavior under stochastic capacity scenarios at freeway bottlenecks, play a significant role in determining the performance of freeway systems. Jia et al. (2010) proposed a robust and easily implementable methodology that identifies freeway bottlenecks and models the queue discharge behavior at the bottlenecks. They formulated the queue discharge rate as a stochastic time-correlated recursive formula. It was implemented in DYNASMART-P using the following functional form:

$$C_t = C_{t-1} + \beta(\mu_c - C_{t-1}) + \varepsilon_t$$

Figure 29. Formula. Queue discharge rate in DYNASMART-P.

Where C_t is the queue discharge rate at time interval t; β is parameter that should be estimated and calibrated; μ_c is the average discharge rate; and ε_t is a normally distributed random term that is added to consider the stochasticity nature of the problem.

Table 4 summarizes the microscopic model parameters introduced in this section that should be calibrated. There are various other models that could be added to the listed microscopic models.

Table 4. Parameters of microscopic models.

Model	Parameter	Description
CAR-FOLLOWING MODELS		
Newell	T	Reaction time.
	d_n	Desired intervehicle spacing.
Gipps	T	Reaction time.
	a	Maximum acceleration.
	d	Maximum desirable deceleration (<0).
	V^*	Desired speed.
	s^*	Minimum net stopped distance from the leader.
	α	Sensitivity factor.
Helly	T	Reaction time.
	C_1	Constant for the relative speed.
	C_2	Constant for the spacing.
	d^*	Desired net stopped distance from the leader.
	γ	Constant for the speed in the desired following distance.

Table 4. Parameters of microscopic models. (continued)

Model	Parameter	Description
IDM	a	Maximum acceleration.
	b	Desired deceleration.
	V_0	Desired speed.
	s_0	Minimum net stopped distance from the leader.
	T_{HW}	Desired safety time headway.
MITSIM	α^i	Parameter for the car-following regime (i represents acceleration/deceleration).
	β^i	Parameter for the car-following regime (i represents acceleration/deceleration).
	γ^i	Parameter for the car-following regime (i represents acceleration/deceleration).
	λ	Parameter for the car-following regime (i represents acceleration/deceleration).
	i	
LANE-CHANGING MODELS		
LMRS Ahmed (1999)	θ_v^{ij}	Voluntary (discretionary) incentive.
	β^{LC}	Vector of parameters in the lane-changing model.
	α^{LC}	Parameter associated with the random term.
Jia et al. (2011)	D^{lead}	Lead gap.
	D^{lag}	Lag gap.
	ΔV^{lead}	Relative speed of the lead vehicle to the subject vehicle.
	ΔV^{lag}	Relative speed of the subject vehicle to the lag vehicle.
	V^{sub}	Subject speed.
	D^{rem}	Remaining distance.
	δ^{sub}	Driver characteristics.
GAP ACCEPTANCE MODELS		
Mahmassani and Sheffi (1981)	T_{cr}	Average critical gap when facing the first gap.
	β	Parameter in the gap acceptance model.
	Δ	Parameter in the gap acceptance model.

Table 4. Parameters of microscopic models. (continued)

Model	Parameter	Description
Taylor and Mahmassani (1998)	β_0	Mean critical gap when all other attributes are zero.
	β_{gap}	Parameter related to the gap and surrounding events.
	β_{int}	Parameter related to the intersection environment.
	β_{man}	Parameter related to the acting vehicle maneuver.
	β_{vo}	Parameter related to the vehicle/operator unit.
	β_{ia}	Parameter related to the interactions between attributes from different groups.
QUEUE ACCEPTANCE MODELS		
Jia et al. (2010)	β	Parameter in the queue discharge model.

IDM = intelligent driver model. LMRS = Lane-Changing Model with Relaxation and Synchronization. MITSIM - MIcroscopic Traffic SIMulator.

CALIBRATION NEEDS FOR OPERATIONAL MODELS: MESOSCOPIC AND MACROSCOPIC

Fundamental diagrams capture the relation between flow, density, and speed. They are widely used to represent flow propagation in mesoscopic simulation-based dynamic network assignment tools. A calibrated fundamental diagram can characterize the performance of the respective facility and/or network in a robust manner.

The Greenshields models in DYNASMART are the traffic flow models that needs to be calibrated based on the desired context (Mahmassani, Hou, et al. 2014). The Type 1 model is a dual-regime model with a constant free-flow speed for the free-flow conditions and modified Greenshields model for congested-flow conditions as shown in figure 30. The Type 1 model is applicable to freeways because it can accommodate dense traffic (up to 23,000 passenger cars per hour per lane (pc/hr/ln)) at near free-flow speeds.

Source: FHWA, 2019.

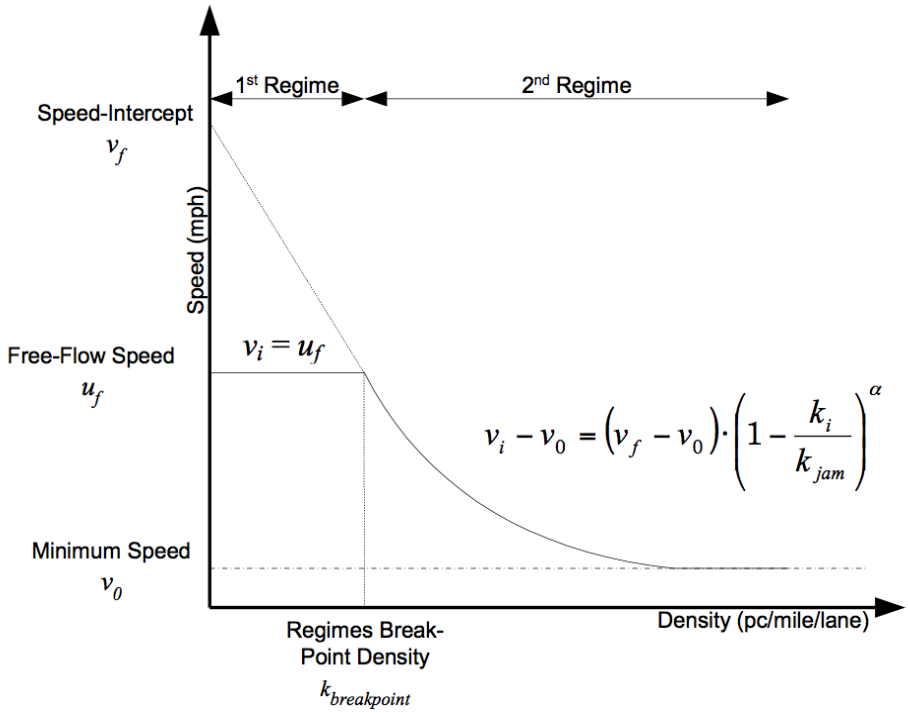


Figure 30. Graph. Type 1 modified Greenshields model (dual-regime model).

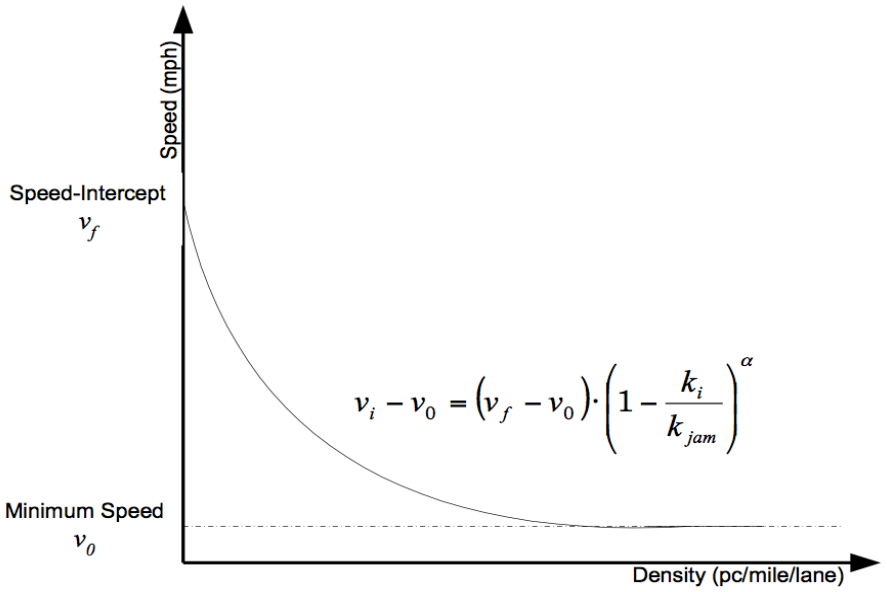
Te following equation mathematically specifies the Type 1 modified Greenshields model:

$$v_i = \begin{cases} u_f & 0 < k_i < k_{breakpoint} \\ v_0 + (v_f - v_0)\left(1 - \frac{k_i}{k_{jam}}\right)^{\alpha} & k_{breakpoint} < k_i < k_{jam} \end{cases}$$

Figure 31. Formula. Type 1 modified Greenshields model.

Where v_i is the speed on link i; v_f is the speed-intercept; u_f is the free-flow speed on link i; v_o is the minimum speed on link i; k_i is the density on link i; k_{jam} is the jam density on link i; $k_{breakpoint}$ is the breakpoint density; and a is the power term.

The Type 2 model is a single-regime Greenshields model that models both the free- and congested- flow conditions using a single relation as shown in figure 32. This model is more appropriate for arterials.



Source: FHWA, 2019.

Figure 32. Graph. Type 2 modified Greenshields model (single-regime model).

The Type 2 modified Greenshields model is expressed in the following mathematical equation:

$$v_i = v_0 + (v_f - v_0)\left(1 - \frac{k_i}{k_{jam}}\right)^{\alpha}$$

Figure 33. Formula. Type 2 modified Greenshields model.

In the above modified Greenshields models, the combination of values and joint probability distributions of v_f, u_f, v_o, k_{jam}, $k_{breakpoint}$, and a constitutes a library of parameters for the modified Greenshields models that should be calibrated.

There are other fundamental diagram relationships whose parameters could be added to the libraries of parameters in the proposed framework. For example:

- Greenberg's logarithmic model defined as

$$v_i = v_0 . ln\left(\frac{k_{jam}}{k_i}\right)$$

Figure 34. Formula. Greenberg's logarithmic model.

- Underwood's exponential model defined as

$$v_i = v_f . e^{-\left(\frac{k_i}{k_o}\right)}$$

Figure 35. Formula. Underwood's exponential model.

Where k_o is the optimal density which is a parameter that needs to be calibrated.

- Pipes' generalized model defined as

$$v_i = v_f \left(1 - \frac{k_i}{k_{jam}}\right)^n$$

Figure 36. Formula. Pipes' generalized model.

Where *n* is a parameter that needs to be calibrated.

One of the features of DYNASMART is the capability to consider the effect of weather condition on the following supply-side parameters (Mahmassani, Hou, et al. 2014):

- Traffic flow model parameters: speed-intercept, minimal speed, density breakpoint, jam density, and the shape term alpha;
- Link performance: maximum service flow rate, saturation flow rate, and posted speed limit adjustment margin;
- Left-turn capacity: g/c ratio;
- 2-way stop sign capacity: saturation flow rate for left-turn vehicles, saturation flow rate for through vehicles, saturation flow rate for right-turn vehicles;
- 4-way stop sign capacity: saturation flow rate for left-turn vehicles, saturation flow rate for through vehicles, saturation flow rate for right-turn vehicles; and
- Yield sign capacity: saturation flow rate for left-turn vehicles, saturation flow rate for through vehicles, saturation flow rate for right-turn vehicles.

The weather-sensitive Traffic Estimation and Prediction System (TrEPS) is implemented in DYNASMART to accurately estimate and predict the traffic status under various weather conditions. The effect of weather condition is captured using a corresponding weather adjustment factor (WAF) as

$$f_i^{weather\ event} = F_i . f_i^{normal}$$

Figure 37. Formula. Weather effect adjustment of model parameters in DYNASMART.

Where $f_i^{weather\ event}$ is the value of parameter i under a certain weather condition, f_i^{normal} represents the value of parameter i under normal weather condition, and F_i is the WAF for parameter i.

The WAF could be defined as a linear function of weather conditions combining the effect of visibility and precipitation intensity as the following mathematical formula:

$$F_i = \beta_{i0} + \beta_{i1} . v + \beta_{i2} . r + \beta_{i3} . s + \beta_{i4} . v . r + \beta_{i5} . v . s$$

Figure 38. Formula. Weather adjustment factor as a function of weather condition.

Where v is visibility; r is precipitation intensity of rain; s is precipitation intensity of snow; and β_{io}, β_{i1}, β_{i2}, β_{i3}, β_{i4}, β_{i5}, are the coefficients within the WAF function that need to be estimated and calibrated based on the context and the study area characteristics. Comprehensive libraries of these coefficients could be implemented in traffic analysis calibration tools. More coefficients could be added to the libraries as more complex correlations between the WAF of a traffic supply-related parameter and the weather-related variables are identified.

CALIBRATION NEEDS FOR STRATEGIC MODELS

Strategic models possess a higher level of complexity compared to operational models. Choice dimensions such as route choice, mode choice, and departure time choice are the main contributors in the complexity of strategic models. There is an extensive body of literature, in the transportation area, on choice models and strategies that could influence the behavior of individuals. The main emphasis of this section is on choice models that considered the effect of different transportation management strategies. Examples of strategic models from the literature are provided in the remainder of this section.

Frei et al. (2014) integrated demand models into weather-responsive network traffic estimation and prediction system methodologies. Therefore, the system encompasses weather-responsive traffic management (WRTM) strategies as well as active travel demand management strategies such as pre-trip information dissemination and rescheduling of school hours. Since the WRTM strategies aim short- to medium-term decisions, the behavioral choice dimensions that could be influenced by travel demand management strategies are route choice, departure time choice, and mode choice. Their proposed framework has three parts:

1. A base travel demand model with travel time and cost skims calibrated using multicriteria dynamic user equilibrium.
2. A network-level scenario manager with weather adjustment factors to update the travel time and cost skims according to the adjusted supply values using multicriteria dynamic user equilibrium.
3. A disaggregated demand model that gets updated based on the updated disaggregated user time and cost skims.

Eventually, the framework output influences the short- and medium-term traveler decisions. One of the models they developed using this framework was a departure time choice model that contains a scheduling cost function (C_s) defined as follow:

$$C_s = const + \beta_1(T + T_w) + \beta_2(SDE) + \beta_3(SDL)$$

Figure 39. Formula. Scheduling cost in a demand model proposed by Frei et al. (2014).

Where T is the travel time defined as the summation of free-flow travel time and extra travel time caused by recurrent congestion; SDE is the schedule delay early; SDL is the schedule delay late; β_1, β_2, and β_3 are cost parameters to be estimated and calibrated, referring to the cost of travel time per unit of time, cost per unit of time for arriving early, and cost per unit of time for arriving late, respectively; and $const$ denotes the alternative specific constant for departure time that should be estimated and calibrated.

Highway tolling and pricing strategies are among the strategies that are extensively studied to capture their effect on travelers' behavior. In the traditional four-step trip-based models, pricing has a first-order impact on the trip mode and trip time of day choice dimensions. It has a second-order impact on trip generation and trip distribution models. On the other hand, for the activity-based models, it has a first-order impact on the trip mode, tour mode, and tour time of day choice dimensions. The second-order impact of pricing for the activity-based models appears in the tour primary destination and tour generation models (Perez, Batac and Vovsha 2012).

Perez et al. (2012) proposed a generalized cost function to evaluate the utility of each network link of a route according to the following formula:

$$G_k = \beta_{1k} T_k + \beta_{2k} C_k$$

Figure 40. Formula. Link-level cost function proposed by Perez et al. (2012).

Where k refers to the vehicle type and auto occupancy class; T_k is the travel time; C_k is the travel cost including the toll; and β_{1k} and β_{2k} are coefficients for travel time and travel cost that should be estimated and calibrated.

A more generalized version of the above cost function could be formulated in order to incorporate other choice dimensions such as the mode choice utility function.

$$U_m^p = \gamma_m^p + \beta_{1m}^p T_m + \beta_{2m}^p C_m + \sum_v \beta_{vm}^p S_v$$

Figure 41. Formula. A generalized mode choice utility function.

Where m represents the set of modes including auto occupancy classes; p is the travel purpose and other possible segments; v refers to the person, household, and zonal variables; T_m is the travel time by mode; C_m is the travel cost by mode; S_v symbolizes the person, household, and zonal variables; γ_m^p is the mode-specific constant for each trip purpose/segment; and the parameters that need estimation and calibration are β_{1m}^p, β_{2m}^p, and β_{vm}^p which are the coefficients for travel time, travel cost, and person/household/zonal variables by travel mode and trip purpose/segment.

The Logsum of the utility function for the mode choice model could be used to generate the impedance functions of upper-level choice dimensions such as the time-of-day choice and destination choice models. The time-of-day choice utility function could be formulated by adding the Logsums of the utility function for the mode choice model in the following manner:

$$V_t^p = \mu \times ln\left[\sum_m \exp\left(U_{mt}^p\right)\right] + \sum_v \beta_{vt}^p S_v$$

Figure 42. Formula. A time-of-day choice utility function.

Where U_{mt}^p is the mode choice utility for the trip purpose p traveled by the mode m at time of day period t, and μ and β_{vt}^p are the parameters that need to be estimated and calibrated. The parameter μ is a scaling factor that possesses a value in the unit interval.

Similar to the time-of-day choice model, the destination choice utility function could be written as:

$$W_{od}^{p} = \eta \times ln\left[\sum_{m} \exp\left(U_{o,d,m,p,t}^{p}\right)\right] + \ln\left(A_{d}^{p}\right)$$

Figure 43. Formula. A destination choice utility function.

Where $U_{o,d,m,p,t}^{p}$ refers to the mode choice utility for the trip purpose *p* traveled from zone *o* to zone *d* using the mode *m* at time of day period *t*; A_{d}^{p} is the destination zone attraction (size variable) for each trip purpose, such as total employment for work purpose, enrollment for school purposes, and retail employment for non-work purpose; and η is a parameter, possessing a value in the unit interval, that needs to be estimated and calibrated.

Reliability is another important aspect that should be incorporated in choice dimension models. It represents the uncertainty level with respect to congestion levels and travel time. More reliable alternatives are generally preferred because of various reasons including but not limited to negative consequences of late arrival at the destination, needing a buffer time to avoid late arrival, and discomfort caused by the uncertainty of travel time at a given time. Therefore, incorporating reliability variables in choice dimension models could potentially improve the model performance.

Vovsha et al. (2013) proposed a generalized model as the highway utility function that includes basic components such as time and distance as well as a reliability-related component. The cost of path is incorporated in a non-linear way to account for the effects of income and vehicle occupancy. They formulated the highway utility function as follow:

$$U = \Delta + \beta_1 \times Time \times (1 + \beta_2 \times D + \beta_3 \times D^2) + \beta_4 \times \left[\frac{Cost}{I^{\beta_5} \times O^{\beta_6}}\right] + \beta_7 \times \frac{STD}{D}$$

Figure 44. Formula. Highway utility function proposed by Vovsha et al. (2013).

Where Δ is the alternative-specific constant for tolled facilities; *Time* is the average travel time; *D* is the travel distance; *Cost* refers to the monetary cost of using the facility including the tolls, parking, and fuel; *I* is the household income of the traveler; *O* indicates the vehicle occupancy; *STD* is the day-to-day standard deviation of the travel time. There are seven parameters ($\beta_1, \beta_2, \ldots, \beta_7$) that should be estimated and calibrated to fully specify the model.

The authors listed 11 key behavioral insights, which are useful for implementing transportation management strategies, that could be derived from the values estimated and calibrated for the parameters (Vovsha, et al. 2013):

- Variation in the value of time (VOT) across highway users.
- Income and willingness to pay.
- Vehicle occupancy and willingness to pay.
- Constraints on time-of-day shifting.
- Importance of value of reliability (VOR) and its association with VOT.
- Effect of travel distance on VOT and VOR.

- Evidence of negative toll bias.
- Hierarchy of likely responses to change in tolls and congestions.
- Summary of user segmentation factors.
- Avoiding simplistic approaches to forecasting.
- Data limitations and Global Positioning System-based data collection methods.

Table 5 summarizes the strategic model parameters introduced in this section that should be calibrated. There are various other models that could be added to the listed strategic models.

Table 5. Parameters of strategic models.

Model	Parameter	Description
SCHEDULING COST FUNCTION		
Frei et al. (2014)	β_1	Cost of travel time per unit of time.
	β_2	Cost per unit of time for arriving early.
	β_3	Cost per unit of time for arriving late.
SIMPLE ROUTE UTILITY FUNCTION		
Perez et al. (2012)	β_{1k}	Coefficient for travel time of vehicle type k.
	β_{2k}	Coefficient for travel cost of vehicle type k.
MODE CHOICE UTILITY FUNCTION		
Perez et al. (2012)	γ_m^p	Mode-specific constant for each travel mode and purpose/segment.
	β_{1m}^p	Coefficient for travel time for each travel mode and purpose.
	β_{2m}^p	Coefficient for travel cost for each travel mode and purpose.
	β_{vm}^p	Coefficient for person, household, and zonal variables for each travel mode and purpose.
TIME-OF-DAY CHOICE UTILITY FUNCTION		
Perez et al. (2012)	μ	Scaling coefficient that should be in the unit interval.
	β_{vt}^p	Coefficient for person, household, and zonal variables for each travel purpose and time of day.
DESTINATION CHOISE UTILITY FUNCTION		
Perez et al. (2012)	η	Scaling coefficient that should be in the unit interval.

Table 5. Parameters of strategic models. (continued)

Model	Parameter	Description
HIGHWAY UTILITY FUNCTION		
Vovsha et al. (2013)	β_1	Coefficient for travel time.
	β_2	Coefficient reflecting the impact of travel distance on the perception of travel time.
	β_3	Coefficient reflecting the impact of travel distance on the perception of travel time.
	β_4	Coefficient for travel cost.
	β_5	Coefficient reflecting the impact of income on the perception of cost.
	β_6	Coefficient reflecting the impact of occupancy on the perception of cost.
	β_7	Coefficient reflecting the impact of travel time reliability.

The strategic models introduced here are a few examples of various models available in the literature. Compared to the operational models discussed in previous sections, development of libraries for these models is more challenging. Various strategies not only impact the value of the variables used in the formulation but also could influence the parameter values. As a result, extensive libraries of strategic models should be produced by combining different management strategies, external conditions (e.g., weather conditions and incidents), and internal behavioral characteristics of simulation agents. These libraries could be used to generate the desired scenarios. In this case, a scenario would be essentially a set of strategies, models, and parameters.

CONCLUSIONS

There are several issues in the conventional calibration processes that reduce the effectiveness of the simulation model, prevent the applicability of the models to new areas and context, and produce unrealistic behavior. These issues could be addressed by introducing new elements in the calibration framework such as libraries of models, parameters, and relevant joint probability distributions capturing correlation effects among the parameters. Besides these elements, using trajectory-based data could improve the quality and accuracy of the model outputs. Last but not least, the adaptive updating process is another feature in the proposed framework that could reduce the computational effort for the calibration process.

Three categories of models were introduced: microscopic operation models, mesoscopic and macroscopic operational models, and strategic models. Developing comprehensive libraries of models and parameters is an essential component of the proposed calibration framework. As a result, a data structure should be constructed to maintain these libraries. The capability of updating the models and parameters of these libraries and the correlation among these libraries would be clearly specified in the proposed data structure.

CHAPTER 5. METHODOLOGY AND FRAMEWORK

Note: Unless accompanied by a citation to statute or regulations, the practices, methodologies, and specifications discussed below are not required under Federal law or regulations.

Table 6 presents the five major components of the proposed calibration framework. These components will be presented throughout the main body of this chapter. For now, a high-level summary is provided. Although these components of the framework could indeed be applied individually or separately, applying the fuller set of components would likely achieve better models.

Table 6. Primary components of the calibration framework.

Calibration Concept	Intent of Concept
Scenarios	Calibrate unique models for typical combinations of demand, weather, and incidents.
Robustness	Weighted sensitivity analysis of future demands and scenarios.
Parameter correlation	Library of realistic driver types for different levels of aggressiveness.
Local density	Develop relationships between driver behavior and degree of saturation.
Vehicle trajectories	Comprehensive and compact representation of system performance.

The first component is ***scenarios***. Any robust calibration framework should address the highly variable nature of traffic demand, weather, incidents, work zones, and special events. Analysis of "typical" conditions is no longer considered safe, or helpful. In essence, there should be a unique simulation model created to represent, at the very least, each of the most common combinations of the above highly variable elements.

The second component is ***robustness***. This component is important because analysis of future conditions is naturally governed by uncertainty. The level of uncertainty often relates to the lack of data. For example, the further into the future one wishes to forecast, the more uncertainty may exist around traffic demand levels. That said, future traffic demand levels may be more certain in some cities than others. Next, when it comes to analysis of emerging technologies (e.g., connected and automated vehicles) and traffic management strategies (e.g., advanced traffic demand management), the amount of available, relevant data may vary widely. For robust model parameters backed up by extensive data, a direct, highly confident calibration is possible. For uncertain model parameters having no data, the best that can be accomplished is a sensitivity analysis that accounts for all possibilities. For those model parameters backed up by, for lack of a better term, "medium" amounts of data, a Bayesian-style approach is needed. In this manner the model parameter is flagged with medium robustness, calibration accuracy is augmented by available data to the maximum possible extent, but a sensitivity analysis is nonetheless needed to forecast future possibilities.

The third component is ***parameter correlation***. This component recognizes the pitfalls of calibrating model parameters independently, as if they were not correlated. If parameter correlations are not explicitly recognized or incorporated into the calibration process, fundamental soundness of the model may be compromised. For example, in a microscopic simulation model, the calibration process may produce driver type characteristics and profiles that do not exist in the real world. To mitigate this risk, a library of congruous model parameter settings should be provided. The prior chapter describes model parameters for microscopic, mesoscopic/macroscopic, and strategic models.

The fourth component is ***local density***. This component recognizes the strong correlation between traffic congestion and driver behavior. The term 'local density' refers to the density of traffic in the local vicinity of a vehicle. Beyond this local spatial area, congestion levels would not affect driver behavior. However within the vicinity of a driver's vehicle, the density of traffic in that area will profoundly affect their aggressiveness. Researchers believe a 500-foot radius around the vehicle is the best spatial area for predicting driver behavior. In cases where it is not possible to obtain densities near vehicle, a link-based or segment-based density may suffice, albeit with some loss of precision. The beauty of the local density concept is that the fundamental relationships between driver behaviors can be developed under existing conditions, and then re-used in future models having different densities on each segment.

The fifth and final component is ***vehicle trajectories***. Use of trajectories in the calibration process is strongly encouraged for many reasons. Chief among these is that simulation models calibrated only to aggregate performance measures have been found to exhibit highly unrealistic driver behaviors.

OVERALL FRAMEWORK

Figure 18 in chapter 4 illustrates the proposed framework for the calibration process. As the first step, after refining the input data and performing an estimation process on the observed data, the appropriate set of parameters are selected/estimated. Simultaneously, the external conditions of the system are incorporated as exogenous input variables. In the proposed framework, the estimated parameters and external conditions are fed to the simulation model in terms of libraries of models, parameters, and related joint probabilities. These library components and probabilities would be used to develop various scenarios and agents as inputs for the simulation model. Using a set of scenarios and agents, the model will be calibrated based on the observations and forecast input variables collected from other sources such as census data.

An important difference between the conventional calibration process and the proposed calibration process is the manner in which the model is updated over time. The conventional approach involves a full re-calibration process when the model is applied to a new context. On the other hand, in the proposed framework, this computationally intensive process is avoided to some extent by updating the scenarios and libraries. Adaptive updating processes such as Bayesian techniques could be used to update the library parameters and their associated joint probability distributions.

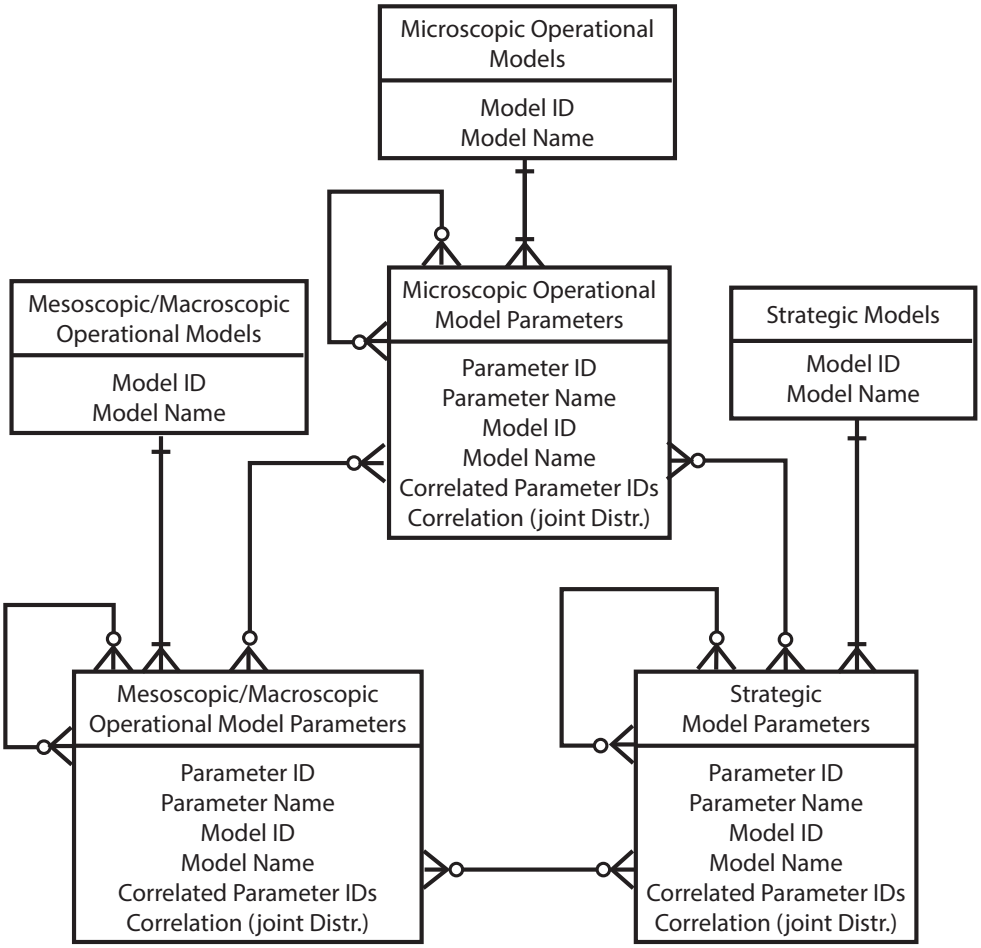
LIBRARY OF PARAMETERS

A key component of the framework is a library of parameters that act as a database accessed by the parameter estimator. Figure 45 shows the database structure for the library of parameters. The libraries could be classified into three major categories:

- Microscopic operational models and their associated parameters.
- Mesoscopic/Macroscopic operational models and their associated parameters.
- Strategic models and their associated parameters.

Each element in a parameter library is a parameter. Therefore, a unique key is assigned to each parameter. Besides the parameter ID, a model ID is also assigned to each parameter which represents the one-to-many relationship between the model library and the parameter library. This relationship is a mandatory relationship because every parameter belongs to a specific model defining the behavior of agents in the transportation system. Therefore, a parameter inherits the model attributes such as the conditions that are used to establish the validity of the model. On the other side of the one-to-many relationship, every model possesses at least one parameter that is stored in the respective parameter library.

One of the most important aspects of the proposed calibration methodology is preserving the correlation among the model parameters by defining joint probability distributions among the parameters. This is another attribute of a variable that should be stored in the parameter libraries. A set of correlated parameter IDs and their respective correlations represented as joint probability distribution functions could be stored in the parameter libraries. The recursive many-to-many relationships and the inter-library relationships symbolize the correlation among the model parameters. The recursive (intra-library) relationship captures the correlation among the parameters that belong to the same model category. For example, values of some parameters of a car following model could be associated with certain values of a lane changing model, when the models are utilized to simulate the behavior of an aggressive (or conservative) driver. The inter-library relationship captures the correlation among the parameters that belong to the different model categories. For instance, this correlation could exist among the parameters of the models used at different stages of the four-step model.



Source: FHWA, 2019.

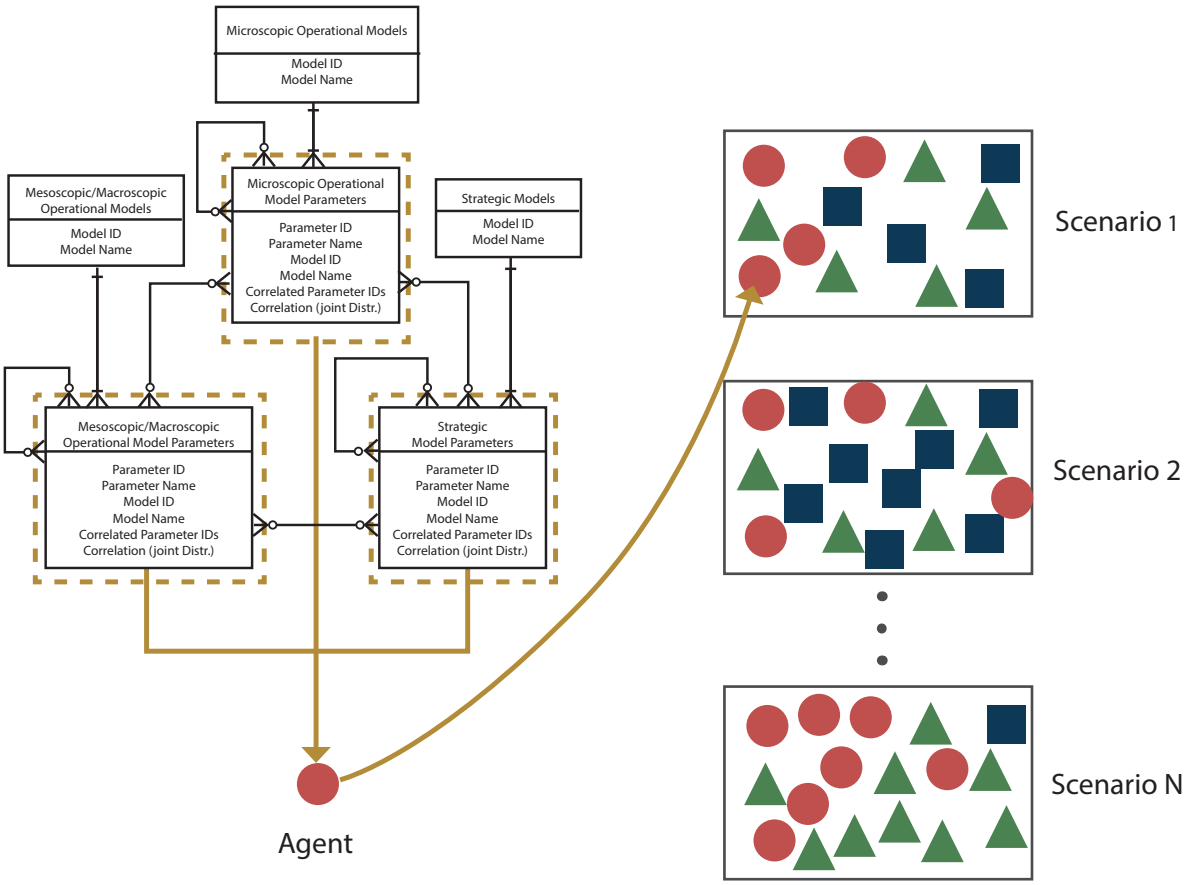
Figure 45. Flowchart. Entity relationship diagram of the model parameter libraries.

The proposed data structure has the capability of introducing new models in the libraries. The model, as well as its parameters, should be added to the appropriate library categories. Once these parameters are entered in the database, they should be mapped to the other correlating parameters and the correlation should be defined using a joint distribution function. Besides adding a new element to the database, the attributes of each model and each parameter could be updated using adaptive processes as new data and observations become available over time. Bayesian techniques offer a formal framework for combining previous values with those based on new observations.

The structure also allows defining sets of deeper parameters that connect the initial parameter values to the scenarios attributes/features which is called reparameterization. The intent is for the deep parameters to remain valid over time. For example, the fraction of automated vehicles (market penetration) in the traffic mix could be made a parameter in the selection of an appropriate fundamental diagram for the corresponding scenario. Moreover, fundamental mechanisms and structural relations that are likely to remain valid as conditions change over time could be specified in the database. This rules out certain machine learning and data-driven approaches that involve training on specific datasets, and therefore have no ability to respond to structural change.

AGENT AND SCENARIO GENERATION PROCESS

Once the libraries of parameters are developed, samples could be drawn from the joint distributions (or marginal distributions) of the parameters in order to generate agents with specified behavior. Based on the models selected for the sampling process, an agent could represent a vehicle, pedestrian, cyclist, or could represent a traffic control entity such as a traffic light, a ramp metering system, or a speed control system, or could be a representation of the condition in which the other agents interact in the specified transportation system; for example, the weather condition, the road network properties, or the road surface conditions, etc. A combination of agents generated in this manner would create a scenario that serves as an input of the simulation tool. This process is schematically shown in figure 46.



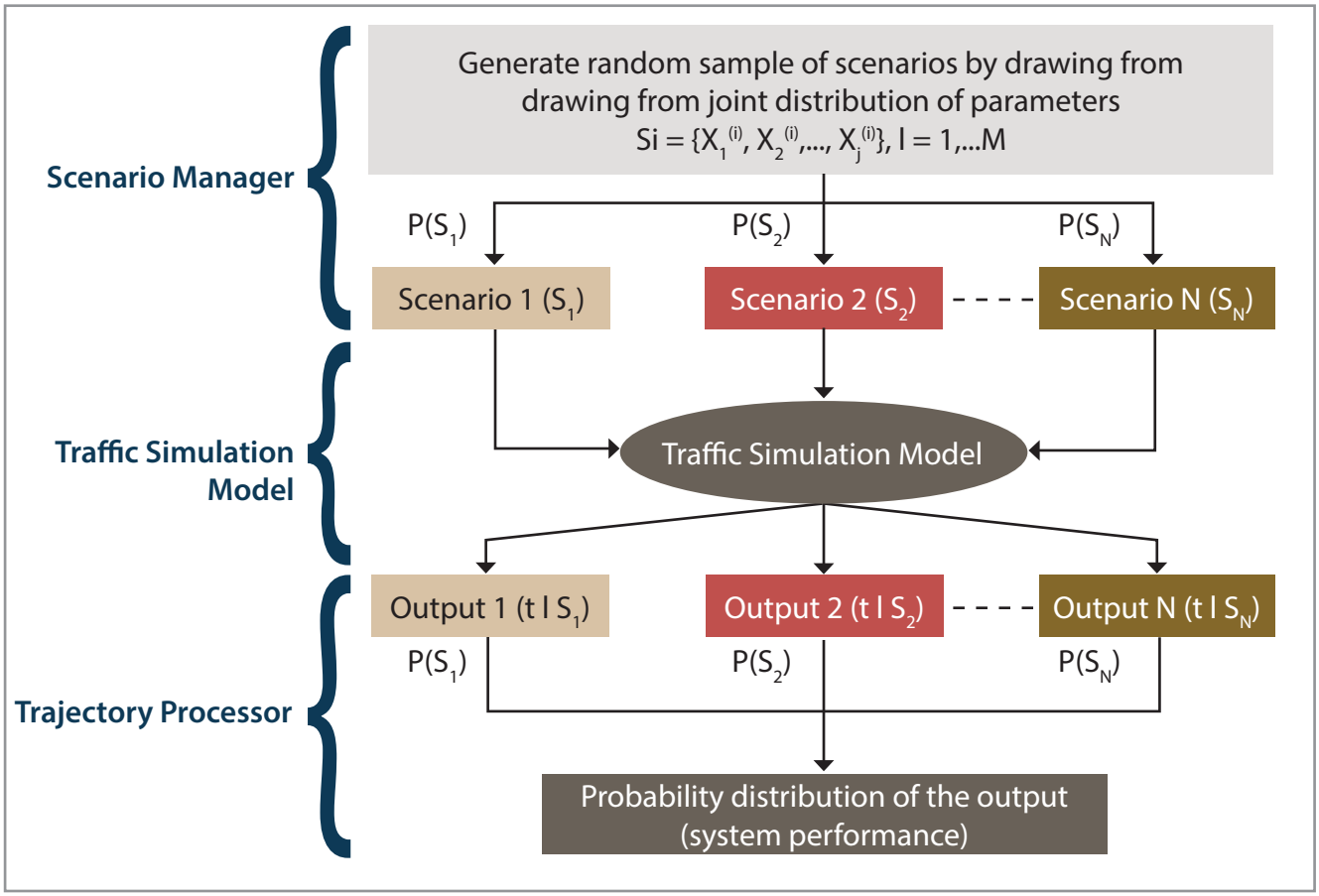
Source: FHWA, 2019.

Figure 46. Illustration. Process of generating agents and scenarios.

Application of a transportation model for a future scenario would entail sampling/selecting from these libraries different agent types reflecting the particular mix expected. Generating different scenarios would essentially reflect the behaviors of these agents under various conditions. If there is certainty regarding the value of a parameter in the future, the parameter value would be the same through all generated scenarios. An example could be the geometric characteristics of the road network under consideration if, until the target year, there is not any major construction planned for the specified road segments.

SCENARIO-BASED ANALYSIS

The method shown in figure 47 is useful when dealing with parameters that have some level of uncertainty.



Source: FHWA, 2019.

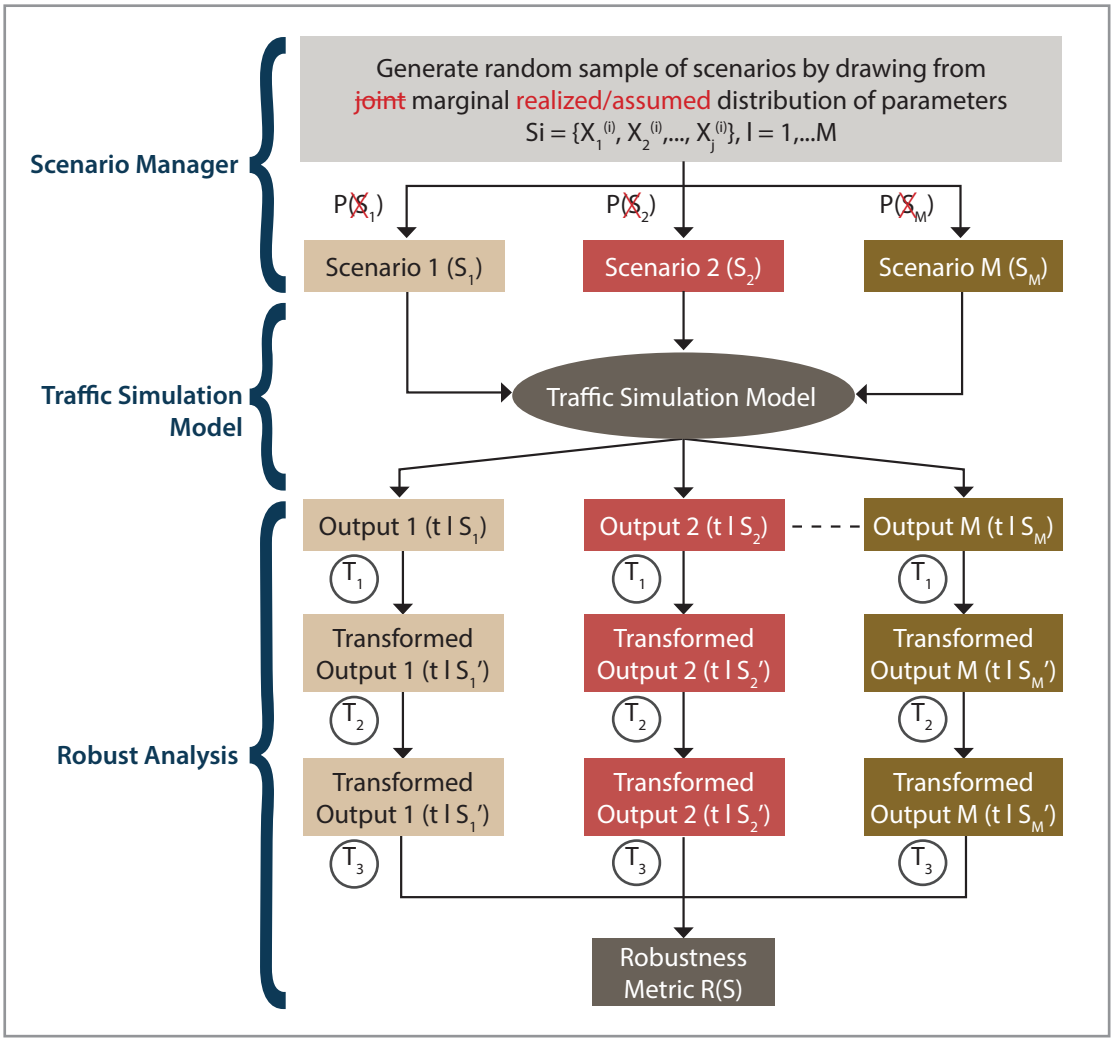
Figure 47. Flowchart. Scenario-based analysis.

In this situation, there is a limited set of values that is applicable to the parameters under consideration. The joint distributions (and correlations) among these parameters are known. The process combines different scenarios to generate the result of analyzing multiple scenarios. When there is enough information regarding the association between different sets of parameters deployed in the simulation tool, a joint distribution of parameter sets could be characterized. Random samples could be drawn from the set of scenarios using random sampling techniques such as the Monte Carlo methods. The sample size depends on the confidence level and anticipated errors in the respective scenarios. Each scenario, comprised of a different set of parameters, would be incorporated in the traffic simulation tool. The outputs of the simulation tool should be coupled with their respective scenario probability. In the trajectory processor step, these outputs should be combined with their associated weights (probabilities) to generate a probability distribution of the output under consideration. This approach is appropriate when there are some uncertainties involved but there is enough knowledge of the scenarios to generate a joint distribution of the scenarios. This approach could be useful to evaluate the transportation system under different demand distributions or under various incident scenarios.

ROBUSTNESS-BASED ANALYSIS

In contrast to the previous method, in some cases, there is a lack of information regarding the possible values of a parameter or the correlation among parameters. Examples of situations with such deep uncertainty could be fundamental changes to the transportation network environment at the target year, new transportation policies that have not been previously tested on the specific context, or introduction of new technologies such as connected and autonomous vehicles. As a result, a different approach should be used to combine the output of different scenarios when deep uncertainty exists. Instead of the joint distribution of the parameters, the marginal distributions might be known or assumed. As a result, the probability by which different scenarios occur is also unknown. As could be seen in figure 47 and figure 48, the main difference is in how the outputs of the simulation tool are treated. In the robustness-based analysis, according to a robustness measure selected by the analyst, the outputs are transformed and combined. The selection of the robustness metric depends on the suitability of the metric for the decision-context, the decision maker's level of risk aversion, and the decision maker's objective function, e.g., maximizing the performance measure or minimizing the variance of the performance measure. There are generally four types of robustness metrics: expected value metrics, metrics of higher orders (variance, skewness, and kurtosis), regret-based metrics, and satisficing metrics. Examples of these metrics and their pertinent transformations are provided in table 7 (McPhail, et al. 2018).

As shown in the graph, three levels of transformation are performed to transform the performance measure output of the simulation tool to a specified robustness metric. First, a system performance transformation (T1) is conducted to derive a value that is more compatible with the robustness metric (for example, change the travel time to regret-based congestion as the difference between the travel time of the selected scenario and the free flow travel time). Then, a scenario subset selection (T2) is performed that selects a subset (one/a few/all) of the scenarios based on the robustness metric and desired level of risk aversion (for example, the 75^{th}, 80^{th}, 90^{th} percentiles among travel time of all scenarios). The robustness metric calculation is the third transformation. This step combines the values derived in the previous steps for the selected scenarios to calculate the robustness metric for the scenarios (for example, the expected value/skewness/kurtosis of the calculated robustness metric for the selected scenarios). A list of sample robustness metrics that would be used in this study is shown in table 6. The "identity" cells mean that no specific transformation is performed at the stage and the performance value is directly passed to the next step.



Source: FHWA, 2019.

Figure 48. Flowchart. Robustness-based analysis.

Table 7. Transformations for robustness metrics.

Robustness Metric	T1	T2	T3
Maximin	Identity	Worst case	Identity
Maximax	Identity	Best case	Identity
Laplace's principle of insufficient reason	Identity	All	Mean
Hurwicz optimism-pessimism rule	Identity	Worst and best cases	Weighted mean
Minimax regret	Regret	Worst case	Identity
N^{th} percentile minimax regret	Regret	N^{th} percentile	Identity
Undesirable deviations	Regret	Worst-half cases	Summation
Percentile-based skewness	Identity	10^{th}, 50^{th}, and 90^{th} percentiles	Skew
Percentile-based peakedness	Identity	10^{th}, 25^{th}, 75^{th}, and 90^{th} percentiles	Kurtosis

Source: McPhail et al. 2018.

CONCLUSIONS

This chapter has presented the framework proposed for calibrating traffic operational models and simulation tools for application to future conditions. The framework is predicated on four key notions: (1) That the models themselves are responsive to the features of the future scenario (i.e., that the descriptors that specify a particular scenario be included in the model specification); (2) the definition of a library of model parameters corresponding to different types of agents under varying conditions; (3) the role of scenarios in both specifying future conditions of interest as well as triggering certain ranges of parameter values for those scenarios; and (4) the potential for robustness analysis in situations where uncertainty about future scenarios or conditions is large.

CHAPTER 6. CASE STUDY

Note: Unless accompanied by a citation to statute or regulations, the practices, methodologies, and specifications discussed below are not required under Federal law or regulations.

The objective of this case study is to conduct a proof-of-concept test of the proposed framework. The focus is on the implementation of the major components of the calibration framework. As explained in the previous chapter there is an association between the library of parameters and the agents introduced into the model. Therefore, instead of sampling from the distribution of parameters and running a microsimulation tool, different types of agents were developed considering the correlation between the behavioral model parameters. The study uses an integrated traffic telecommunication tool that was developed at Northwestern University to evaluate the operational performance impacts of different agents in a mixed traffic environment at different market penetrations of the agents (Talebpour, Mahmassani and Bustamante 2016).

The microsimulation tool is a special-purpose platform for simulating mixed traffic conditions on freeways with the possibility of including connected vehicles and autonomous vehicles in the system. In the current case study, three types of drivers (agents) were considered: regular vehicles, connected vehicles, and automated vehicles. As will be discussed later, two agent types, namely, aggressive drivers and conservative drivers, were introduced by changing some of the parameters of the regular vehicles. The testbed uses a 3.5-mile section of I-290 in Chicago, Illinois (shown in figure 49). The following sections discuss the behavioral models embedded in the simulation tool.

Source: FHWA, 2019.

Figure 49. Illustration. I-290E study segment in Chicago, IL.

ACCELERATION (CAR FOLLOWING) FRAMEWORK

In the simulation platform, distinct car-following models are defined to specify the behavior of each agent: (1) isolated-manually driven vehicles (regular vehicles), (2) connected-manually driven vehicles (connected vehicles), and (3) isolated-automated vehicles (automated vehicles).

ISOLATED-MANUALLY DRIVEN VEHICLES

In the microsimulation platform, isolated-manually driven vehicles use the acceleration model first developed by Hamdar et al. (2008) and extended by Talebpour et al. (2011). The model was formulated based on Kahneman and Tversky's prospect theory. Two value functions, one for modeling driver behavior in congested regimes and the other for modeling driver behavior in uncongested regimes, were introduced. The following formula shows the value function for the uncongested regime:

$$U_{PT}^{UC}(a_n)=\frac{\left[w_m+(1-w_m)\left(\tanh\left(\frac{a_n}{a_0}\right)+1\right)\right]}{2}\left[\frac{\left(\frac{a_n}{a_0}\right)}{\left(1+\left(\frac{a_n}{a_0}\right)\right)^{\left(\frac{\gamma-1}{2}\right)}}\right]$$

Figure 50. Formula. Value function for the uncongested regime.

Where U_{PT}^{UC} denotes the value function for the uncongested traffic conditions. $\gamma>0$ and w_m are parameters to be estimated and calibrated and $a_o = 1\ m/s^2$ is used to normalize the acceleration. On the other hand, the following formula shows the value function for the congested regime:

$$U_{PT}^{C}(a_n)=\frac{\left[w'_m+(1-w'_m)\left(\tanh\left(\frac{a_n}{a_0}\right)+1\right)\right]}{2}\left(\frac{a_n}{a_0}\right)^{\gamma'}$$

Figure 51. Formula. Value function for the congested regime.

Where U_{PT}^{C} denotes the value function for the congested traffic conditions. $\gamma' > 0$ and w'_m are parameters to be estimated and calibrated. At each evaluation time step, the driver evaluates the gain from a candidate acceleration selected from a feasible set of values. The surrounding traffic condition is taken into consideration by the driver throughout the acceleration evaluation process. The driver utilizes the following binary probabilistic regime selection model to evaluate each acceleration value:

$$U_{PT}(a_n)=P(C).U_{PT}^{C}+P(UC).U_{PT}^{UC}$$

Figure 52. Formula. Binary probabilistic regime selection model.

Where U_{PT}, $P(C)$, and $P(UC)$ denote the expected value function, the probabilities of driving in a congested traffic condition, and the probability of driving in uncongested traffic condition, respectively. After calculating the expected value function, the total utility function for acceleration could be written as follows:

$$U(a_n) = (1 - p_{n,i})U_{PT}(a_n) - p_{n,i} w_c k(v, \Delta v)$$

Figure 53. Formula. Total utility function for the choice of acceleration.

Where *Pn,i* is the crash probability. Finally, the following probability density function is used to evaluate the stochastic response of the drivers:

$$f(a_n) = \begin{cases} \dfrac{e^{\beta_{PT} U(a_n)}}{\int_{a_{\min}}^{a_{\max}} e^{\beta_{PT} U(a')}\, da'} & a_{\min} < a_n < a_{\max} \\ 0 & Otherwise \end{cases}$$

Figure 54. Formula. Probability density function for the evaluation of drivers' stochastic response.

Where β_{PT} is the sensitivity of choice to the utility $U(a_n)$.

CONNECTED-MANUALLY DRIVEN VEHICLES

These vehicles are capable of exchanging information with other vehicles and infrastructure-based equipment. The information is exchanges through the vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications networks. As a result, the driver receives information about other connected vehicles as well as updated information containing transportation management center decisions (e.g., real-time changes in speed limit). The drivers' behavior may change based on the information conveyed to the driver. The reliability and the frequency of the information received by the driver plays a significant role in the drivers' behavior and on the overall performance of the traffic network.

An active V2V communication network allows the drivers to be aware of other drivers' behavior, the driving environment, road condition, and weather condition. As a result, the driving behavior could be modeled using a deterministic acceleration modeling framework. The simulation tool utilizes the Intelligent Driver Model (IDM) to model this connected environment. Because the IDM is able to capture various congestion dynamics and provides greater realism than most of the deterministic acceleration modeling frameworks.

The acceleration model specified by the IDM entails the vehicle's current speed, the ratio of the current spacing to the desired spacing, the difference between the leading and the following vehicles' velocities, and subjective parameters such as desired acceleration, desired gap size, and comfortable deceleration.

$$a_{IDM}^{n}(s_n, v_n, \Delta v_n) = \bar{a}_n \left[1 - \left(\frac{v_n}{v_0^n}\right)^{\delta_n} - \left(\frac{s^*(v_n, \Delta v_n)}{s_n}\right)^2\right]$$

$$s^*(v_n, \Delta v_n) = s_0^n + T_n v_n + \frac{v_n \Delta v_n}{2\sqrt{\bar{a}_n \bar{b}_n}}$$

Figure 55. Formula. The intelligent driver acceleration model.

Where δ_n is the free acceleration exponent; T_n is the desired time gap; a_n is the maximum acceleration; b_n is the desired deceleration; s_0^n is the jam distance; and v_0^n is the desired speed. These parameters need to be calibrated to better capture the behavior of connected vehicles.

If the V2V communication network is inactive, the driving behavior of connected vehicles would be similar to that of isolated-manually driven vehicles. In the presence of V2I communications, the TMC decisions, such as the speed limits in the case of speed harmonization, could be transferred to the drivers. However, their reaction times would still be like regular drivers.

ISOLATED-AUTOMATED VEHICLES

Automated vehicles can continuously monitor other vehicles in their vicinity, which results in a deterministic behavior in interacting with other drivers. Furthermore, they can quickly react to any perturbations in the driving environment. Therefore, the car-following behavior of automated vehicles could be specified by a deterministic modeling framework. Talebpour and Mahmassani (2016) developed a car-following model for automated vehicles based on the previous simulation studies by Van Arem et al. (2006) and Reece and Shafer (1993). They simulated similar individual sensors installed on the automated vehicles in order to generate the input data for the acceleration model. Figure 56 shows the sensor formation of an automated vehicle with the following specifications:

- Smart Micro Automotive Radar (UMRR-00 Type 30) with 90m±2.5 percent detection range and ±35 degrees horizontal field of view.
- The sensing information is updated every 50 ms.
- The sensors are capable of tracking 64 objects.

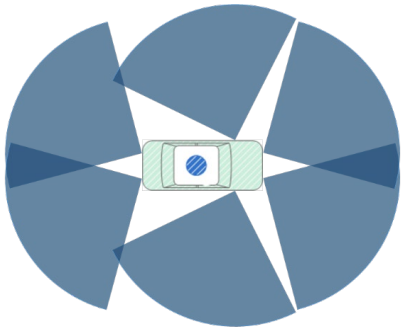


Figure 56. Illustration. Radar sensor formation on an automated vehicle.

Considering the sensor range and limitations in accuracy, it is important for automated vehicles to be ready to react to any situation outside of their sensing range once it is detected (e.g., a vehicle at a complete stop right outside of the sensors detection range). Furthermore, if a leader is spotted, the speed of the automated vehicle should be adjusted in a way that allows it to stop if the leader decides to decelerate with its maximum deceleration rate and reach a full stop. Considering different situations that involve immediate reaction of the automated vehicle, the maximum safe speed can be calculated using the following equations:

$$\Delta x_n = (x_{n-1} - x_n - l_{n-1}) + v_n \tau_n + \frac{v_{n-1}^2}{2a_{n-1}^{decc}}$$

$$\Delta x = min\{SensorDetectionRange, \Delta x\}$$

$$v_{max} = \sqrt{-2a_i^{decc}\Delta x}$$

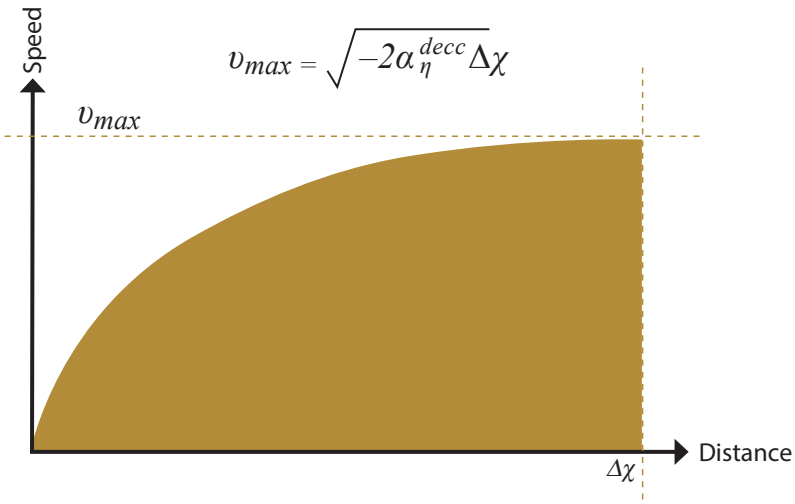
Figure 57. Formula. Maximum speed of automated vehicles.

Where n and n-1 represent the automated vehicle and its leader, respectively; x_n is the position of vehicle n; l_n is the length of vehicle n; is the speed of vehicle n; τ_n is the reaction time of vehicle n; and a_n^{decc} is the maximum deceleration of vehicle n. Figure 59 represents the concept of maximum safe speed; any speed below the curve is considered to be safe.

Besides the safety constraint, the following formula, adopted from the model proposed by Van Arem et al. (2006), updates the acceleration of the automated vehicle at every decision point:

$$a_n^d(t) = k_a a_{n-1}(t-\tau) + k_v(v_{n-1}(t-\tau) - v_n(t-\tau)) + k_d(s_n(t-\tau) - s_{ref})$$

Figure 58. Formula. Acceleration model for automated vehicles.



Adapted from Reece and Shafer, 1993.

Figure 59. Diagram. Maximum safe speed curve.

Where a_n^d is the acceleration of vehicle n, and k_a, k_v, and k_d are model parameters that need to be calibrated. s_n is the spacing and s_{ref} is the maximum between the minimum distance (s_{min}), following distance based on the reaction time (s_{system}), and safe following distance (s_{safe}). In this study, the minimum distance is set at 2.0 m and is calculated according to the following formula:

$$s_{safe} = \frac{v_{n-1}^2}{2}\left(\frac{1}{a_n^{decc}} - \frac{1}{a_{n-1}^{decc}}\right)$$

Figure 60. Formula. Safe following distance formula.

Finally, the acceleration of the automated vehicle can be calculated using the following equation:

$$a_n(t) = min(a_n^d(t), k(vn_{max}()))$$

Figure 61. Formula. Acceleration of automated vehicles.

Where k is a model parameter. Van Arem et al. (2006) suggested using the following values for the model parameters: $k = 1$, $k_a = 1$, , $k_v = 0.58$ and $k_d = 0.1$

LANE-CHANGING MODEL

The lane-changing behavior in a multilane traffic, as an essential element in a microsimulation model, is prone human error, particularly at high speed in high-density environments. In a connected and automated environment, the lane changing behavior is expected to occur with lower risk and to incorporate fewer abrupt maneuvers than human-negotiated cases.

Talebpour et al. (2016) developed a game-theoretic lane-changing model that captures the dynamic interactions between drivers during discretionary and mandatory lane-changing maneuvers and introduced a game structure to model the behavior when drivers are not aware of the nature of the lane-changing maneuver (i.e., mandatory vs. discretionary). They proposed two game types:

- A two-person non-zero-sum non-cooperative game under complete information to model lane-changing decisions when drivers/automated vehicles are aware of the nature of lane-changing maneuver.
- A two-person non-zero-sum non-cooperative games under incomplete information to model lane-changing decisions in the absence of such knowledge.

The target vehicle (i.e., that is changing lane) is assumed to have two pure strategies (change lane, wait), while the lag vehicle (i.e., the new follower after the lane-changing maneuver) has three pure strategies (accelerate, decelerate, and change lane). Table 8 and table 9 define the structure of the mandatory and discretionary lane-changing games.

Table 8. Discretionary lane-changing game with inactive vehicle-to-vehicle communication in normal form.

Action		Target Vehicle	
		A_1 (Change Lane)	A_2 (Do not Change Lane)
Lag Vehicle	B_1 (Accelerate)	(P_{11}, R_{11})	(P_{12}, R_{12})
	B_2 (Decelerate)	(P_{21}, R_{21})	(P_{22}, R_{22})
	B_3 (Change Lane)	(P_{31}, R_{31})	(P_{32}, R_{32})

Source: Talebpour, Mahmassani, and Bustamante, 2016.

Table 9. Mandatory lane-changing game with inactive vehicle-to-vehicle communication in normal form.

Action		Target Vehicle	
		A_1 (Change Lane)	A_2 (Do not Change Lane)
Lag Vehicle	B_1 (Accelerate)	(P_{11}, Q_{11})	(P_{12}, Q_{12})
	B_2 (Decelerate)	(P_{21}, Q_{21})	(P_{22}, Q_{22})
	B_3 (Change Lane)	(P_{31}, Q_{31})	(P_{32}, Q_{32})

Source: Talebpour, Mahmassani, and Bustamante, 2016.

Under uncertainty about the nature of the lane-changing maneuver, the approach developed by Harsanyi (1967) was utilized in order to transform a game of in incomplete information to a game of imperfect information. This method introduces "nature" as a player who characterizes the nature of the lane-changing maneuver with a specific probability. Figure 62 illustrates the extended form of the transformed game. Further information about the calibration and validation of these game structures could be found in Talebpour et al. (2016).

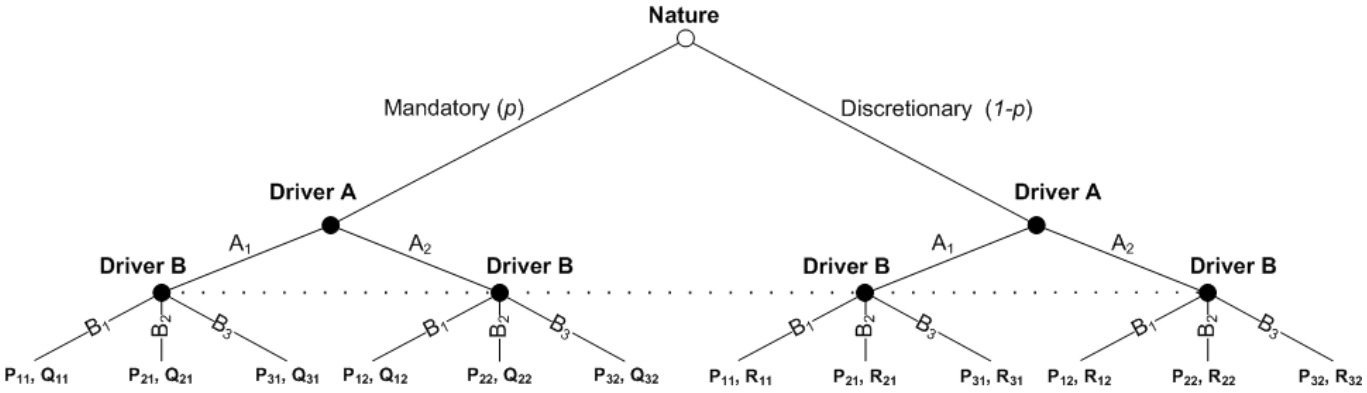


Figure 62. Chart. Extended form of the lane-changing game with inactive vehicle-to-vehicle communication.

CASE STUDY SCENARIOS AND AGENTS

In the following sections, three sets of scenarios were analyzed. The scenario sets serve as a small-scale experiment for implementing the proposed methodology. The proposed scenario sets are as follows:

- Demand scenarios
- Driver aggressiveness scenarios
- Mixed traffic scenarios

Besides generating the scenarios, an important step in the methodology is to generate the agents that are interacting in the simulated environment. As mentioned earlier, five different agents were considered in the study. The behavior of each agent should be determined by selecting appropriate microsimulation models. Here the models that should be selected for each agent are the car-following and lane-changing models. Once the behavioral models are selected for an agent, in order to generate an instance of the agent type, the associated model parameters are sampled from the joint probability distributions of the model parameters stored in the model libraries.

DEMAND SCENARIOS

Based on the layout of the road segment shown in figure 49, there are 15 paths that could be specified for the vehicles who enter the study area. Within the simulation tool, vehicles were assumed to enter the simulated environment following a Poisson process (O-D pair) that was calibrated based on the I-290 traffic flow data. In the current case study, in order to produce different demand levels, the interarrival time of the vehicles in each path was changed to produce different demand scenarios. A probability of occurrence was assigned to each demand scenario. In order to develop an accurate analysis of different demand scenarios under the proposed methodology, the selected highway segment should be monitored over a longer time period to acquire information about the traffic flow during peak hours of different days. Such observation would help to construct a better probability distribution for the vehicle entry and more accurate probabilities for the occurrence of the peak demand scenarios.

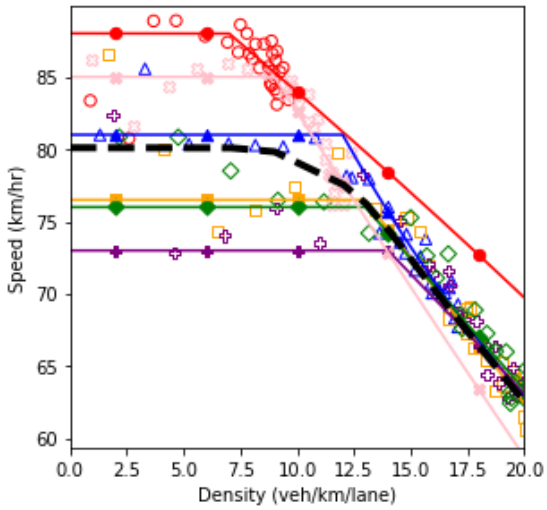
The following demand scenarios were examined in this section:

- Scenario 1 – The interarrival time increased by 20 percent for each path (with a probability of 0.1).
- Scenario 2 – Base case scenario (interarrival process calibrated based on the I-290 traffic flow data) (with a probability of 0.25).
- Scenario 3 – The interarrival time decreased by 20 percent for each path (with a probability of 0.2).
- Scenario 4 – The interarrival time decreased by 40 percent for each path (with a probability of 0.2).
- Scenario 5 – The interarrival time decreased by 60 percent for each path (with a probability of 0.15).
- Scenario 6 – The interarrival time decreased by 80 percent for each path (with a probability of 0.1).

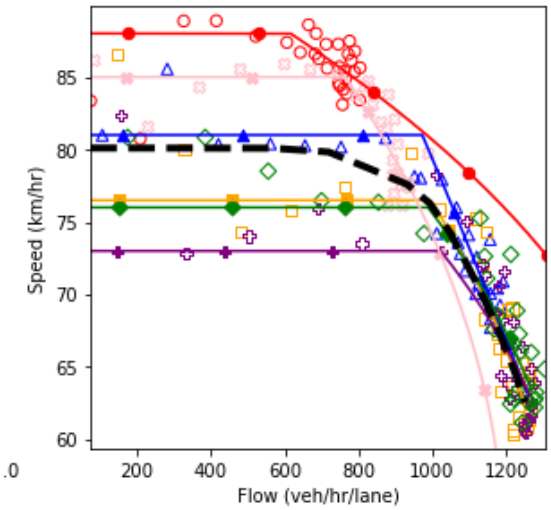
In this case, it was assumed that enough information is available to establish the probability distribution of the different scenarios. Therefore, the proposed scenario-based approach could be utilized to analyze the system and combine the performance metrics of the system under the specified scenarios.

Figure 63 shows the fundamental diagrams corresponding to the abovementioned scenarios. As expected, as the demand level increases, the system performance deteriorates; however, the decline in the system performance is not exactly proportional to the changes in the interarrival time. This disproportion occurred because of the complexity of the system as there are 15 different paths that are not fully independent of each other. Furthermore, some of the interarrival times generated by the model were overwritten based on the safe interarrival time between two consecutive vehicles. The dashed line on the plot was derived by combining the fundamental diagram of the scenarios using their respective probabilities.

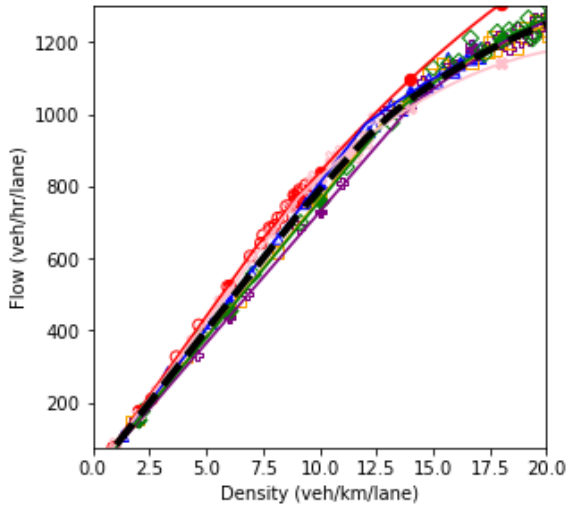
Figure 65 represents the travel time distributions on the mainline of the highway for each scenario. A decrease in the demand level results in an improvement in the travel time as seen by the shift in the distribution towards the lower travel time bins. From the fifth scenario, the improvement becomes insignificant because the vehicle should obey the safety headway instead of the headway generated according to the interarrival distribution. Moreover, as the demand on the highway segment increases, the system becomes more unstable by possessing a larger standard deviation in the travel time values. The overall performance of the system could be defined by combining the average travel time of the scenarios using the scenario probability of occurrence.



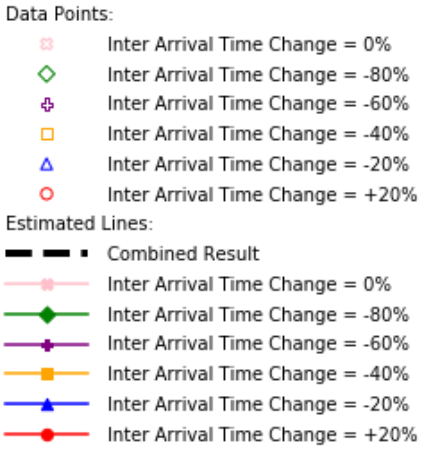
a) Speed and density.



b) Speed and flow.



c) Flow and density.

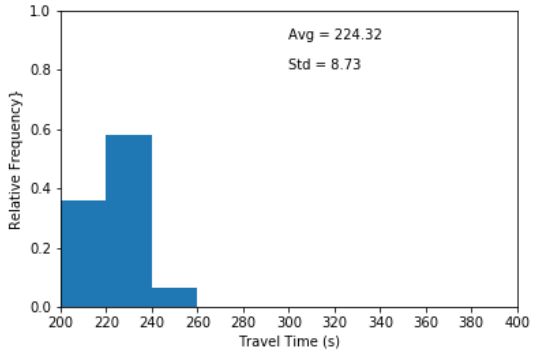


Source: FHWA, 2019.

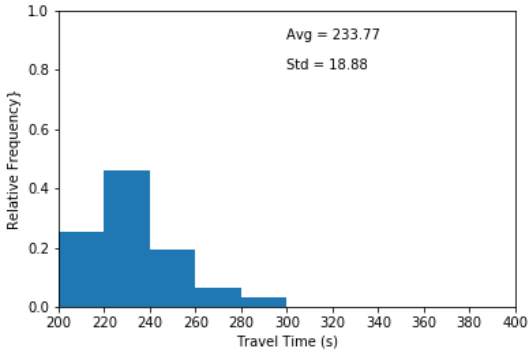
Figure 63. Diagrams. Compound figure depicts fundamental diagrams for different demand levels.

$$224.32 \times 0.1 + 233.77 \times 0.25 + 250.03 \times 0.2 + 273.77 \times 0.2 + 276.11 \times 0.15 + 275.43 \times 0.1 = 254.59$$

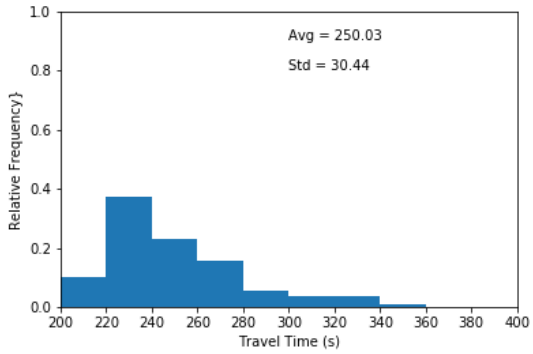
Figure 64. Equation. Weighted average travel time of the scenarios.



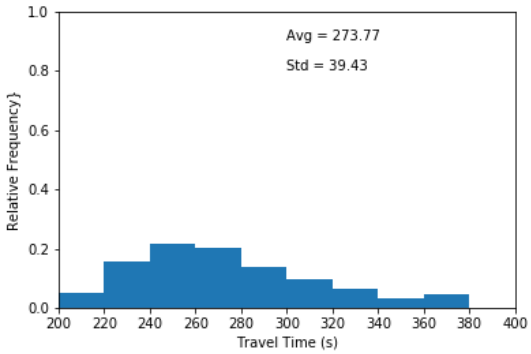
a) 20 percent increase in interarrival time. (scenario 1)



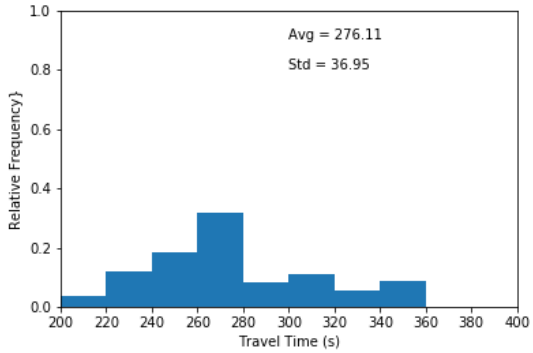
b) Base case. (scenario 2)



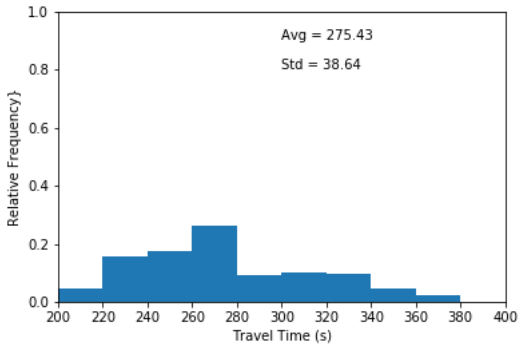
c) 20 percent decrease in interarrival time. (scenario 3)



d) 40 percent decrease in interarrival time. (scenario 4)



e) 60 percent decrease in interarrival time. (scenario 5)



f) 80 percent decrease in interarrival time. (scenario 6)

Source: FHWA, 2019.

Figure 65. Charts. Compound figure depicts travel time distribution for mainline vehicles under different interarrival time scenarios.

DRIVER AGGRESSIVENESS SCENARIOS

One of the major components in the proposed calibration methodology is the library of model parameters and the capability to generate simulation agents using the values stored in these libraries. Aggressiveness is one of the behavioral aspects of driving that has been extensively studied by modeling different types of drivers such as aggressive drivers, conservative drivers and neutral drivers (Papaioannou 2007, Herman, Malakhoff and Ardekani 1988, Tang, et al. 2012, Chen and Zhan 2008).

There are several ways to develop agents with different level of aggressiveness. The first approach is to initially determine the model parameters for a given dataset. Then, agents with different levels of aggressiveness could be generated by tweaking the model parameters. Another approach is to collect data from different groups of drivers knowing that each group, in general, behaves differently than the other groups in terms of the aggressiveness level. For the third approach, one might observe the vehicle trajectories and based on a predefined set of criteria distinguish between aggressive and conservative drivers. In this section the first and the third approaches were utilized.

As shown in section 2.1, there are various parameters that need to be estimated and calibrated when a regular vehicle is modeled. Two parameters that were tweaked in this study are the weighing factor for accidents (w_c) and the maximum anticipation time horizon (τ_{max}) which is used by a driver when estimating the probability of a rear-end collision. Basic statistics of the values of these parameters which were calibrated based on the NGSIM vehicle trajectory dataset are listed in table 10.

Table 10. Basic statistics of the weighing factor for accidents (w_c) and the maximum anticipation time horizon (τ_{max}) parameters.

Basic Statistics	τ_{max} (s)	w_c
Average	5.11	115657
Standard Deviation	2.43	94195

For the first approach, these parameters were changed in order to create new agents which would be categorized as conservative or aggressive. All the other parameters of the drivers were kept unchanged. This means that as a simplifying assumption it was assumed that the two selected parameters are independent of all other parameters in the model. In order to generate aggressive drivers, one second was subtracted from the maximum anticipation time horizon and the weighing factor for accidents was divided by 10. In contrast, to create conservative drivers, one second was added to the maximum anticipation time horizon and the weighing factor for accidents was multiplied by 10.

After characterizing the agents' behavioral model framework and their associated parameters, a set of feasible scenarios should be defined. Four different scenarios were considered in this example as follows:

- Scenario 1 – Fully conservative driving environment.
- Scenario 2 – Partially aggressive driving environment (67 percent conservative drivers and 33 percent aggressive drivers).
- Scenario 3 – Base case calibrated based on the NGSIM data.
- Scenario 4 – Fully aggressive driving environment.

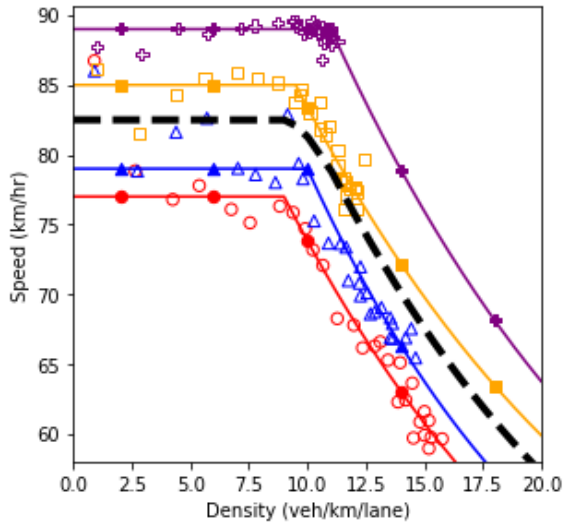
Figure 66 shows the fundamental diagrams under the four specified scenarios. The dashed line is a simple average of all the fundamental diagrams. This line is equivalent to the output of the scenario-based analysis if the scenarios occur with equal probability. It is also equivalent to the output of the robustness-based analysis if the Laplace's principle of insufficient reason metric is utilized.

The results show that as more conservative drivers are introduced in the system, the speed throughout the road segment decreases. The aggressive drivers can traverse the segment faster due to their higher speed and their ability to react more rapidly to perturbation in the system. Therefore, without any changes in the demand level, as more aggressive drivers enter the system, the range of values for density decreases. Furthermore, the less scatter in the data points for the fully aggressive driving condition indicates the stability of the system.

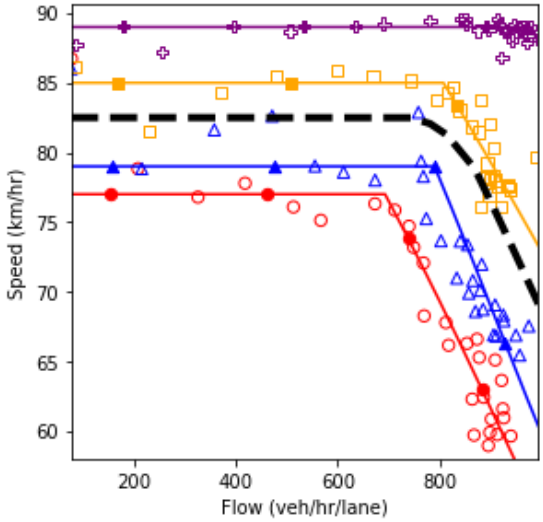
Figure 67 illustrates the travel time distribution on the mainline of the highway for each scenario. As could be seen in the figures, as the proportion of aggressive divers increases, the distribution shifts more toward lower travel time bins, which results in a decrease in the average and variance of the travel time on the mainline. The average travel time of all the scenarios is equal to 245.45 seconds.

An issue with the approach of generating artificial agents by changing the model parameters is that the results of analyzing the scenarios developed when these agents are in the system could be unreliable. To address this reliability concern, a different method of generating drivers should be selected.

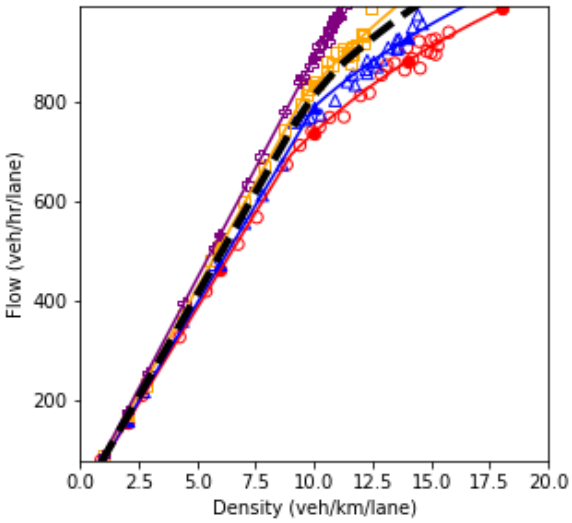
The other approach which was mentioned as a way of generating drivers with various level of aggressiveness is to separate the distribution of the NGSIM drivers into regular, conservative, and aggressive subgroups. In order to split the drivers, the weighing factor for accidents (w_c) and the maximum anticipation time horizon (τ_{max}) were sorted. The drivers who possessed the highest (lowest) 15 percent combined values of these two parameters were considered in the conservative (aggressive) subgroup. Such categorization of the drivers resulted in a 22 percent decrease in the average w_c and a 56 percent decrease in the average τ_{max} for aggressive drivers. On the other hand, for conservative drivers, the average w_c was increased by 11 percent and the average τ_{max} was increased by 72 percent.



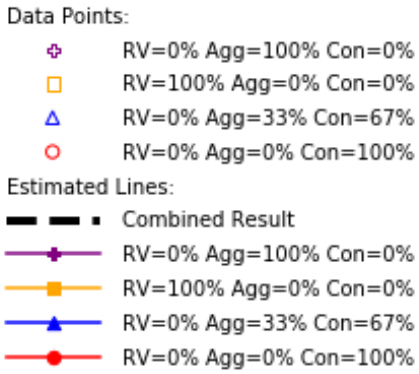
a) Speed and density.



b) Speed and flow.

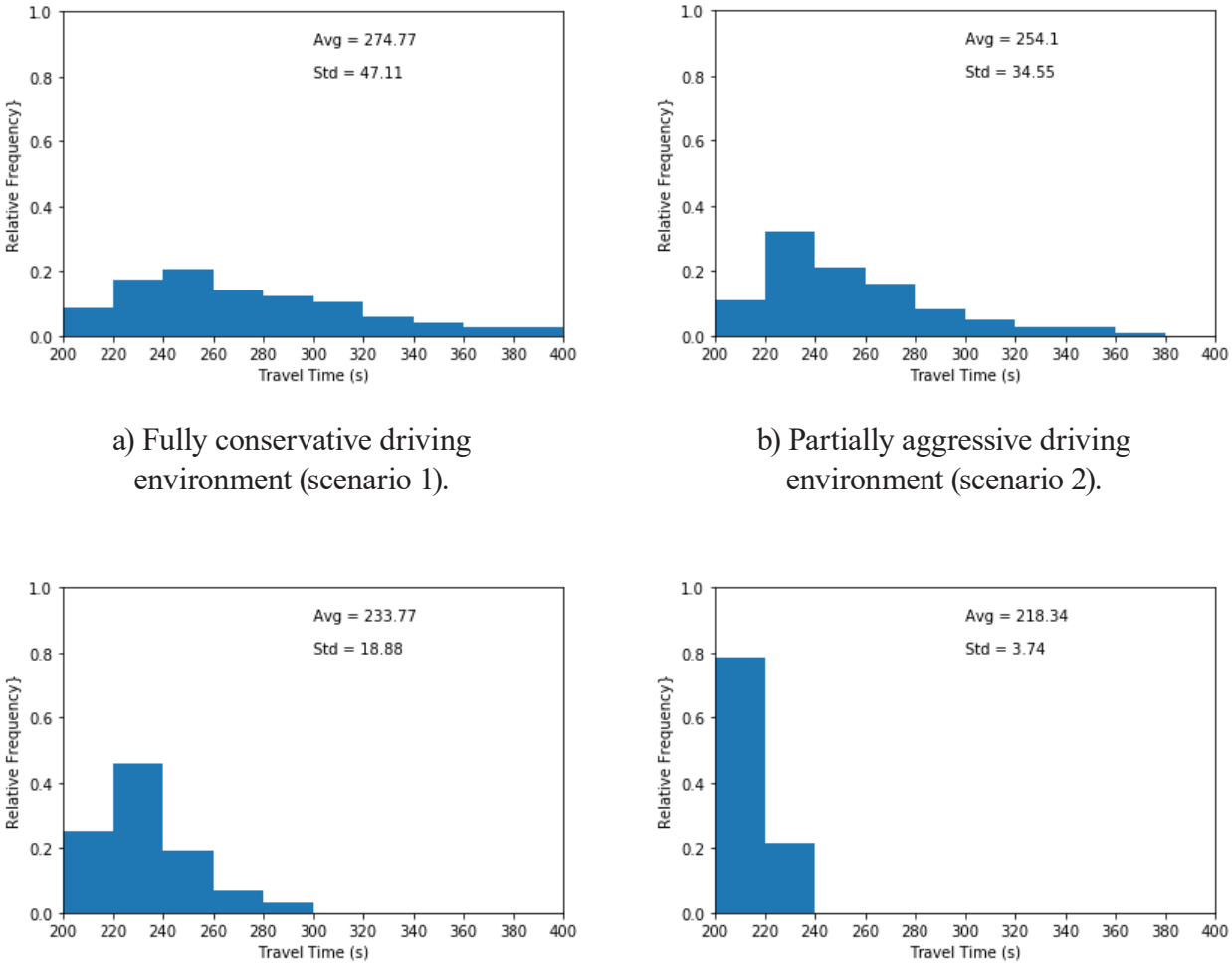


c) Flow and density.



Source: FHWA, 2019.

Figure 66. Charts. Compound figure depicts fundamental diagrams for different levels of aggressiveness in driving behavior.



a) Fully conservative driving environment (scenario 1).

b) Partially aggressive driving environment (scenario 2).

c) Base case (Scenario 3).

d) Fully aggressive driving environment (scenario 4).

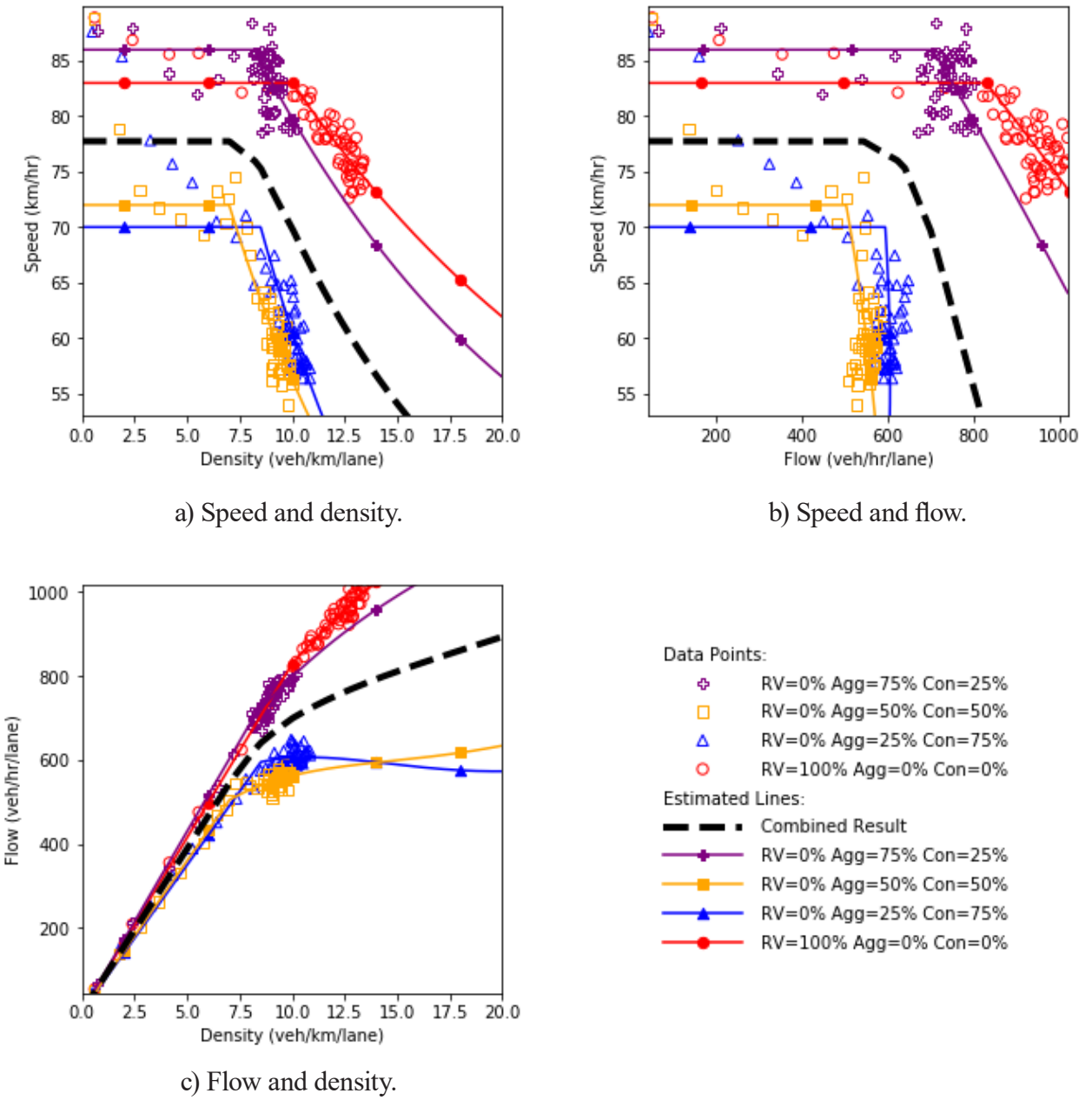
Source: FHWA, 2019.

Figure 67. Charts. Compound figure depicts travel time distributions for mainline vehicles under different aggressive driving scenarios.

After characterizing the agents' behavioral model framework, a set of feasible scenarios should be defined. Four different scenarios were considered in this example as follows:

- Scenario 1 – Absence of the conservative and aggressive drivers (NGSIM data with 30 percent of the agents removed).
- Scenario 2 – 25 percent aggressive drivers and 75 percent conservative drivers.
- Scenario 3 – 50 percent aggressive drivers and 50 percent conservative drivers.
- Scenario 4 – 75 percent aggressive drivers and 25 percent conservative drivers.

Figure 68 shows the fundamental diagrams under the four specified scenarios. The dashed line is a simple average of all the fundamental diagrams. This line is equivalent to the output of the scenario-based analysis if the scenarios occur with equal probability. It is also equivalent to the output of the robustness-based analysis if the Laplace's principle of insufficient reason metric is utilized. Based on the fundamental diagrams, as more conservative drivers are introduced in the system, the speed of the vehicles in the uncongested regime decreases. As could be seen, the fundamental diagram of scenario 3 is closer to the diagram of scenario 2 than the diagram of scenario 4. Therefore, it could be concluded that the conservative drivers impose a higher impact in the system performance than the aggressive drivers. This makes the accumulation of conservative drivers in the system to act as a dynamic bottleneck in the system that deteriorates the system performance. Comparing the diagrams of scenario 1 and scenario 2 shows that in the presence of a high proportion of aggressive drivers, the system performs better in the uncongested regime. However, the system enters the congested regime at a lower density value compared to the case where there is no extreme behavior (aggressiveness or conservativeness) in the system. The existence of conservative drivers in scenario 2 causes the system to have a lower performance in the congested regime compared to scenario 1. The less scatter in the data points for the scenario with high level of aggressive driving behavior indicates the higher stability in the system and the higher robustness against perturbations.



a) Speed and density.

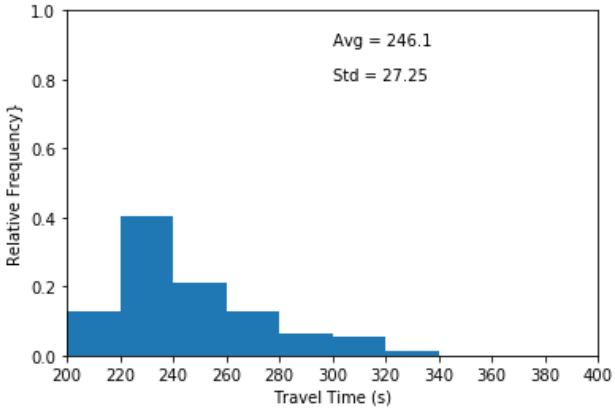
b) Speed and flow.

c) Flow and density.

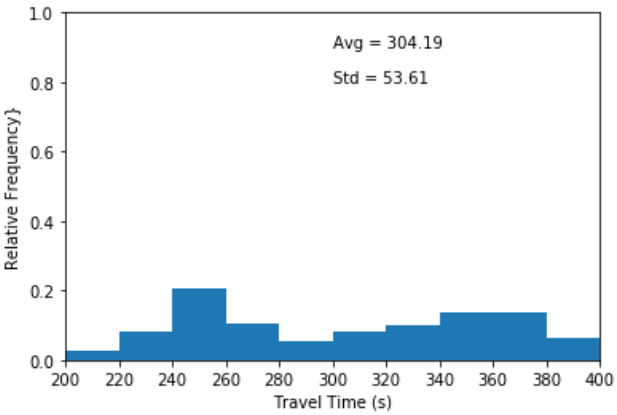
Source: FHWA, 2019.

Figure 68. Diagram. Fundamental diagrams for different levels of aggressiveness in driving behavior.

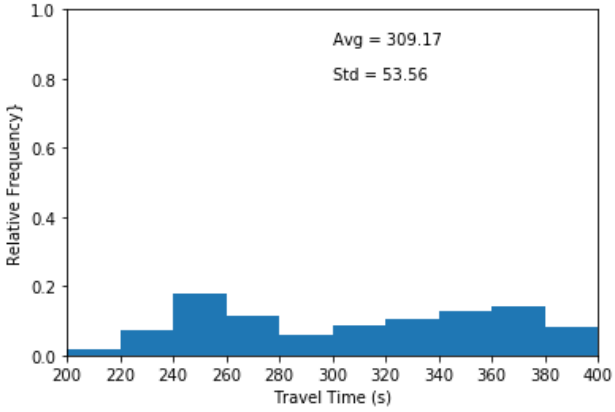
The mainline travel time histograms for each scenario are shown in figure 69. As could be seen in the figures, introducing conservative drivers in the system increases the average travel time and the travel time variance significantly. As the proportion of aggressive divers increases, the distribution shifts more toward lower travel time bins which results in a decrease in the average and variance of the travel time on the mainline. The effect of aggressive drivers on the travel time becomes more tangible once the proportion of aggressive drivers is higher than the proportion of conservative drivers. The average travel time of all the scenarios is equal to 272.73 seconds.



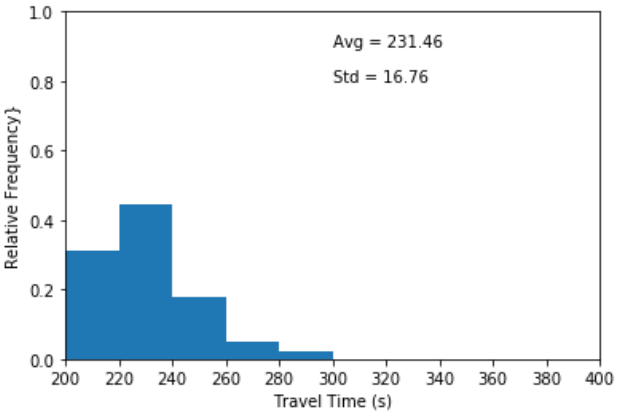
a) Aggressive = 0% and Conservative = 0% (Scenario 1).



b) Aggressive = 25% and Conservative = 75% (Scenario 2).



c) Aggressive = 50% and Conservative = 50% (Scenario 3).



d) Aggressive = 75% and Conservative = 25% (Scenario 4).

Source: FHWA, 2019.

Figure 69. Charts. Compound figure depicts travel time distribution for the mainline vehicles under various aggressive driver and conservative driver mix scenarios.

MIXED TRAFFIC SCENARIOS

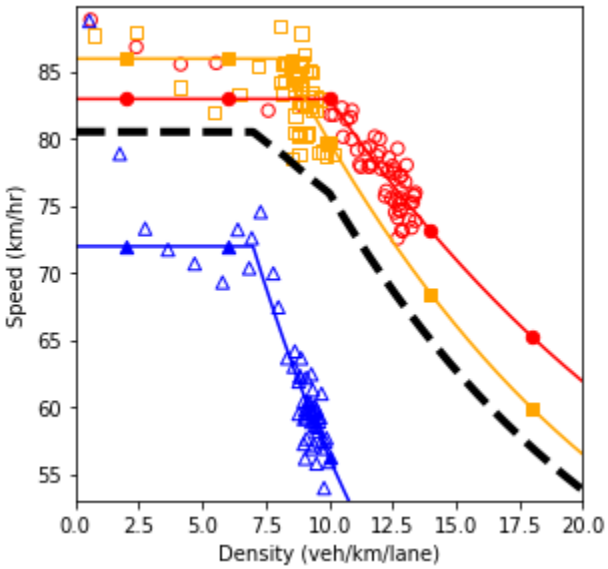
This section showcases the bi-level approach that combines the scenario-based and robustness-based analyses. Five different agents were introduced in the system, three of which were regular vehicles with various levels of aggressiveness. These agents enable the application of the scenario-based analysis assuming there is enough information to construct the probability of realizing each scenario. The connected and automated vehicles are the two agent types that impose deep uncertainty in the problem. Therefore, a robustness-based approach becomes useful. In total, 35 scenarios were considered as shown in table 11. The scenarios were generated by changing the market penetration rate of each agent type. As a simplifying assumption, it was assumed that the market penetration rates of connected vehicles and autonomous vehicles do not influence the regular vehicles' behavior such as their aggressiveness parameters. Similar to the previous section, it was assumed that 65 percent of the times the regular vehicles behave like the vehicles in the NGSIM data without the aggressive and conservative vehicles. In 25 percent and 10 percent of the occasions, the regular vehicles were considered as a 50/50 and 75/25 percentage proportions of aggressive and conservative, respectively. The scenario identifier in the table was assigned in a way that scenarios with a "-" character would be combined in the robust analysis. For example, scenarios 1-1, 1-2, and 1-3 would be combined using their assigned probabilities in order to generate scenario 1.

Figure 70 to figure 84 show the fundamental diagrams for the scenarios specified in table 10. Figure 74, figure 78, figure 78, figure 78, and figure 78 represent the scenarios where only connected and/or automated vehicles are in the system. In the remainder of the figures, the three different levels of aggressiveness were plotted. The dashed line shows the result of the scenario-based analysis by combining the three scenarios using their relevant probabilities (65 percent, 25 percent, and 10 percent). Up to this point, the analysis performed for each pair of market penetration rates for connected vehicles and automated vehicles is similar to the analysis performed in the "driver aggressiveness scenarios" section.

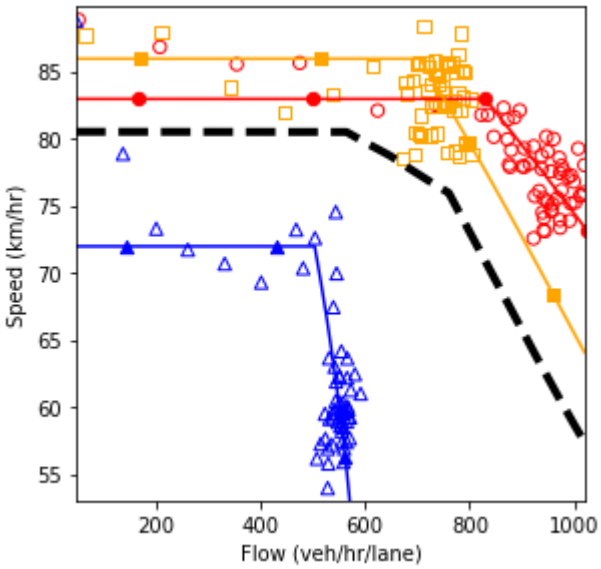
Based on the fundamental diagrams, in the absence of connectivity in the system, the scenario with regular vehicles possesses a higher speed in a congested environment. As connected vehicles start entering the highway segment, the scenario which has a high proportion of aggressive drivers dominates the other scenarios at different congestion levels. Throughout all plots in the figures, the scenarios with equal market penetration rates of aggressive and conservative drivers have the lowest performance. The only exception is when 75 percent of the traffic flow is composed of connected vehicles and there are no automated vehicles in the system. The results show that in this case, the connectivity feature can overcome the negative influence of slow conservative vehicles to some extent.

Table 11. Scenarios with different market penetration percentages by agent.

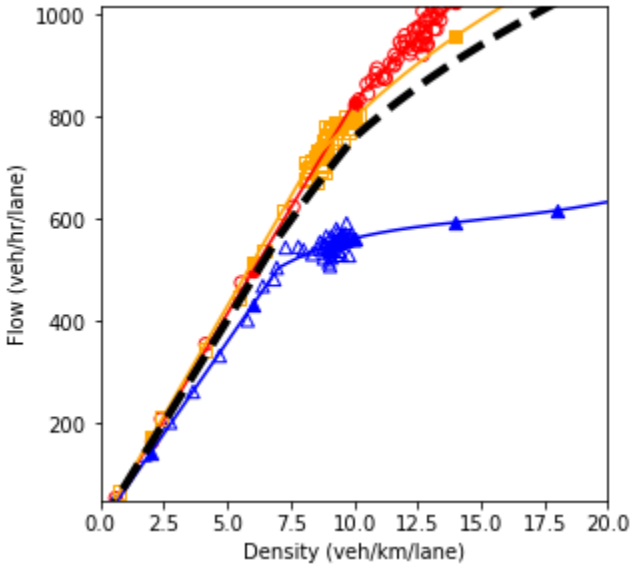
Scenario ID	Regular Vehicles	Aggressive Vehicles	Conservative Vehicles	Connected Vehicles	Automated Vehicles
1-1	100	0	0	0	0
1-2	0	50	50	0	0
1-3	0	75	25	0	0
2-1	75	0	0	0	25
2-2	0	37.5	37.5	0	25
2-3	0	56.25	18.75	0	25
3-1	50	0	0	0	50
3-2	0	25	25	0	50
3-3	0	37.5	12.5	0	50
4-1	25	0	0	0	75
4-2	0	12.5	12.5	0	75
4-3	0	18.75	6.25	0	75
5	0	0	0	0	100
6-1	75	0	0	25	0
6-2	0	37.5	37.5	25	0
6-3	0	56.25	18.75	25	0
7-1	50	0	0	25	25
7-2	0	25	25	25	25
7-3	0	37.5	12.5	25	25
8-1	25	0	0	25	50
8-2	0	12.5	12.5	25	50
8-3	0	18.75	6.25	25	50
9	0	0	0	25	75
10-1	50	0	0	50	0
10-2	0	25	25	50	0
10-3	0	37.5	12.5	50	0
11-1	25	0	0	50	25
11-2	0	12.5	12.5	50	25
11-3	0	18.75	6.25	50	25
12	0	0	0	50	50
13-1	25	0	0	75	0
13-2	0	12.5	12.5	75	0
13-3	0	18.75	6.25	75	0
14	0	0	0	75	25
15	0	0	0	100	0



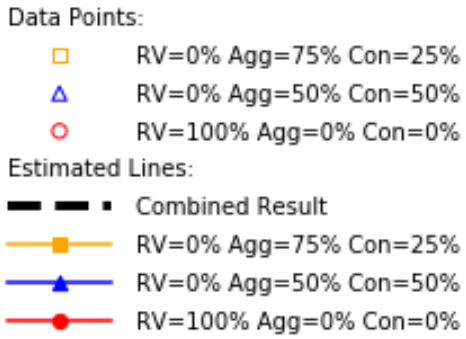
a) Speed and density.



b) Speed and flow.

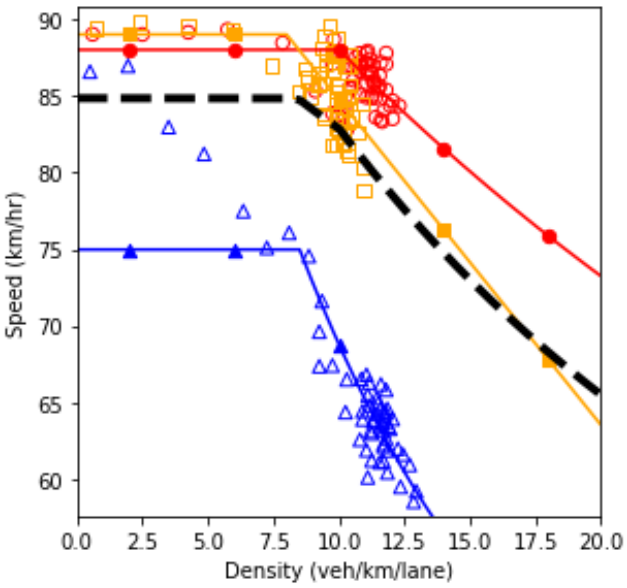


c) Flow and density.

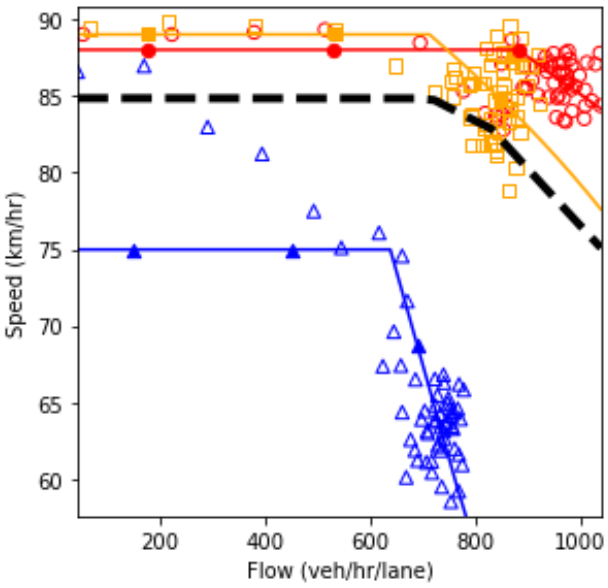


Source: FHWA, 2019.

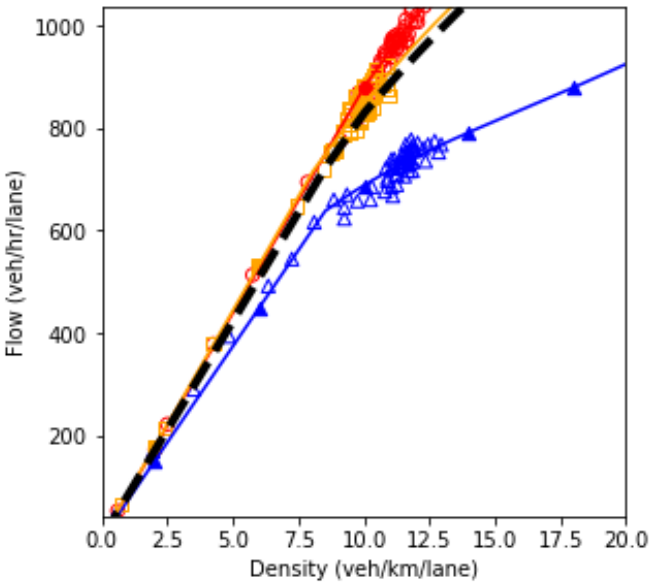
Figure 70. Charts. Compound figure depicts fundamental diagrams for a scenario in which the connected vehicle market penetration rate is 0 percent and the automated vehicle market penetration rate is 0 percent.



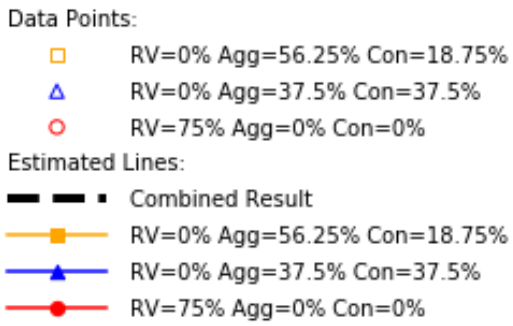
a) Speed and density.



b) Speed and flow.

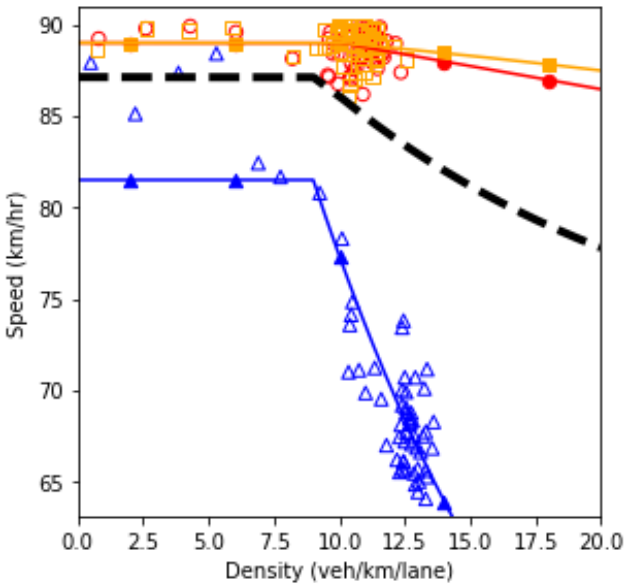


c) Flow and density.

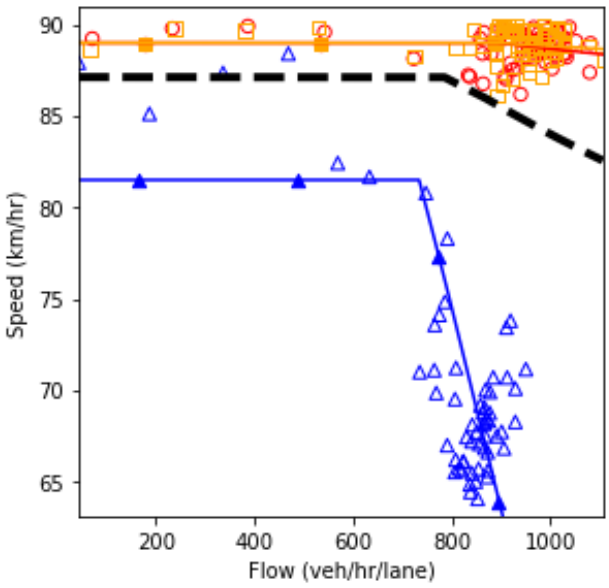


Source: FHWA, 2019.

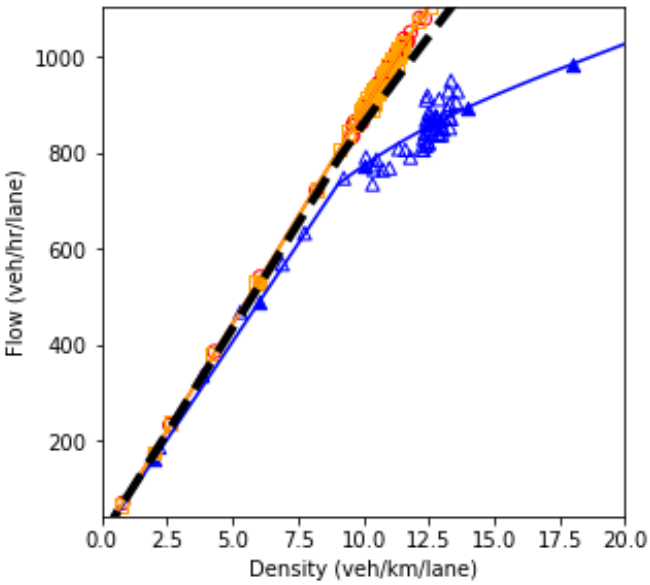
Figure 71. Charts. Compound figure depicts fundamental diagrams for a scenario in which the connected vehicle market penetration rate is 0 percent and the automated vehicle market penetration rate is 25 percent.



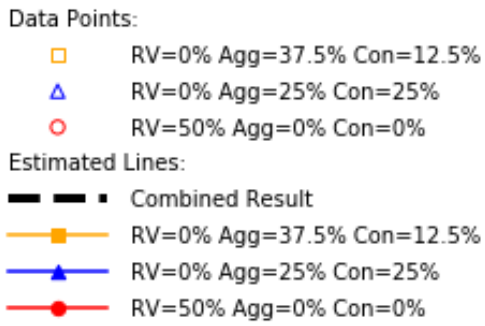
a) Speed and density.



b) Speed and flow.

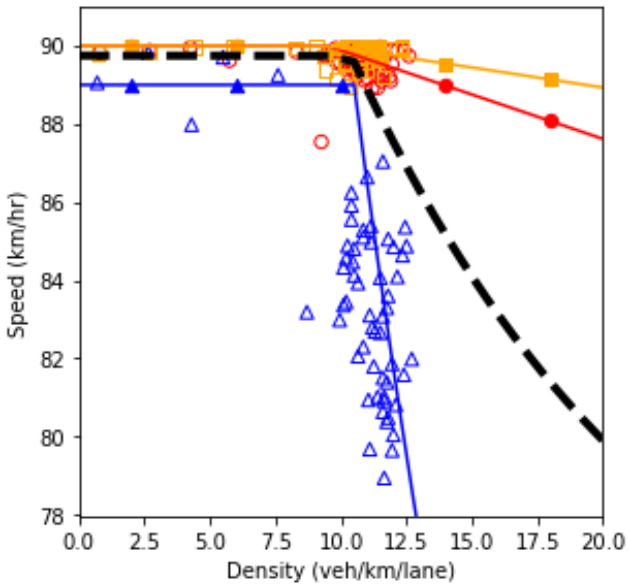


c) Flow and density.

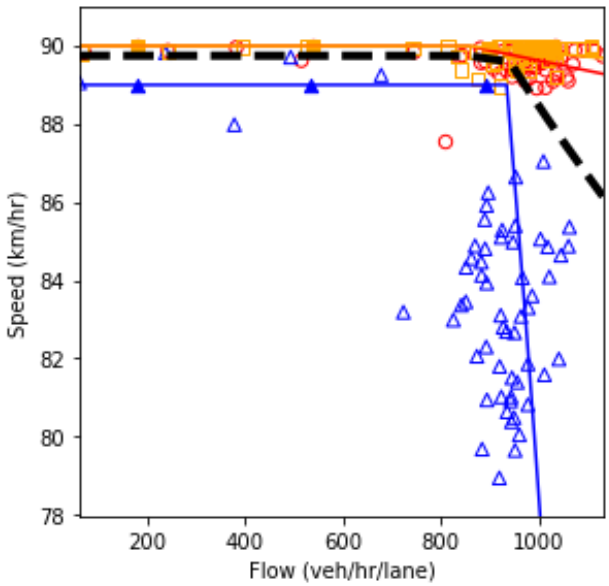


Source: FHWA, 2019.

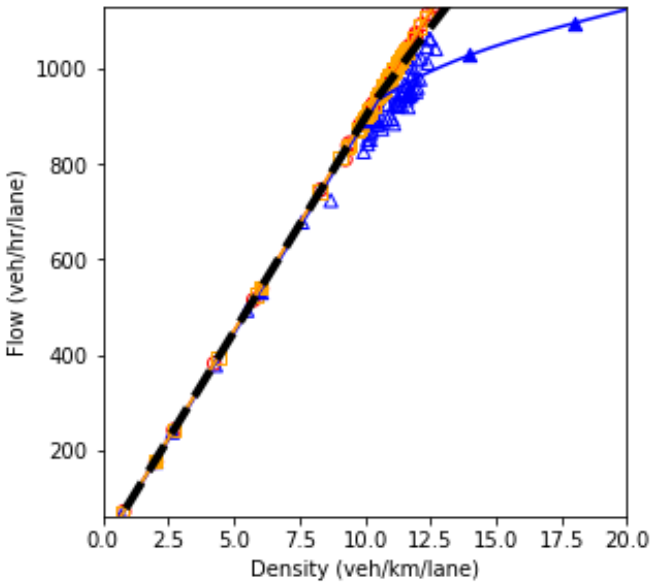
Figure 72. Charts. Compound figure depicts fundamental diagrams for a scenario in which the connected vehicle market penetration rate is 0 percent and the automated vehicle market penetration rate is 50 percent.



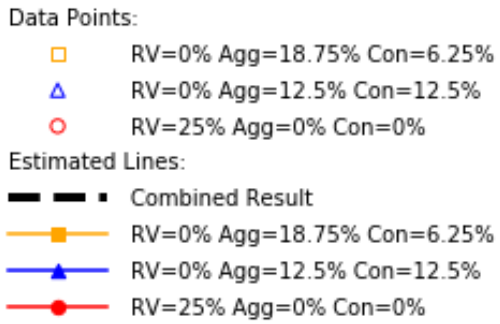
a) Speed and density.



b) Speed and flow.

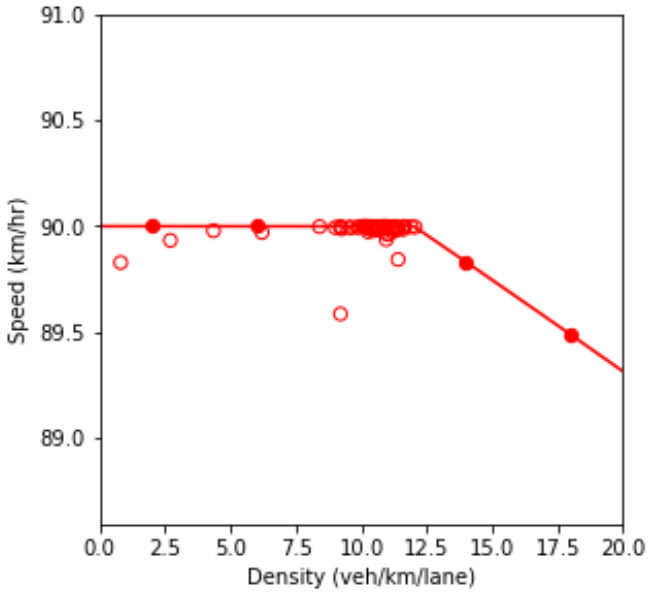


c) Flow and density.

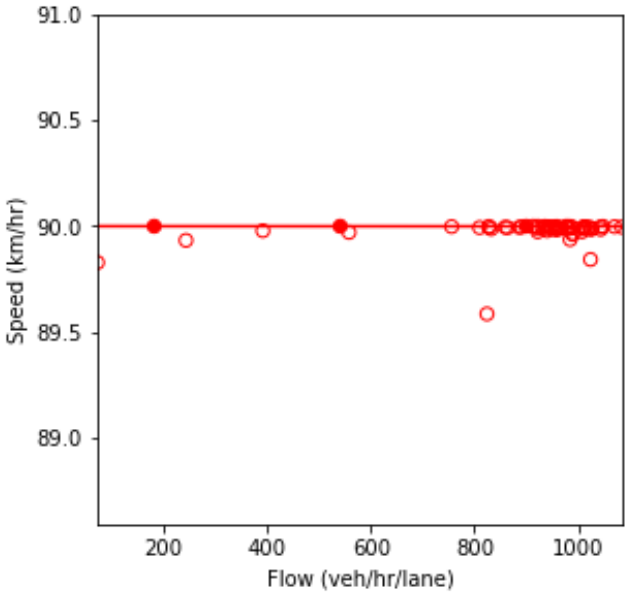


Source: FHWA, 2019.

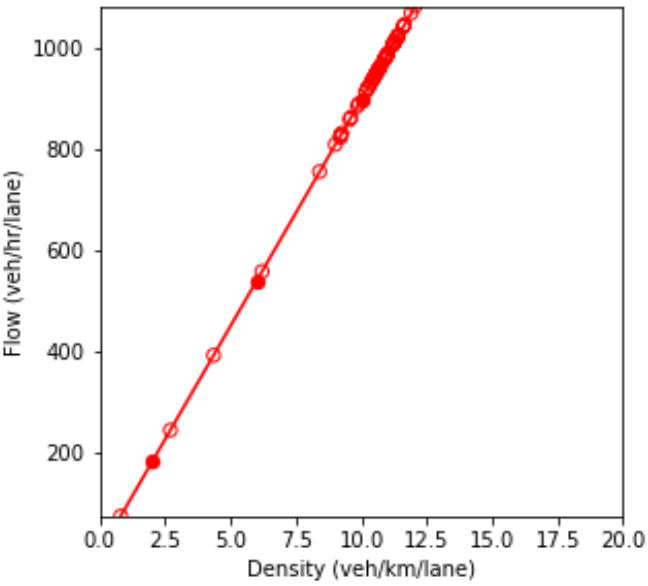
Figure 73. Charts. Compound figure depicts fundamental diagrams for a scenario in which the connected vehicle market penetration rate is 0 percent and the automated vehicle market penetration rate is 75 percent.



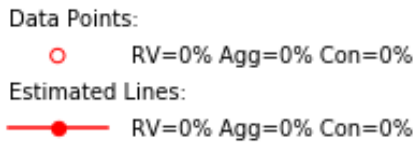
a) Speed and density.



b) Speed and flow.

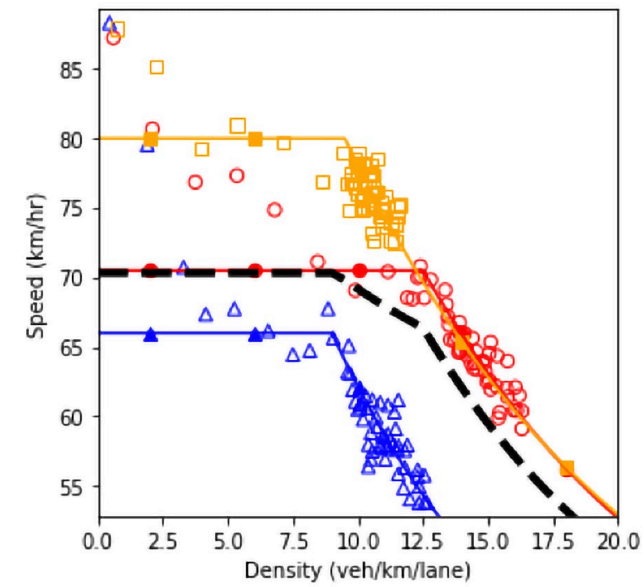


c) Flow and density.

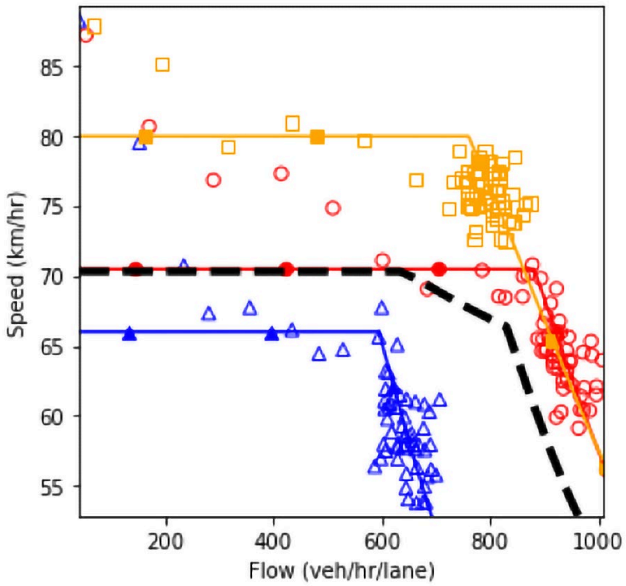


Source: FHWA, 2019.

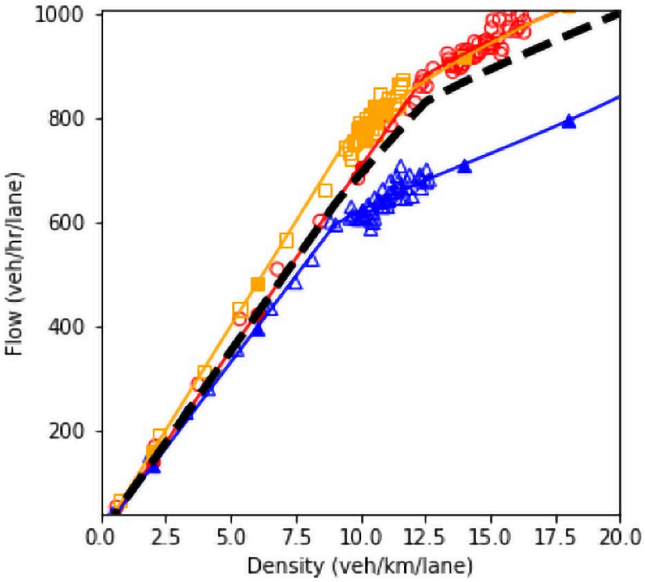
Figure 74. Charts. Compound figure depicts fundamental diagrams for a scenario in which the connected vehicle market penetration rate is 0 percent and the automated vehicle market penetration rate is 100 percent.



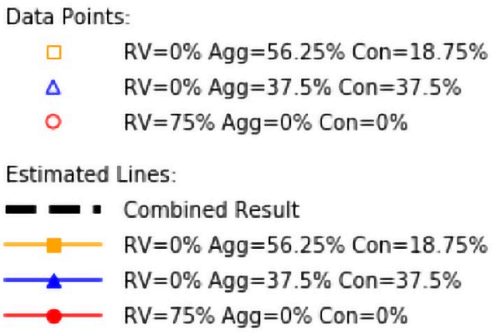
a) Speed and density.



b) Speed and flow.

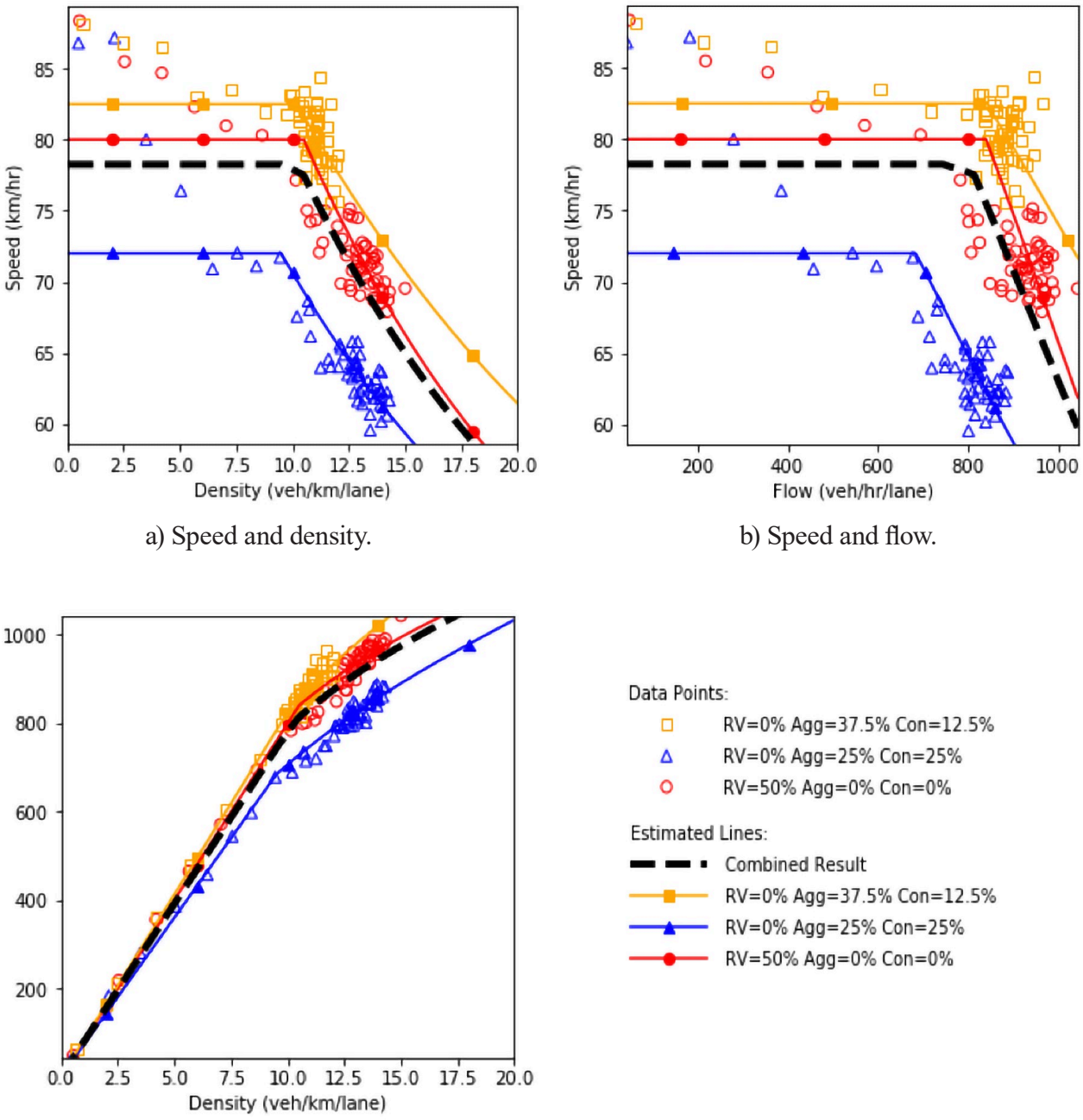


c) Flow and density.



Source: FHWA, 2019.

Figure 75. Charts. Compound figure depicts fundamental diagrams for a scenario in which the connected vehicle market penetration rate is 25 percent and the automated vehicle market penetration rate is 0 percent.



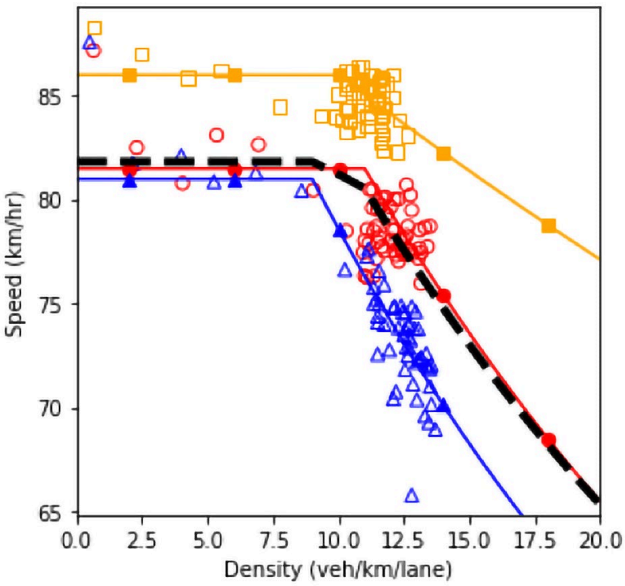
a) Speed and density.

b) Speed and flow.

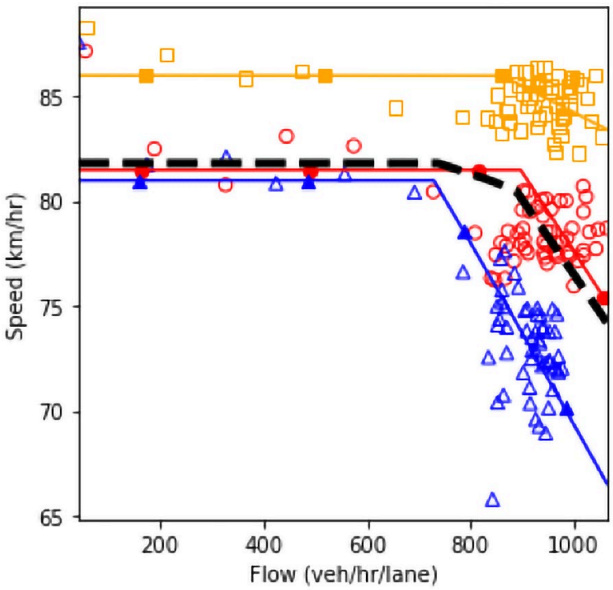
c) Flow and density.

Source: FHWA, 2019.

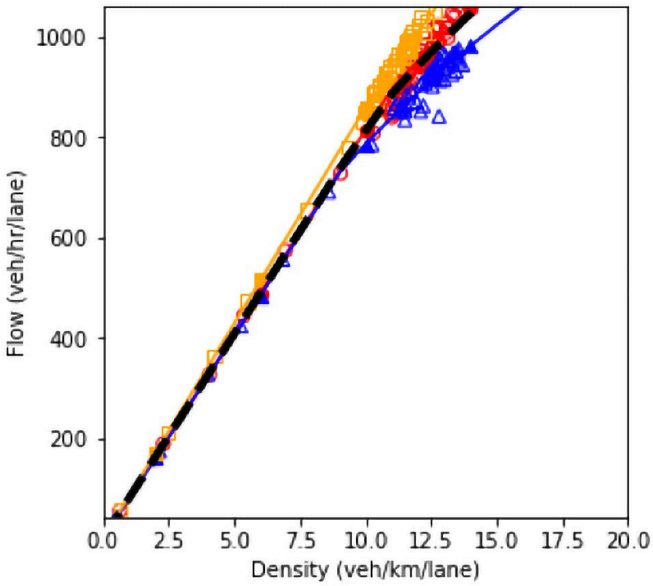
Figure 76. Charts. Compound figure depicts fundamental diagrams for a scenario in which the connected vehicle market penetration rate is 25 percent and the automated vehicle market penetration rate is 25 percent.



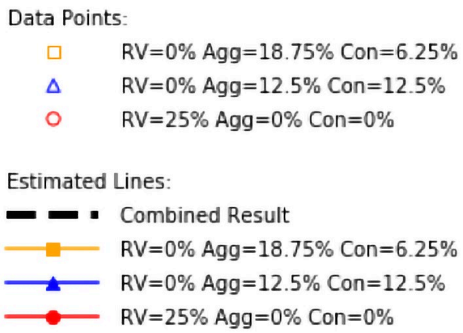
a) Speed and density.



b) Speed and flow.

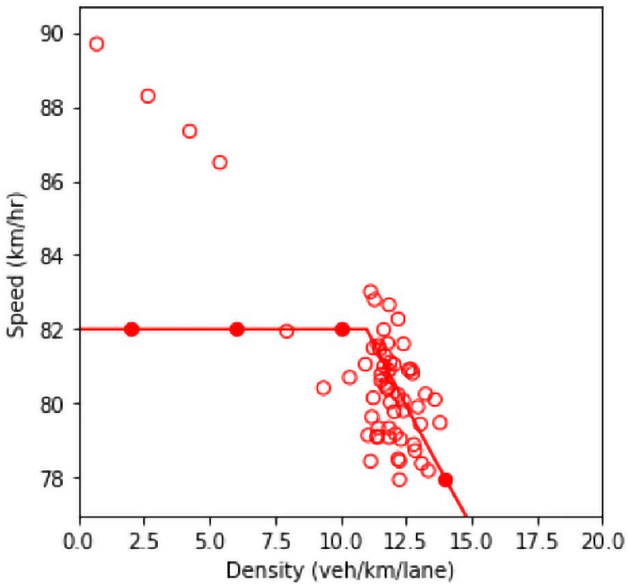


c) Flow and density.

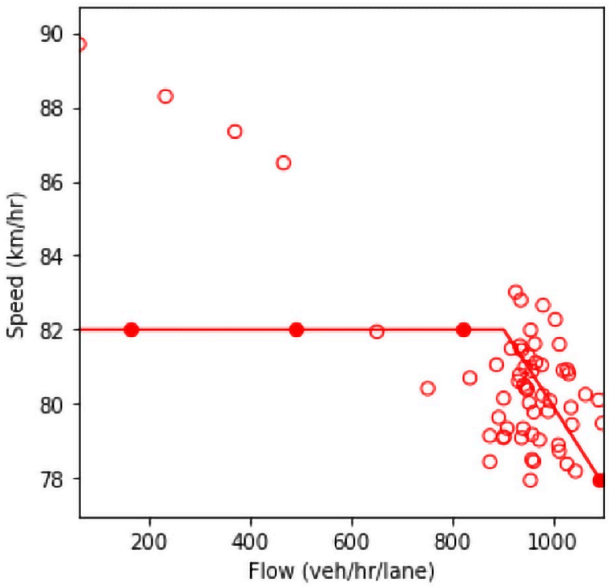


Source: FHWA, 2019.

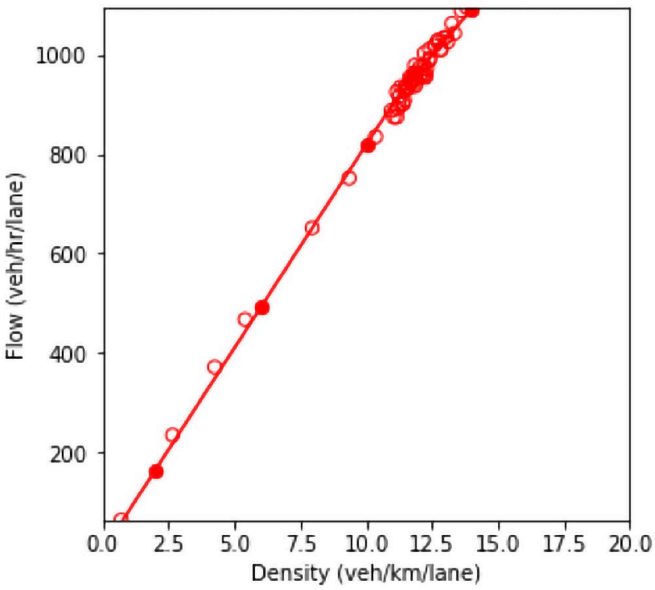
Figure 77. Charts. Compound figure depicts fundamental diagrams for a scenario in which the connected vehicle market penetration rate is 25 percent and the automated vehicle market penetration rate is 50 percent.



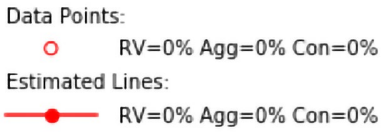
a) Speed and density.



b) Speed and flow.

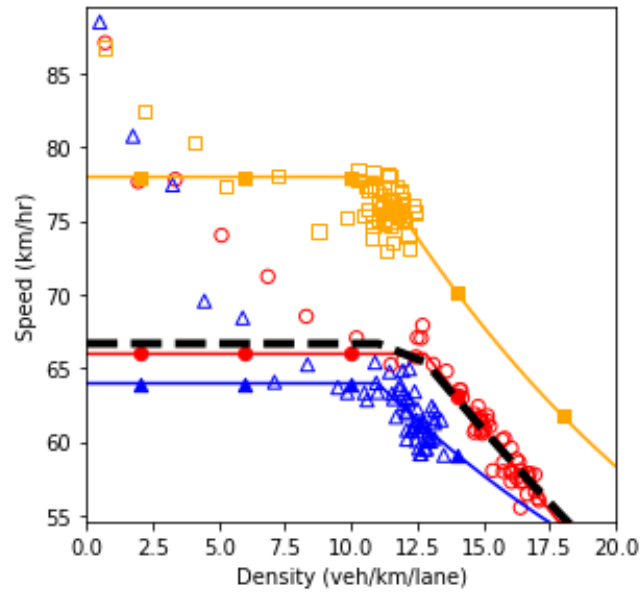


c) Flow and density.

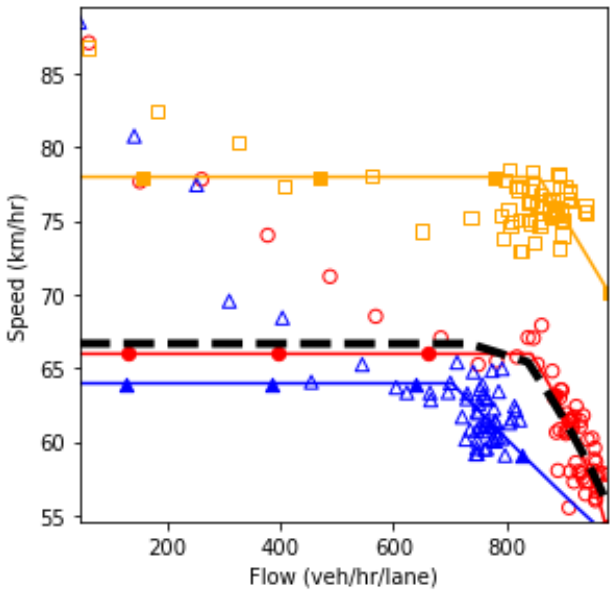


Source: FHWA, 2019.

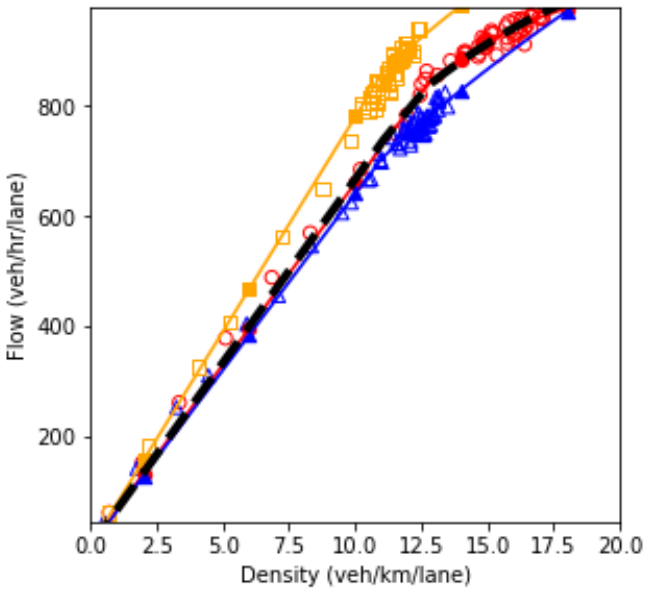
Figure 78. Charts. Compound figure depicts fundamental diagrams for a scenario in which the connected vehicle market penetration rate is 25 percent and the automated vehicle market penetration rate is 75 percent.



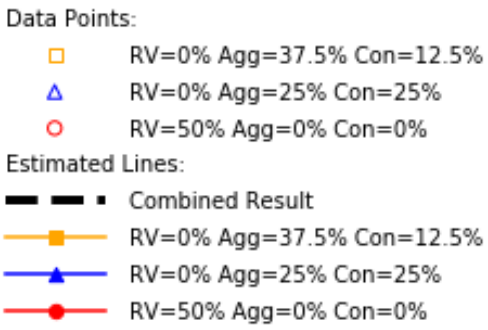
a) Speed and density.



b) Speed and flow.

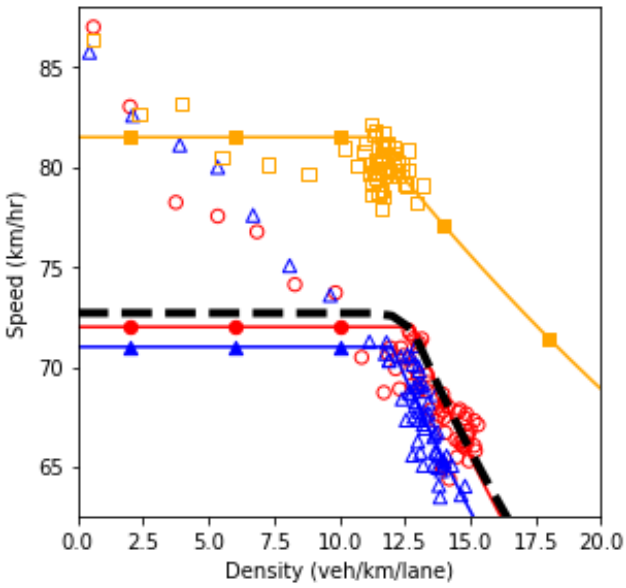


c) Flow and density.

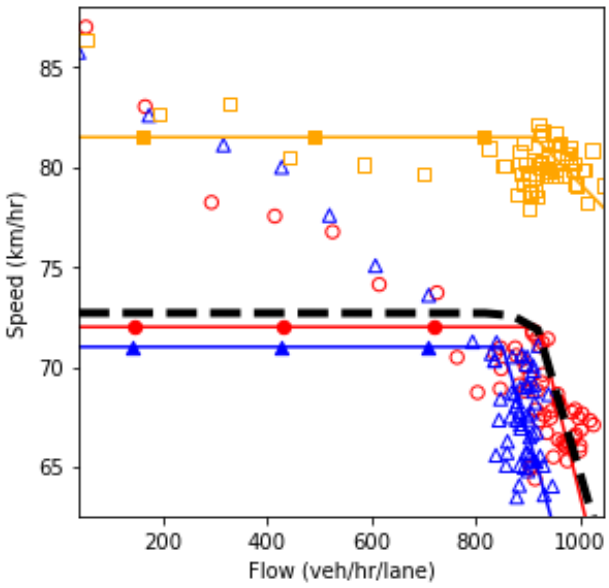


Source: FHWA, 2019.

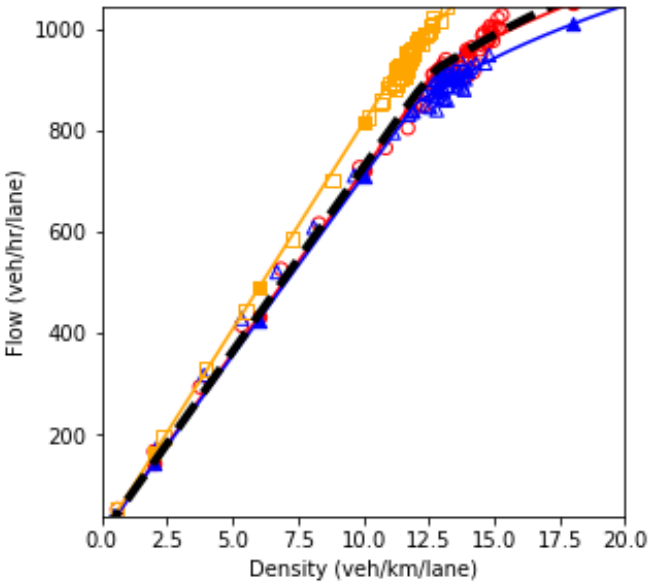
Figure 79. Charts. Compound figure depicts fundamental diagrams for a scenario in which the connected vehicle market penetration rate is 50 percent and the automated vehicle market penetration rate is 0 percent.



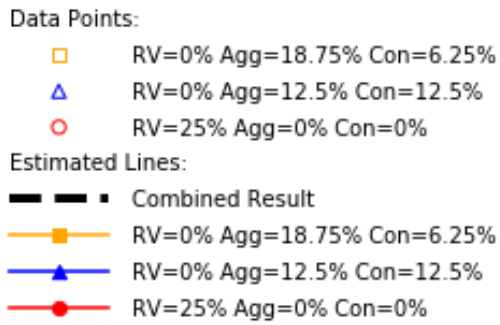
a) Speed and density.



b) Speed and flow.

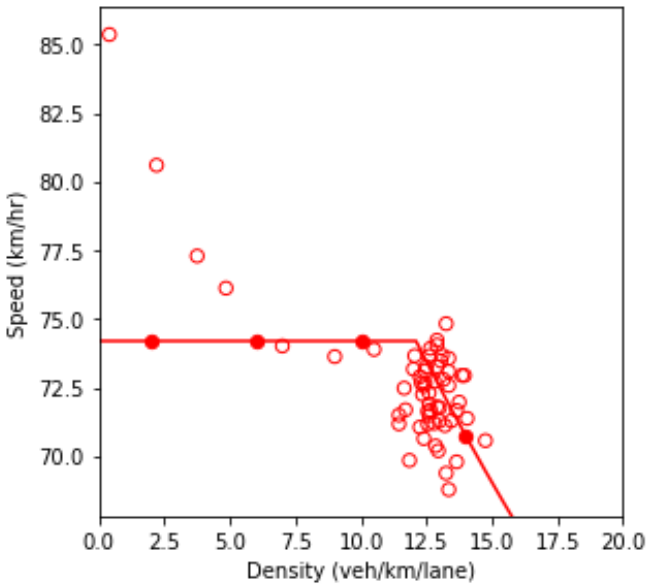


c) Flow and density.

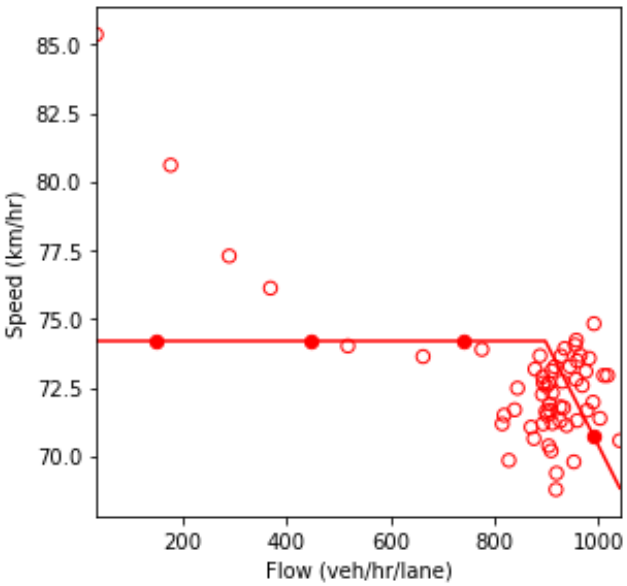


Source: FHWA, 2019.

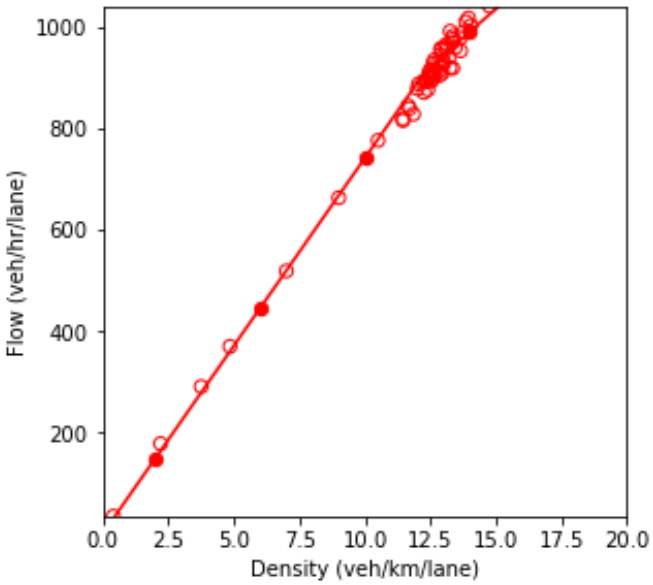
Figure 80. Charts. Compound figure depicts fundamental diagrams for a scenario in which the connected vehicle market penetration rate is 50 percent and the automated vehicle market penetration rate is 25 percent.



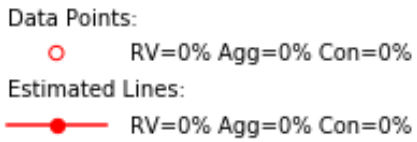
a) Speed and density.



b) Speed and flow.

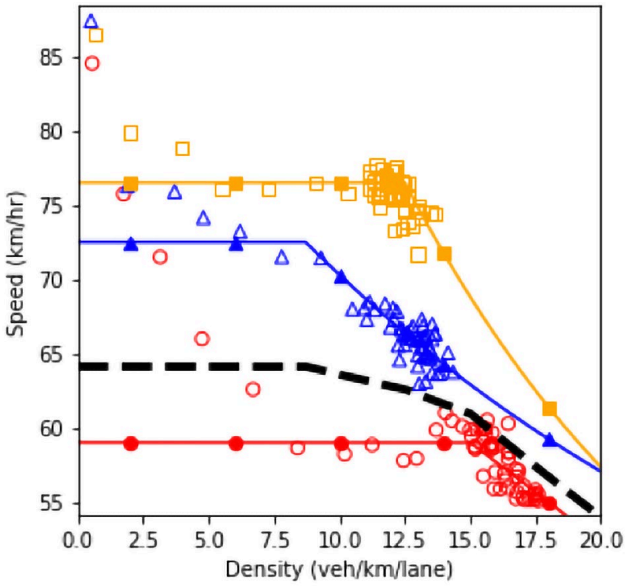


c) Flow and density.

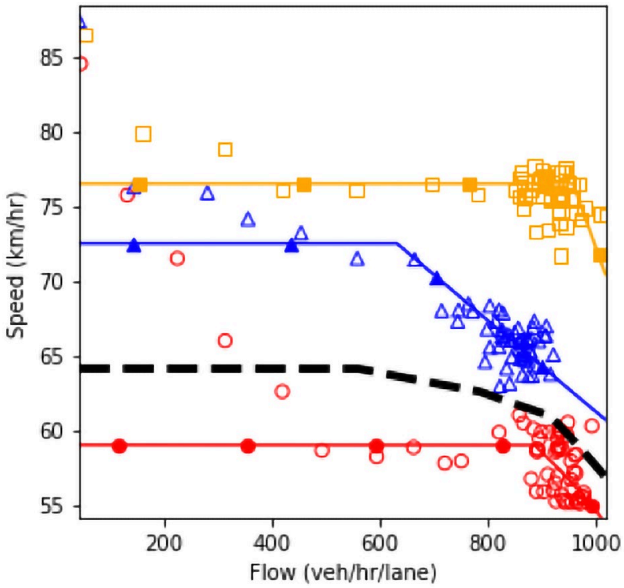


Source: FHWA, 2019.

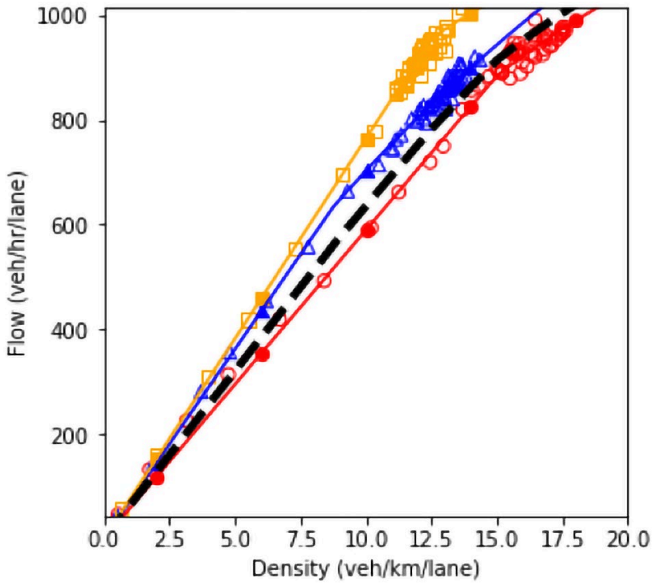
Figure 81. Charts. Compound figure depicts fundamental diagrams for a scenario in which the connected vehicle market penetration rate is 50 percent and the automated vehicle market penetration rate is 50 percent.



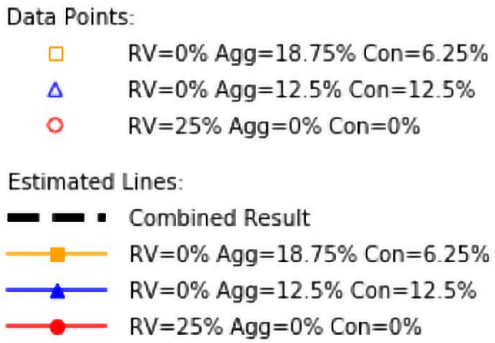
a) Speed and density.



b) Speed and flow.

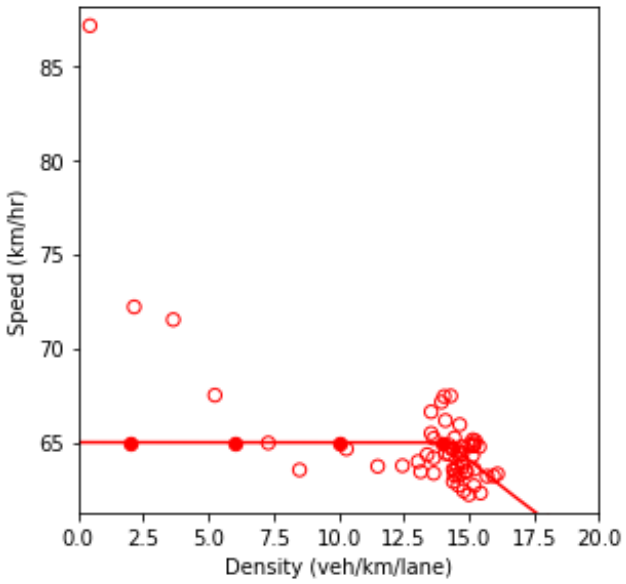


c) Flow and density.

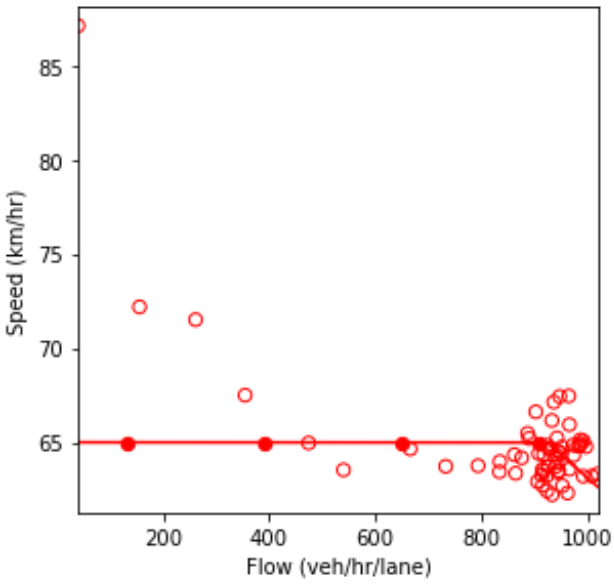


Source: FHWA, 2019.

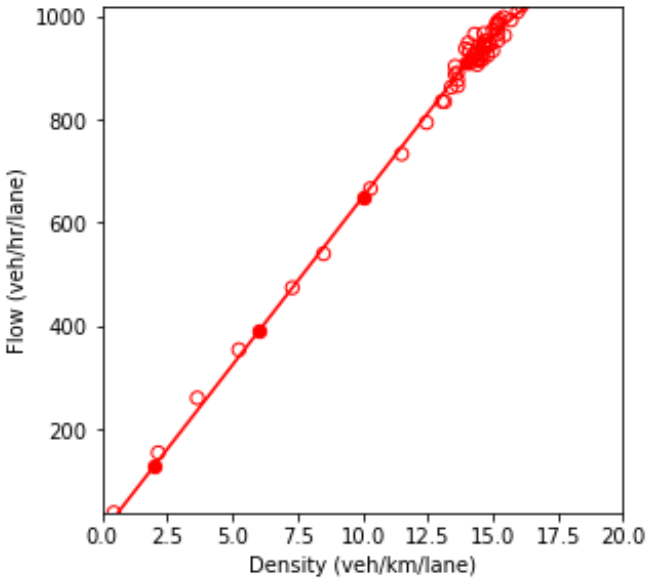
Figure 82. Charts. Compound figure depicts fundamental diagrams for a scenario in which the connected vehicle market penetration rate is 75 percent and the automated vehicle market penetration rate is 0 percent.



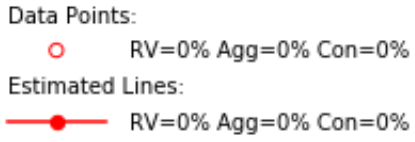
a) Speed and density.



b) Speed and flow.

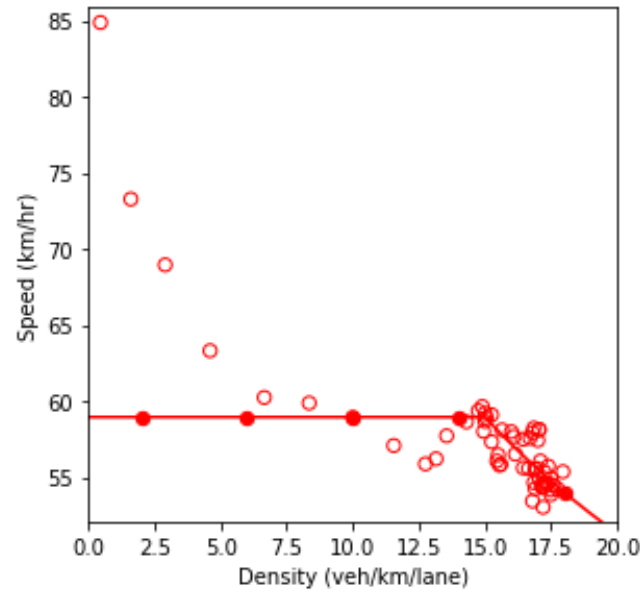


c) Flow and density.

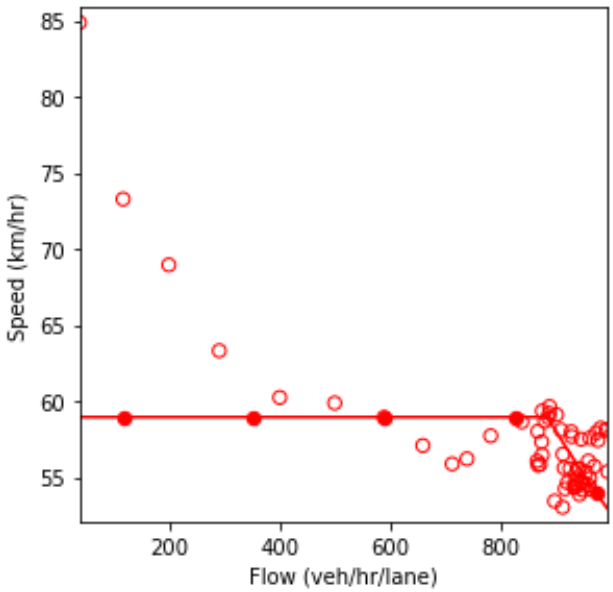


Source: FHWA, 2019.

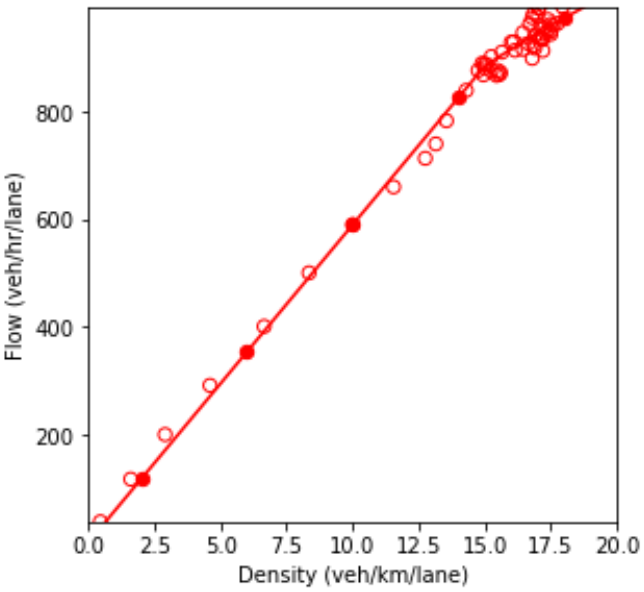
Figure 83. Charts. Compound figure depicts fundamental diagrams for a scenario in which the connected vehicle market penetration rate is 75 percent and the automated vehicle market penetration rate is 25 percent.



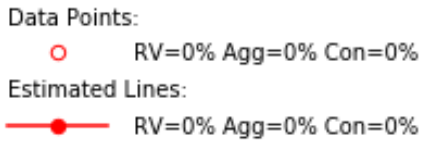
a) Speed and density.



b) Speed and flow.



c) Flow and density.



Source: FHWA, 2019.

Figure 84. Charts. Compound figure depicts fundamental diagrams for a scenario in which the connected vehicle market penetration rate is 100 percent and the automated vehicle market penetration rate is 0 percent.

The second step in the bi-level approach is the robustness-analysis. In order to perform a robust analysis, the outputs of the scenario-based analysis should be plotted together. Therefore, all the dashed lines in the previous figures that possessed more than one scenario were plotted together in figure 86. The figure shows the fundamental diagram for various market penetration rates of connected vehicles and autonomous vehicles.

In an environment with only regular and automated vehicles, as the proportion of automated vehicles increases, the system performance enhances. However, this improvement occurs with a diminishing rate; resulting in the scenario with 75 percent automated vehicle proportion to possess a performance that is similar to the performance of a fully automated environment. The solid red line pertains to the scenario where all vehicles are manually driven without any connectivity feature. It could be seen that compared to the base case (red solid line), introducing connected vehicles in the system causes decay in the system performance. The main reason for this impact of connected vehicles is the different car-following model defined for this type of vehicle. Such a negative impact on system performance could be addressed by adding automated vehicles to the system. Although the inclusion of connected vehicles caused a decrease in the system performance when the system is in the free-flow regime, their presence becomes more beneficial as the system reaches the congested state. For a fixed level of automation in the system, as the market penetration rate of connected vehicles increases, the curvature of the speed-density and speed-flow diagram decrease. As a result, the robustness of the system against congestion increases and the breakdown density occurs at higher density values.

As could be seen in the figure, at an automation level of 25 percent, an increase in connected vehicles market penetration rate from 25 percent to 50 percent has a negligible effect on the system performance once the density on the mainline exceeds 12.5 vehicles per kilometer per lane (veh/km/lane). Similarly, beyond a density of 11 veh/km/lane, the system performance difference between scenario 2 and scenario 8 is negligible.

After a general comparison of the graphs, in order to conduct the robustness-based analysis, the robustness measure should be selected. A variety of measures were introduced in the previous chapter including maximin, maximax, Laplace's principle of insufficient reason, Hurwicz optimism pessimism rule, minimax regret, undesirable deviations, skewness, and peakedness. Some of these measures would be used in the current example. Based on the fundamental diagrams of figure 86, the scenario with full connectivity and the scenario with full automation are the scenarios selected using the maximin and maximax robustness measures, respectively. For the Laplace's principle of insufficient reason measure, a simple average of all the 15 diagrams could be plotted. This would be similar to performing the scenario-based analysis with equal probabilities for the scenarios. Based on the definition of the Hurwicz optimism pessimism rule, the representative fundamental diagram is a weighted average of the full connectivity scenario and the full automation scenario. The weight is determined using the coefficient of pessimism.

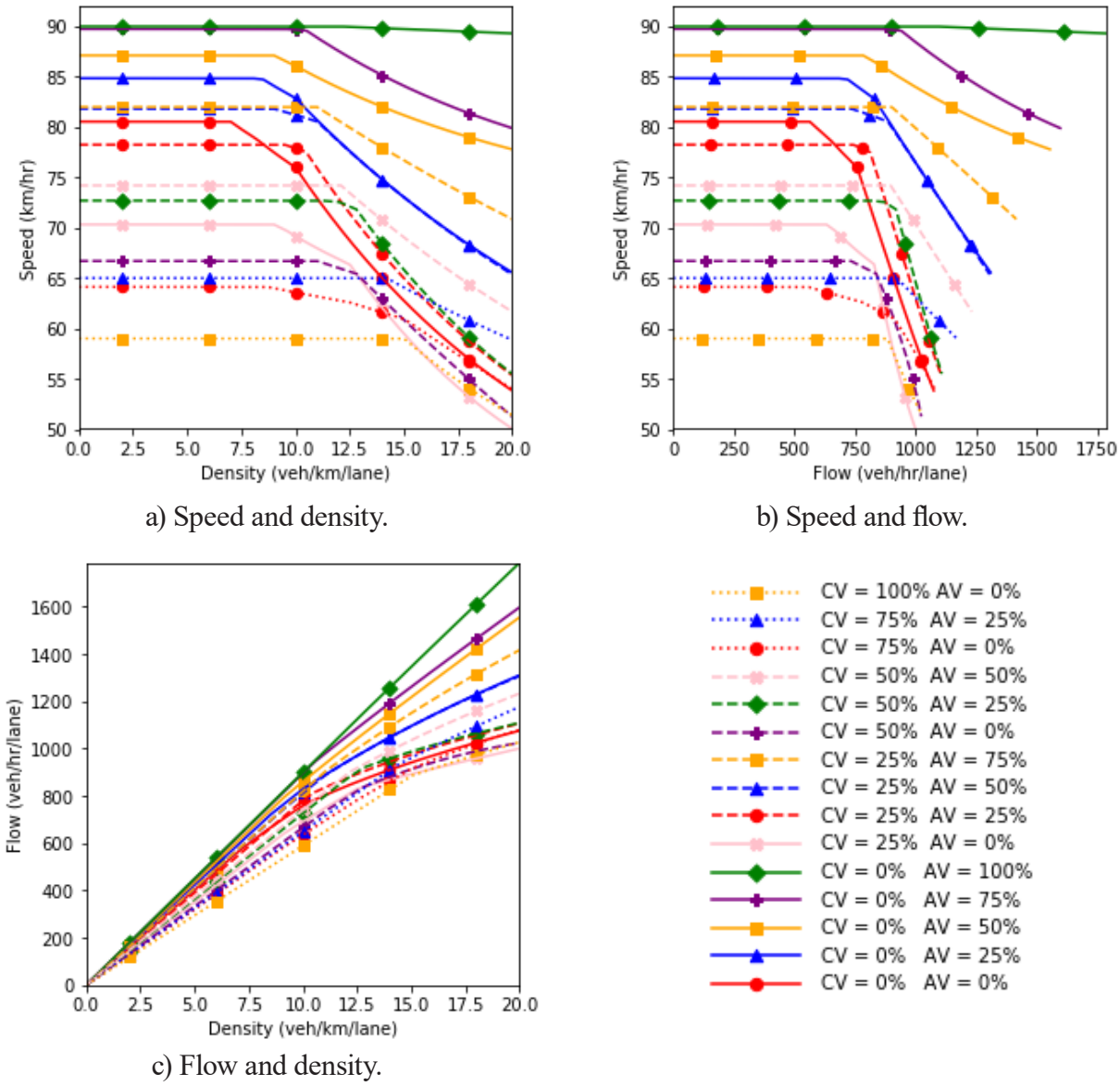
Travel time is another performance measure that could be used to evaluate the system under different scenarios by performing the bi-level analysis. Comparing the distributions shows that in the absence of connected vehicles, the distribution is skewed toward lower values. As the proportion of connected vehicles increases, the mode of the distribution moves toward higher values of travel time and the distribution flattens resulting in an increase in the variance of travel time.

The scenario-based analysis was performed on the average travel time of the scenarios. The results are depicted in figure 87. Then, based on a selected robustness measure, the second step of the bi-level approach could be performed on the average travel time. For instance, the scenario with a 100 percent market penetration rate of connected vehicles is the worst-case scenario corresponding to the maximin robustness metric. Similarly, the scenario where all vehicles are automated possesses the lowest average travel time. Therefore, it represents the best-case scenario which relates to the maximax robustness metric. The average travel time for the Laplace's principle of insufficient reason measure, as the average of all the values in the figure, is equal to 266.2 seconds. The average travel time under the Hurwicz optimism pessimism rule could be determined by the following formula:

$$330.9\alpha + 215.97(1 - \alpha)$$

Figure 85. Equation. Travel time based on the Hurwicz optimism pessimism rule.

Where is the coefficient of pessimism varying between 0 and 1.



a) Speed and density.

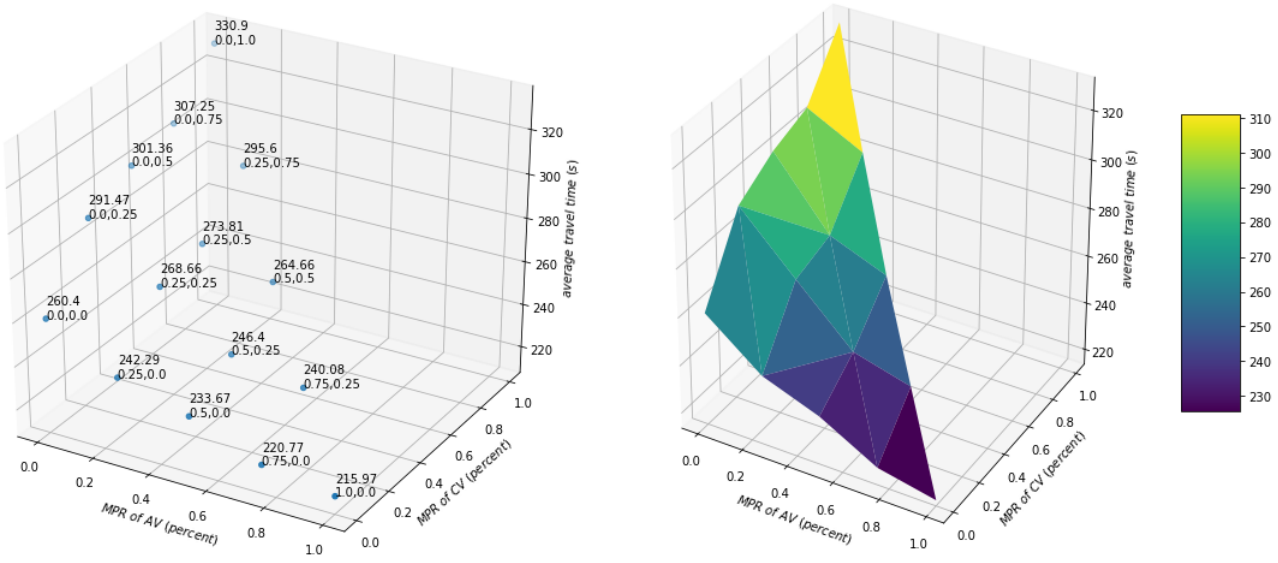
b) Speed and flow.

c) Flow and density.

Source: FHWA, 2019.

Figure 86. Diagram. Fundamental diagrams of mixed traffic scenarios.

The two performance measures (fundamental diagram and average travel time) could also be compared using regret-based metrics to show how not only the selection of the robustness metric but also the choice of the performance metric influences the result of the bi-level analysis. Therefore, the maximum regret of the speed for each scenario was calculated by taking the maximum difference between the speed in the scenario under consideration and the minimum speed in all scenarios. The maximum difference occurs at the highest density value shown on the plot (20 veh/km/lane). On the other hand, the regret summation equals the total area between the speed-density function of the scenario under consideration and the minimum values of speed for all scenarios (figure 88). For the average mainline travel time, both robustness metrics would deliver similar results which would be the difference in the average travel time of the desired scenario and the maximum average travel time among all the scenarios.

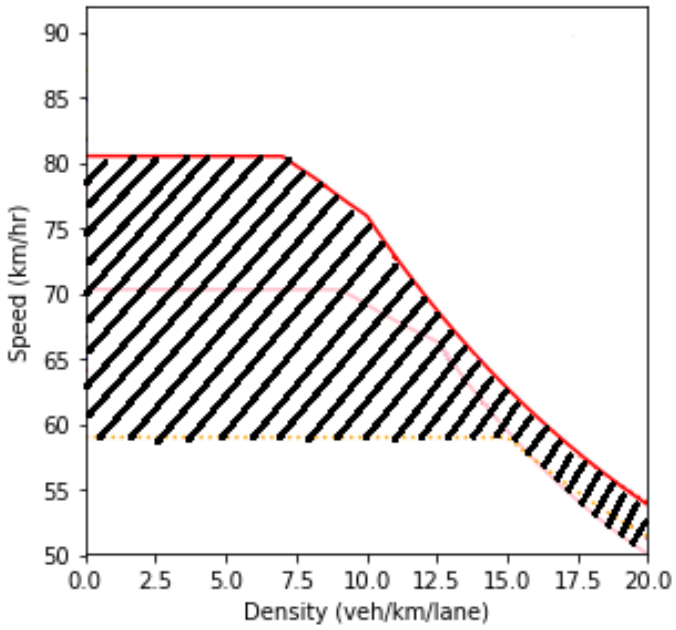


a) Data presented in chart format.

b) Data presented in heat map format.

Source: FHWA, 2019.

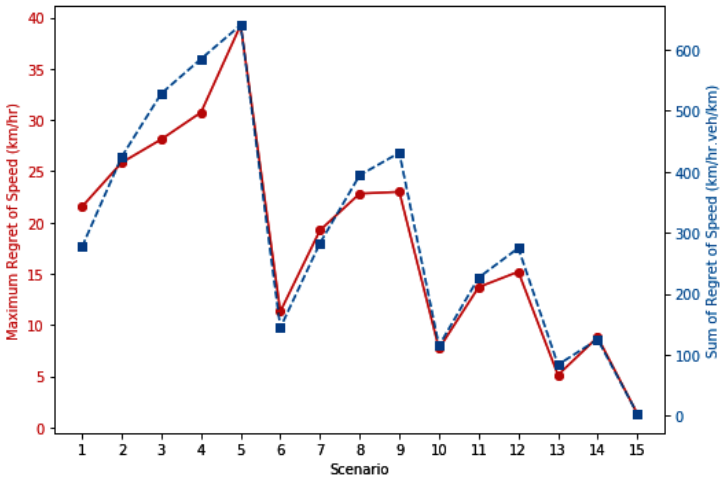
Figure 87. Diagrams. Compound figure depicts average travel time at different market penetration rates for autonomous vehicles and connected vehicles on a selected segment of I-290.



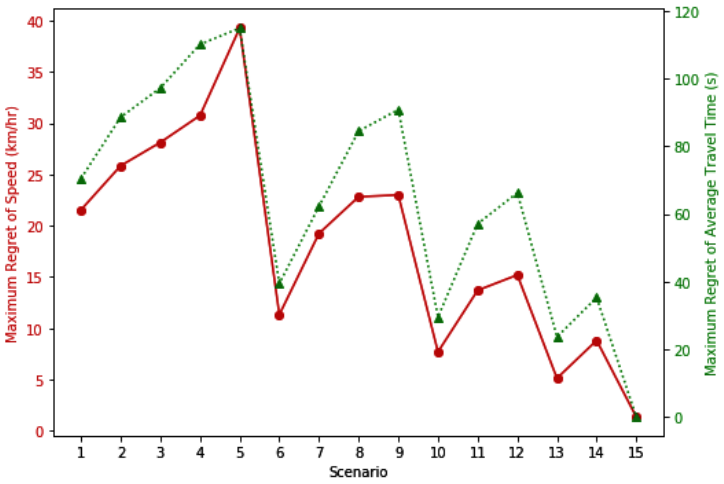
Source: FHWA, 2019.

Figure 88. Diagram. Example for calculating the regret summation for the speed in the speed-density profile.

Figure 89 illustrates the two types of regret-based metrics (maximum regret and regret summation) calculated for the two performance measures. The new scenario IDs are also tabulated in table 12. Even though the general trends of the diagrams are similar to each other, the choice of the performance measure and the robustness metric is important in determining the order of preference between the scenarios.



a) Regret summation.



b) Maximum regret summation.

Source: FHWA, 2019.

Figure 89. Diagram. Regret-based robustness metrics for different performance measures.

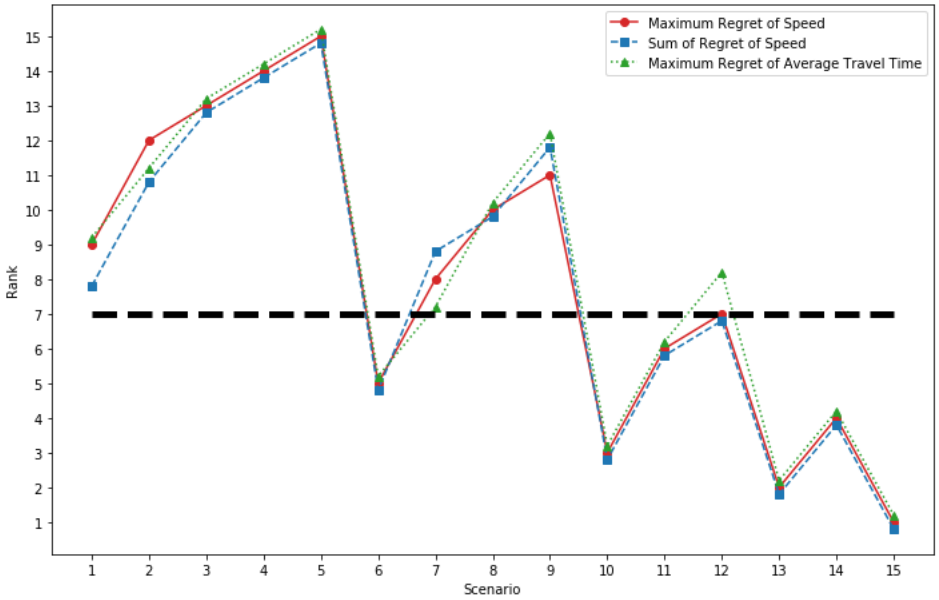
Table 12. New scenario identifiers.

Scenario ID	Connected Vehicles	Automated Vehicles
1	0	0
2	0	25
3	0	50
4	0	75
5	0	100
6	25	0
7	25	25
8	25	50
9	25	75
10	50	0
11	50	25
12	50	50
13	75	0
14	75	25
15	100	0

Figure 90 summarizes the ranking found for each scenario based on the selected set of performance measure and robustness metric. The rankings were assigned in a manner that a higher rank is equivalent to a higher value calculated based on the robustness metric. In other words, the scenario that has the worst performance (least regret) in terms of the robustness value is ranked 1st and the scenario that has the best performance (highest regret) is ranked 15th.

Under the minimax regret and minimum of regret summation, the recommended scenario is scenario 15 which refers to the scenario where all vehicles are fully connected (but without any automation feature). Compared to other scenarios, considering the system to be in a fully connected status would be a conservative choice with the least possible regret. In other words, if the highway segment was designed based on a fully connected environment, then the segment would be overdesigned once autonomous vehicles enter the system.

Comparing the three order sets generated by using the robustness metrics shows that the rankings are mostly consistent for different combinations of regret-based metrics and performance measures. However, there are some differences in the rankings. For example, if a 50th percentile minimax regret is considered (the horizontal dashed line on the graph in figure 90), scenario 7 would be selected based on the average travel time. In contrast, under the same percentile minimax regret, scenario 12 would be selected if the system is evaluated based on speed-density relationship. This shows the difference in the suggested scenario under the robust analysis if different performance measures were used. For scenario 1, the slight difference between the ranking provided by the maximum regret metric for both performance measures and the ranking assigned by the summation of regret metric shows the effect of the robustness metric choice on the result of the analysis. The combined influence of the choices for the regret-based robustness metric and the performance measure is visible in scenario 8 where the rankings provided by the three graphs are different from each other.



Source: FHWA, 2019.

Figure 90. Diagram. Scenario rankings for various regret-based robustness metrics and performance measures.

CONCLUSIONS

This chapter provided case study examples for the concepts of the proposed calibration methodology. Separate examples and combined examples of the scenario-based and robustness-based approaches were exhibited. Within the examples, the process of generating simulation agents and scenarios using the library of model parameters was explained. Agent trajectories constitute another important component of the methodology as well as the main output of the simulation platform. The trajectories were further processed to generate the performance measures such as the fundamental diagrams and the travel time distributions.

Once the feasible scenarios are evaluated, and a complete scenario-based and robustness-based analyses are performed, the outputs of the simulation tool could be stored. As more information becomes available through observing the real-world system, the probability distributions assigned to the scenarios could be updated using various methods such as Bayesian techniques. Furthermore, for cases where a robust analysis was performed due to lack of information regarding the correlations among the scenarios, as data becomes available, the joint probability distribution for the scenario set could be established. As a result, the scenarios could be combined using the scenario-based approach rather than the robustness-based approach. In summary, with the proposed calibration methodology, instead of recalibrating the entire model specifications and parameters, if a sufficient set of scenarios were considered in the calibration process, only the joint probability distribution of the scenarios would need to be adjusted over time.

CHAPTER 7. STEP-BY-STEP APPROACH

The primary components introduced in chapter 5 could be used to develop a step-by-step calibration process so that the analytical/simulation models and tools to data that are reflective of the target future conditions.

Note: Unless accompanied by a citation to statute or regulations, the practices, methodologies, and specifications discussed below are not required under Federal law or regulations.

STEP 1

As the first step, prior to the commencement of any modeling, the study goals and objectives should be determined. At this stage, the granularity level (i.e., micro, meso, or macro) of the problem is specified. A list of performance measures consistent with the study goals should be defined. The relevant field data is another important element that should be identified. The following lists suggested data categories that could serve as input to the simulation tool:

- Network coding data such as link length, number of lanes, lane connectivity, lane use restrictions, lane stripping, turn prohibitions, signal locations, and signal timing.
- Simulation agent behavior data such as fundamental diagrams and driver's reaction time.
- Demand data such as link traffic flows and turning movements.
- Incident data such as location, date, time of clearance, specification of the affected network segments, and severity.
- Work zone data such as work zone activity type, intensity, location, date, time of occurrence, specification of the affected network segments, duration, restrictions during the work zone and activity.
- Special event data such as the type and description of the event, location and area of impact, date, time of occurrence, and duration.
- Weather data such as weather station location (longitude and latitude), date, time of weather record, visibility, precipitation type, precipitation intensity, and temperature.

A list of possible policies tied to the study goals that are being examined by the simulation tool should be provided. These policies would be used throughout the scenario generation.

STEP 2

After the first step, the models that would be used in the simulation should be selected. The models include microscopic models such as car-following, lane-changing, gap acceptance, queue discharge models; mesoscopic and macroscopic models such as fundamental diagrams at the network level and at the link level; and strategic models such as route choice, mode choice, and departure time choice models. Once the models are selected, the input data specified in step 1 could be transformed into variables that are used by the model. Next, the model parameters should be identified. Based on the problem specifications in step 1 and the available data, the parameters could be classified into the categories introduced in chapter 3 (parameters with the least level of uncertainty; parameters with some level of uncertainty; and parameters with deep uncertainty). Categorization of the parameters would help to determine the type of analysis that should be used.

At this stage, some parameters could be estimated from the data collected from the study area or could be prespecified using assumptions consistent with the problem. The remainder of the parameters could be extracted from the library of parameters. These libraries essentially provide a set of possible values for each model parameter. The parameter correlations should be considered when the parameter values are extracted from the libraries. The correlation between parameters is reflected in the joint probability distribution of the model parameters. Once all the parameters are extracted, the set of parameters that need to be calibrated should be determined. As a result, this step could be summarized into answers of the following questions:

- Which models should be used in the simulation tool?
- What are the variables and parameters of the selected model?
- How should the available data be translated into the model variables?
- What category does each model parameter belong to? (Type 1: parameters with the least level of uncertainty; Type 2: parameters with some level of uncertainty; and Type 3: parameters with deep uncertainty)
- How could the available data be used to estimate some of the model parameters?
- What are the parameters that should be extracted from the available libraries?
- What are the possible set of values and joint probability distributions for the parameters that are taken from the libraries?
- Which parameters need to be calibrated for the problem being studied?

STEP 3

In this step, the simulation agents should be generated. The parameters and models specified in the previous step characterize the behavior of the agents. The agents could be created by selecting parameter values from the joint probability distributions of the model parameters. Each set of values chosen for the parameters could be incorporated into the models to define the behavior of an agent.

STEP 4

Once the different agent types are generated in step 3, the various scenarios that would be analyzed by the simulation tool should be created. Scenarios are generated on the basis of the agents created in the previous step and the policies specified in the first step. Scenarios could be classified into three groups:

- External event scenarios such as different weather conditions, incidents, and work zones;
- Traffic control scenarios such as signal control, pricing, ramp metering, variable message signs, managed lanes; and
- Travel demand scenarios such as market penetration rates of different agents, day-to-day demand variation, visitors demand, demand of special events, and closure of alternative modes.

The three classes of scenarios could be combined to generate new scenarios.

STEP 5

After fully characterizing the scenarios, the possibility of constructing a joint probability distribution for the scenarios should be examined. The joint distribution explains the relative importance of each scenario. Therefore, the joint distribution could be characterized based on:

- The joint probability distribution of the model parameters.
- Predictions for the realization of the scenarios under potential future conditions.
- The relative importance of various scenarios according to an expert opinion.

If the joint probability distribution could be constructed, then the step 6-1 should be performed. Otherwise, in the presence of parameters with deep uncertainty, due to lack of information, the joint probability distribution for the scenario set could not be specified. In this case, step 6-2 should be conducted.

An important question to be answered is the number of scenarios that should be generated so that the simulation outputs represent the reality at an acceptable level. Recommendations regarding the number of scenarios are provided in appendix A.

STEP 6-1

Run each scenario in the simulation tool. Then, perform the scenario-based analysis to combine the simulation output using the probabilities assigned to the scenarios. As a result, the combined output is a weighted average of the simulation outputs with the weights equal to the scenario probabilities. The combined output could be changed to become a reasonable representation of the reality by changing the probabilities assigned to the scenarios. A possible approach for updating the probabilities is to Bayesian inference which is explained in appendix B.

The number of simulations for each scenario is an important question that should be addressed in this step. Depending on the available resources, recommendations are provided in appendix A.

Once the simulation output is calibrated by adjusting the probabilities of the analyzed scenarios, the probability distribution of each parameter and the scenario probabilities could be used to define a calibrated probability density function for each parameter.

STEP 6-2

Determine the desired robustness metric to use for combining the outputs of the simulations. Run each scenario in the simulation tool. Perform the robustness-based analysis on the simulation output based on the selected robustness metric. The number of simulations for each scenario is an important question that should be addressed in this step. Depending on the available resources, recommendations are provided in appendix A.

According to the second transformation column in table 6, based on the selected robustness metric, a single scenario or a set of scenarios would be selected to perform the analysis. Once the simulation output under the robustness-based analysis is specified, the probability distribution of each model parameter and the selected scenario(s) could be used to define a proposed probability density function for each parameter.

CHAPTER 8. CONCLUSIONS

Note: Unless accompanied by a citation to statute or regulations, the practices, methodologies, and specifications discussed below are not required under Federal law or regulations.

Models incorporated in simulation tools typically contain mathematical relations. The relations are specified by a set of parameters. Calibration entails the process of estimating values of the model parameters to provide the best representation of the context and study area. Building on previous studies on calibration of simulation tools, this project provides a framework to develop models that are capable to produce realistic predictions of the available problem under potential future conditions and contemplated policy/operational interventions. The motivation for this study was to relax the assumption that the model parameters values remain constant over time. Parameter values may be different under future conditions than they are under the conditions for which the model was calibrated.

In recent years, some research studies on the calibration of traffic analysis tools have provided more information about the potential effectiveness of certain heuristic methods and fitness functions. However, they have not provided robust models that are capable of correctly representing the system behavior for future conditions. On the other hand, other studies attempted to address this challenge by proposing an overall methodology or framework for modeling future conditions. One of the commonly used methodologies is scenario-based simulation. A scenario is defined by a set of operational conditions, interventions, as well as characteristics of the general activity system and associated technologies. In order to improve the accuracy of models in forecasting network behavior under various situations, a set of scenarios could be generated that reflect the expected future conditions of the network. Furthermore, reasonable probabilistic distributions for the scenarios would be used in order to combine the effect of the scenarios. As a result, scenarios play a significant role in depicting a realistic view of the future status of the system. Therefore, the scenario-based simulation methodology constitutes a major component in this study.

Other components of the proposed framework were defined based on other possible sources of error that were identified in previous studies. Besides the scenario-based simulation, parameter correlation, vehicle (and other simulation agents) trajectories, robustness-based simulation, and local density are other major components of the framework.

Recognizing the correlation among model parameters could produce a more reliable model by more accurately capturing the behavior of actors in the transportation network. To preserve the correlation among parameters, there is a need for developing library of parameters. These libraries could be used in different studies as a source of extracting parameter values for different models. Besides the significant value of parameter libraries, trajectories provide the most complete description of the simulation agent's behavior and the system state by retaining the ability to extract stochastic properties of both individual behaviors and performance metrics. The library of parameters and trajectories could provide a comprehensive understanding of the context and study area. Local density recognizes the strong correlation between traffic congestion and driver behavior.

It is defined as the density of traffic in the local vicinity of a vehicle perceived by the drive. Based on the local density, fundamental relationships between driver behaviors can be developed under existing conditions, and then re-used in future models having different densities on each segment. These relationships could be stored as a set of models and parameters in the parameter libraries.

Vehicle trajectories are still not regularly available to traffic analysts, though the situation is rapidly changing with greater willingness by system integrators and data vendors to share this information (albeit with all kinds of limitations on use). However, greater deployment of connected vehicle systems promises to dramatically increase the availability and accuracy of this type of data. Furthermore, the data collected by the sensors of connected vehicles could be used to identify trajectories of other actors (manually-driven vehicles, bicyclists, and pedestrians) in the transportation environment. A complete knowledge of trajectories could be used to specify the feasible set of values for parameters of behavioral models. Such a complete knowledge can support development of model and parameter libraries that could be used for simulation practices.

While the scenario-based simulation, as one of the main concepts in this study, helps to calibrate model parameters when a fully specified set of scenarios are imported to the simulation tool, the notion of robustness becomes helpful when the uncertainty level of parameters and scenarios increases. Since future conditions is naturally governed by uncertainty especially due to lack of data, a sensitivity analysis that accounts for all possibilities could be helpful in developing a reliable model. The robustness-based simulation incorporated into the framework is provides a systematic approach toward such sensitivity analysis.

The proposed framework is predicated on four key notions:

1. That the models themselves are responsive to the features of the future scenario (i.e., that the descriptors that specify a particular scenario be included in the model specification).
2. The definition of a library of model parameters corresponding to different types of agents under varying conditions.
3. The role of scenarios in both specifying future conditions of interest as well as triggering certain ranges of parameter values for those scenarios.
4. The potential for robustness analysis in situations where uncertainty about future scenarios or conditions is large.

Based on the proposed framework and its main components, a step-by-step approach is devised. Follow-on studies could apply the framework and the information to develop robust models that present meaningful and reliable predictions of the system performance under future conditions and contemplated policy/operational interventions. Although the case studies of this report show the application of the proposed calibration framework to a microsimulation problem, the framework could be used for simulations with any granularity level (i.e., micro, meso, or macro).

APPENDIX A. SUGGESTED NUMBER OF SCENARIOS AND NUMBER OF RUNS

Note: Unless accompanied by a citation to statute or regulations, the practices, methodologies, and specifications discussed below are not required under Federal law or regulations.

The number of simulations performed to achieve a desired level of accuracy in the results is an important element of studies that involve simulation. The time allocated to running the simulation tool and the available computational resources play a major role in determining this number. There are two numbers that should be determined for the framework proposed in this report: The number of scenarios to be analyzed, and the number of simulation runs for each scenario.

As discussed in chapter 7, if a joint probability distribution could be defined for the scenarios, the scenario-based analysis would be performed, otherwise, a robustness-based analysis is suggested. The type of analysis also influences the number of scenarios and simulations to be run. Under the scenario-based analysis, there is more knowledge about the scenarios and their probabilities. Therefore, a limited number of scenarios could be analyzed but a higher number of simulation runs could be performed to derive more accurate results for each scenario. On the other hand, in the robustness-based analysis, due to lack of information about the possible set of the scenarios and no/limited information about their associated probabilities, it is better to define a large number of scenarios. Analyzing a relatively comprehensive set of scenarios with a limited number of simulation runs provides an approximate output for each scenario but a better overall picture of the problem.

In most problems, it is not possible to test every possible scenario. As a result, the goal is to derive a distribution of the simulation tool outputs using a limited set of scenarios that closely matches the target distribution for the output. A smoothing method such as the multivariate kernel density estimation method could be utilized to approximate the target simulation output function by smoothing the curve developed by the outputs of the simulated scenarios. The following function represents the smoothed simulation output curve calculated as a function of a multivariate kernel density function such as the standard normal distribution:

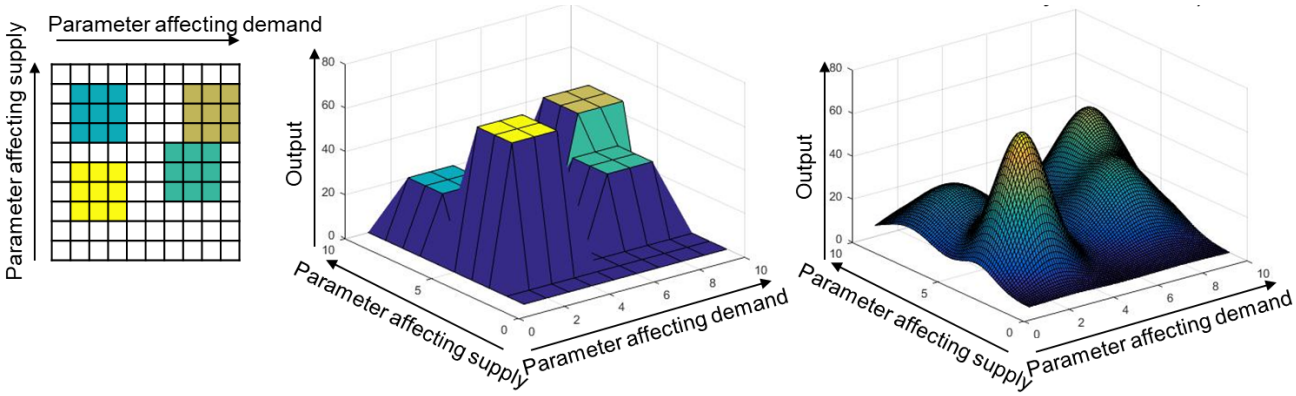
$$\hat{f}(X) = \frac{1}{nh^d}\sum_{i=1}^{n} K\left\{\frac{1}{h}(X - X_i)\right\}$$

$$K\left\{\frac{1}{h}(X - X_i)\right\} = (2\pi)^{-d/2}\exp\left(-\frac{1}{2h^2}(X - X_i)^T(X - X_i)\right)$$

Figure 91. Formula. Multivariate kernel density estimation formula.

Where $\hat{f}(.)$ is the smoothed simulation output curve; X is the output of the simulation tool; X_i is the output of the simulation tool for the ith scenario; K is the kernel function; h is the smoothing parameter; d refers to the number of dimensions (number of varying parameters used to generate the scenarios); and n is the number of scenarios.

Figure 92 schematically represents how to smooth the simulation output by applying the kernel density estimation method. As shown in the figure, variations of two parameters were considered as the basis for developing different scenarios. Using these two dimensions, four distinct scenarios were generated. The three-dimensional diagram in the middle of the figure shows the outputs of the simulation tool for all the scenarios. By applying the kernel density estimation method, a smoothed curve such as the one shown on the right-hand side of the figure could be generated as an approximation of the target output over all feasible scenarios.



a) Scenario space. b) Simulation output. c) Smoothed simulation output (applying kernel density estimation).

Source: FHWA, 2019.

Figure 92. Illustration. Compound figure depicts the process of smoothing the simulation output using the multivariate kernel density estimation method.

A measure that could be used to evaluate the quality of the approximated curve is the relative mean integrated square error defined as follows:

$$\hat{u}^2 = \frac{E\left[\int (\hat{f} - f)^2\right]}{\int f^2}$$

Figure 93. Formula. Relative mean integrated square error.

Where $\hat{f}$ is the smoothed simulation output curve, and f is the target function for the output if all the feasible scenarios are simulated. If both curves ($\hat{f}$ and f are scaled so that the volume under the curve equals one, then the relative mean integrated square error varies between zero and one. Epanechnikov (1969) derived the required sample size (number of scenarios) to ensure that the relative mean square error at zero is less than a specified threshold, when estimating a standard multivariate normal density using a normal kernel and a smoothing parameter that minimizes the mean square error at zero (table 13). These values serve as a starting point to identify the minimum number of scenarios that should be simulated. Based on the general shape and the roughness of the target simulation output curve, the minimum required sample size could differ. Based on table 13, for example, if the simulation tool is used to analyze the effect of demand variation on a network and a relative mean integrated square error of less than 0.3 is acceptable, then it is suggested to

define six scenarios with different demand levels. As another example, under the condition that a relative mean integrated square error of less than 0.2 is acceptable, if the effect of connected vehicles (CVs) and automated vehicles (AVs) on a network is studied, at least 21 scenarios should be generated. Therefore, five different market penetration rates for CVs and five for AVs could be determined. A total of 25 scenarios could be developed based on these two parameters.

Table 13. Suggested number of scenarios to simulate.

Minimum Number of Scenarios		Relative Mean Integrated Square Error ()				
		0.1	0.2	0.3	0.4	0.5
d	1	22	11	6	4	3
	2	58	21	11	7	5
	3	175	52	26	16	11
	4	600	150	67	38	24
	5	2220	470	190	98	59

Source: Epanechnikov 1969.

The number of simulation runs for each scenario is a function of:

- Variance in the simulation outputs (S^2).
- Desired level of confidence (1-α).
- Desired range of confidence interval ($CI_{1-\alpha}$).

These three factors could vary for each scenario. As a result, the number of simulation runs could vary across the scenarios. Based on the above factors the following formula could be used to determine the minimum number of simulations per scenario.

$$N \geq \left(\frac{2 \times t_{1-\frac{\alpha}{2},N-1}}{CI_{1-\alpha}/S} \right)^2$$

Figure 94. Formula. Minimum number of simulation runs for each scenario.

Table 14 provides suggested number of simulation runs calculated by the formula.

Table 14. Minimum number of simulation runs for each scenario.

Minimum Number of Simulation Runs		Level of Confidence (1 – *a*)		
		90%	95%	99%
CI_{1-a}/S	0.5	64	84	131
	1	18	23	36
	1.5	10	12	19
	2	6	8	12
	2.5	5	6	9
	3	4	5	8

APPENDIX B. USING BAYESIAN INFERENCE TO UPDATE SCENARIO PROBABILITIES

Note: Unless accompanied by a citation to statute or regulations, the practices, methodologies, and specifications discussed below are not required under Federal law or regulations.

The joint probability distribution developed over the scenarios in step 5 of the approach could be updated using different techniques. As new information related to the scenarios become available (e.g., through field observations) and/or the analyst belief regarding the relative importance of various scenarios changes, the predefined joint probability distribution may need to be updated. One of the commonly used methods for updating probabilities is the Bayesian inference method. As a statistical inference method, Bayesian inference derives a posterior joint probability distribution by applying the Bayes' theorem to the prior joint probability distribution of the scenarios. The following formula shows the Bayes' rule.

$$P\left(S_k|\{S'_i,\dots,S'_j\}\right) = \frac{P\left(\{S'_i,\dots,S'_j\}|S_k\right) \times P(S_k)}{\sum_{S_m\in\{S_1,S_2,\dots,S_n\}} P\left(\{S'_i,\dots,S'_j\}|S_m\right) \times P(S_m)}$$

Figure 95. Formula. Bayes' rule.

Where S_i is the ith scenario in the prior set of scenarios, S'_i is the ith scenario in the posterior set of scenarios, $P(S_k)$ refers to the prior probability of the kth scenario, and $P\left(S_k|\{S'_i,\dots,S'_j\}\right)$ refers to the posterior probability of the kth scenario.

If there is complete information about each scenario in the prior and posterior stages, and if all the scenarios are mutually exclusive, the following relationships would be applicable.

$$P(S'_l|S_k) = 0 \quad \forall k \neq l$$

$$P\left(\{S'_i,\dots,S'_j\}|S_m\right) = \sum_{l=i}^{j} P(S'_l|S_m) = P(S'_m|S_m)$$

Figure 96. Formula. Relationship between prior and posterior states of mutually exclusive scenarios.

With these two assumptions the Bayes' rule could be revised as follow:

$$P\left(S_k|\{S'_i,\dots,S'_j\}\right) = \frac{P(S'_k|S_k) \times P(S_k)}{\sum_{S_m\in\{S_1,S_2,\dots,S_n\}} P(S'_m|S_m) \times P(S_m)}$$

Figure 97. Formula. Revised Bayes' rule.

In this section, the following two types of examples for the probability updating process are provided using the first case study discussed in chapter 6.

The first case study was selected for these examples because of the mutual exclusiveness of the scenarios:

- Example 1 – One of the scenarios of the prior stage is removed in the posterior stage.
- Example 2 – The probabilities for a set of scenarios are updated in the posterior stage.

The prior probabilities are listed in table 15.

Table 15. Prior probabilities for different demand scenarios.

Scenario	Description	Prior Probability
1	Interarrival time increased by 20% for each path	0.1
2	Interarrival process calibrated based on the I-290 traffic flow	0.25
3	Interarrival time decreased by 20% for each path	0.2
4	Interarrival time decreased by 40% for each path	0.2
5	Interarrival time decreased by 60% for each path	0.15
6	Interarrival time decreased by 80% for each path	0.1

EXAMPLE 1

Let's assume that over time it was realized that scenario 1 should be removed from the scenario set since it would not appear in future conditions of the system. Assuming that the ratio between the probabilities of any two scenarios (other than scenario 1) does not change in the prior and posterior stage, then the following relationships exist between the prior and posterior stages of the scenarios:

$$P(S_1'|S_1) = 0$$

$$P(S_k'|S_k) = 1 \quad \forall k \in \{2,3,4,5,6\}$$

$$P(S_k|\{S_2', \dots, S_6'\}) = \frac{P(S_k'|S_k) \times P(S_k)}{\sum_{S_m \in \{S_1, S_2, \dots, S_6\}} P(S_m'|S_m) \times P(S_m)}$$

Figure 98. Equation. Relationship between prior and posterior probabilities in example 1.

The posterior probabilities calculated based on the above equations are shown in table 16.

Table 16. Prior and posterior probabilities for different demand scenarios in example 1.

Scenario	Prior Probability	Posterior Probability
1	0.1	0
2	0.25	0.2778
3	0.2	0.2222
4	0.2	0.2222
5	0.15	0.1667
6	0.1	0.1111

EXAMPLE 2

Now let's assume that during the first three scenarios correspond to off-peak hour time periods and the last three scenarios correspond to peak hour time periods. If a set of observations during the peak hour shows that scenario 4, 5, and 6 occur 30 percent, 30 percent, and 40 percent of the times, respectively, then the probabilities of these three scenarios should be updated. In this situation, the off-peak period probabilities remain the same since no data is collected during the off-peak hours. The following relationships exist between the prior and posterior stages of the scenarios:

$$P(\{S'_4, S'_5, S'_6\}|S_4) = 0.3$$
$$P(\{S'_4, S'_5, S'_6\}|S_5) = 0.3$$
$$P(\{S'_4, S'_5, S'_6\}|S_6) = 0.4$$
$$P(\{S'_4, S'_5, S'_6\}|\{S'_1, S'_2, S'_3, S'_4, S'_5, S'_6\}) = 0.2 + 0.15 + 0.1 = 0.45$$

$$P(S_k|\{S'_4, S'_5, S'_6\}) = \frac{P(\{S'_4, S'_5, S'_6\}|S_k) \times P(S_k)}{\sum_{S_m \in \{S_4, S_5, S_6\}} P(\{S'_4, S'_5, S'_6\}|S_m) \times P(S_m)} \qquad \forall k \in \{4,5,6\}$$

$$P(S_k|\{S'_1, S'_2, S'_3, S'_4, S'_5, S'_6\}) = P(S_k|\{S'_4, S'_5, S'_6\}) \times P(\{S'_4, S'_5, S'_6\}|\{S'_1, S'_2, S'_3, S'_4, S'_5, S'_6\})$$
$$\forall k \in \{4,5,6\}$$

Figure 99. Equation. Relationship between prior and posterior probabilities in example 2.

The posterior probabilities calculated based on the above equations are shown in table 17.

Table 17. Prior and posterior probabilities for different demand scenarios in example 2.

Scenario	Prior Probability	Posterior Probability
1	0.1	0.1
2	0.25	0.25
3	0.2	0.2
4	0.2	0.186
5	0.15	0.14
6	0.1	0.124

REFERENCES

Ahmed, K. I. 1999. *Modeling drivers' acceleration and lane changing behavior.* Doctoral dissertation, Massachusetts Institute of Technology.

Artimy, M. 2007. "Local density estimation and dynamic transmission-range assignment in vehicular ad hoc networks." *IEEE Transactions on Intelligent Transportation Systems* 8 (3): 400-412.

Balakrishna, R., C. Antoniou, M. Ben-Akiva, H. Koutsopoulos, and Y. Wen. 2007. "Calibration of microscopic traffic simulation models: Methods and application." *Transportation Research Record: Journal of the Transportation Research Board* 1999 (1): 198-207.

Brackstone, M., M. Montanino, W. Daamen, C. Buisson, and V. Punzo. 2012. "Use, Calibration, and Validation of Traffic Simulation Models in Practice: Results of Web-Based Survey." *91st Annual Meeting of the Transportation Research Board.*

Chen, X., and F. B. Zhan. 2008. "Agent-based modeling and simulation of urban evacuation: relative effectiveness of simultaneous and staged evacuation strategies." *Journal of the Operational Research Society* 59 (1): 25-33.

Cheu, R. L., X. Jin, K. C. Ng, Y. L. Ng, and D. Srinivasan. 1998. "Calibration of FRESIM for Singapore expressway using genetic algorithm." *Journal of Transportation Engineering* 124 (6): 526-535.

Chiappone, S., O. Giuffrè, A. Granà, R. Mauro, and A. Sferlazza. 2016. "Traffic simulation models calibration using speed–density relationship: an automated procedure based on genetic algorithm." *Expert systems with applications* 44: 147-155.

Daamen, W., C. Buisson, and S. P. Hoogendoorn. 2014. *Traffic Simulation and Data: Validation Methods and Applications.* CRC Press.

Dowling, R., A. Skabardonis, and V. Alexiadis. 2004. *Traffic analysis toolbox volume III: guidelines for applying traffic microsimulation modeling software.* Washington, D.C.: FHWA-HRT-04-040. Federal Highway Administration.

Epanechnikov, V. A. 1969. "Non-parametric estimation of a multivariate probability density." *Theory of Probability & Its Applications* 14 (1): 153-158.

Frei, A., H. S. Mahmassani, A. Zockaie, and C. Frei. 2014. "Integrating Behavioral Models in Network Operations: Evaluating Traveler Information and Demand Management for Weather-Related Events." *Transportation Research Record: Journal of the Transportation Research Board* 2434 (1): 80-88.

Gao, L., and B. A. Bryan. 2016. "Incorporating deep uncertainty into the elementary effects method for robust global sensitivity analysis." *Ecological modelling* 321 (1): 1-9.

Gao, L., B. A. Bryan, M. Nolan, J. D. Connor, X. Song, and G. Zhao. 2016. "Robust global sensitivity analysis under deep uncertainty via scenario analysis." *Environmental Modelling & Software* 76 (1): 154-166.

Hale, D. K. 2015. "System and method for automated model calibration, sensitivity analysis, and optimization." *US 2015/0039278 A1: Google Patents.*

Hale, D. K., C. Antoniou, M. Brackstone, D. Michalaka, A. T. Moreno, and K. Parikh. 2015. "Optimization-based assisted calibration of traffic simulation models." *Transportation Research Part C: Emerging Technologies* 55: 100-115.

Hamdar, S., M. Treiber, H. Mahmassani, and A. Kesting. 2008. "Modeling driver behavior as sequential risk-taking task." *Transportation Research Record: Journal of the Transportation Research Board* 2088 (1): 208-217.

Harsanyi, J. C. 1967. "Games with incomplete information played by "Bayesian" players, I–III Part I. The basic model." *Management science* 14 (3): 159-182.

Herman, R., L. A. Malakhoff, and S. A. Ardekani. 1988. "Trip time-stop time studies of extreme driver behaviors." *Transportation Research Part A: General* 22 (6): 427-433.

Hou, T., H. Mahmassani, R. Alfelor, J. Kim, and M. Saberi. 2013. "Calibration of traffic flow models under adverse weather and application in mesoscopic network simulation." *Transportation Research Record: Journal of the Transportation Research Board* 2391 (1): 92-104.

Jia, A., B. Williams, and N. Rouphail. 2010. "Identification and calibration of site-specific stochastic freeway breakdown and queue discharge." *Transportation Research Record: Journal of the Transportation Research Board* 2188 (1): 148-155.

Jia, H., Y. Tan, and L. Yang. 2011. "Modeling vehicle merging behavior in urban expressway merging sections based on logistic model." *2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE).* 656-659.

Kim, J., and H. Mahmassani. 2011. "Correlated parameters in driving behavior models: Carfollowing example and implications for traffic microsimulation." *Transportation Research Record: Journal of the Transportation Research Board* 2249 (1): 62-77.

Kim, J., and H. S. Mahmassani. 2015. "Spatial and temporal characterization of travel patterns in a traffic network using vehicle trajectories." *Transportation Research Part C: Emerging Technologies* 59: 375-390.

Kim, J., H. Mahmassani, P. Vovsha, Y. Stogios, and J. Dong. 2013. "Scenario-based approach to analysis of travel time reliability with traffic simulation models." *Transportation Research Record: Journal of the Transportation Research Board* 2391 (1): 56-68.

Kim, S. J., W. Kim, and L. Rilett. 2005. "Calibration of microsimulation models using nonparametric statistical techniques." *Transportation Research Record: Journal of the Transportation Research Board* 1935 (1): 111-119.

Lee, J. B., and K. Ozbay. 2009. "New calibration methodology for microscopic traffic simulation using enhanced simultaneous perturbation stochastic approximation approach." *Transportation Research Record: Journal of the Transportation Research Board* 2124 (1): 233-240.

Luo, X., X. Wang, P. Wang, F. Liu, and N. N. Van. 2017. "Local Density Estimation Based on Velocity and Acceleration Aware in Vehicular Ad-Hoc Networks." *International Conference on Machine Learning and Intelligent Communications.* Springer, Cham. 463-471.

Ma, J., H. Dong, and H. Zhang. 2007. "Calibration of microsimulation with heuristic optimization methods." *Transportation Research Record: Journal of the Transportation Research Board* 1999 (1): 208-217.

Ma, T., and B. Abdulhai. 2002. "Genetic algorithm-based optimization approach and generic tool for calibrating traffic microscopic simulation parameters." *Transportation Research Record: Journal of the Transportation Research Board* 1800 (1): 6-15.

Mahmassani, H. S. 2016. "50th anniversary invited article—autonomous vehicles and connected vehicle systems: flow and operations considerations." *Transportation Science* 50 (4): 1140-1162.

Mahmassani, H. S., J. Kim, Y. Chen, Y. Stogios, A. Brijmohan, and P. Vovsha. 2014. *Incorporating Reliability Performance Measures into Operations and Planning Modeling Tools.* Transportation Research Board.

Mahmassani, H. S., T. Hou, J. Kim, Y. Chen, Z. Hong, H. Halat, and R. Haas. 2014. *Implementation of a weather responsive traffic estimation and prediction system (TrEPS) for signal timing at Utah DOT.* United States: FHWA-OP-02-031. Joint Program Office for Intelligent Transportation Systems.

Mahmassani, H., and Y. Sheffi. 1981. "Using gap sequences to estimate gap acceptance functions." *Transportation Research Part B: Methodological* 15 (3): 143-148.

McPhail, C., H. R. Maier, J. H. Kwakkel, M. Giuliani, A. Castelletti, and S. Westra. 2018. "Robustness Metrics: How Are They Calculated, When Should They Be Used and Why Do They Give Different Results?" *Earth's Future* 6 (2): 169-191.

Menneni, S., C. Sun, and P. Vortisch. 2008. "Microsimulation calibration using speed-flow relationships." *Transportation Research Record: Journal of the Transportation Research Board* 2088 (1): 1-9.

Ossen, S., and S. Hoogendoorn. 2008. "Validity of trajectory-based calibration approach of car-following models in presence of measurement errors." *Transportation Research Record: Journal of the Transportation Research Board* 2088 (1): 117-125.

Papaioannou, P. 2007. "Driver behaviour, dilemma zone and safety effects at urban signalised intersections in Greece." *Accident Analysis & Prevention* 39 (1): 147-158.

Park, B., and H. Qi. 2005. "Development and Evaluation of a Procedure for the Calibration of Simulation Models." *Transportation Research Record: Journal of the Transportation Research Board* 1934 (1): 208-217.

Paz, A., V. Molano, Martinez, Gaviria, C. E., and C. Arteaga. 2015. "Calibration of traffic flow models using a memetic algorithm." *Transportation Research Part C: Emerging Technologies* 55 (1): 432-443.

Perez, B. G., T. Batac, and P. S. Vovsha. 2012. *Assessing Highway Tolling and Pricing Options and Impacts. Volume 2.* Washington D.C.: The National Academies Press.

Prakash, A. A., Seshadri R., C. Antoniou, F. C. Pereira, and M. E. Ben-Akiva. 2018. "Improving scalability of generic online calibration for real-time dynamic traffic assignment systems." *97th Annual Meeting of the Transportation Research Board.*

Punzo, V., and F. Simonelli. 2005. "Analysis and comparison of microscopic traffic flow models with real traffic microscopic data." *Transportation Research Record: Journal of the Transportation Research Board* 1934 (1): 53-63.

Punzo, V., M. T. Borzacchiello, and B. Ciuffo. 2011. "On the assessment of vehicle trajectory data accuracy and application to the Next Generation SIMulation (NGSIM) program data." *Transportation Research Part C: Emerging Technologies* 19 (6): 1243-1262.

Rahman, M., M. Chowdhury, Y. Xie, and Y. He. 2013. "Review of microscopic lane-changing models and future research opportunities." *IEEE transactions on intelligent transportation systems* 14 (4): 1942-1956.

Reece, D. A., and S. A. Shafer. 1993. "A computational model of driving for autonomous vehicles." *Transportation Research Part A: Policy and Practice* 27 (1): 23-50.

Saberi, M., H. Mahmassani, T. Hou, and A. Zockaie. 2014. "Estimating network fundamental diagram using three-dimensional vehicle trajectories: extending edie's definitions of traffic flow variables to networks." *Transportation Research Record: Journal of the Transportation Research Board* 2422 (1): 12-20.

So, J., M. Ostojic, D. Jolovic, and A. Stevanovic. 2016. "Building, Calibrating, and Validating a Large-scale High Fidelity Microscopic Traffic Simulation Model - a Manual Approach." *95th Annual Meeting of the Transportation Research Board.*

Talebpour, A., and H. S. Mahmassani. 2016. "Influence of connected and autonomous vehicles on traffic flow stability and throughput." *Transportation Research Part C: Emerging Technologies* 71 (1): 143-163.

Talebpour, A., H. Mahmassani, and S. Hamdar. 2011. "Multiregime sequential risk-taking model of car-following behavior: specification, calibration, and sensitivity analysis." *Transportation Research Record: Journal of the Transportation Research Board* 2260 (1): 60-66.

Talebpour, A., H. S. Mahmassani, and F. E. Bustamante. 2016. "Modeling driver behavior in a connected environment: Integrated microscopic simulation of traffic and mobile wireless telecommunication systems." *Transportation Research Record: Journal of the Transportation Research Board* 2560 (1): 75-86.

Talebpour, A., H. S. Mahmassani, and S. H. Hamdar. 2015. "Modeling lane-changing behavior in a connected environment: A game theory approach." *Transportation Research Part C: Emerging Technologies* 59 (1): 216-232.

Tang, T., C. Li, H. Huang, and H. Shang. 2012. "A new fundamental diagram theory with the individual difference of the driver's perception ability." *Nonlinear Dynamics* 67 (3): 2255-2265.

Taylor, D., and H. Mahmassani. 1998. "Bicyclist and motorist gap acceptance behavior in mixed traffic." *78th Annu. Meeting Transportation Research Board and Publication in Transportation Research Board.*

Van Arem, B., C. J. Van Driel, and R. Visser. 2006. "The impact of cooperative adaptive cruise control on traffic-flow characteristics." *IEEE Transactions on Intelligent Transportation Systems* 7 (4): 429-436.

Vovsha, P., R. Donnelly, M. Bradley, J. Bowman, H. Mahmassani, T. Adler, K. Small, et al. 2013. *Improving our Understanding of How Highway Congestion and Price Affect Travel Demand* . Washington D.C.: The National Academies Press. SHRP 2 Report S2-C04-RW-1.

Wegmann, F., and J. Everett. 2012. "Minimum Travel Demand Model Calibration and Validation Guidelines for the State of Tennessee Update 2012."

Wunderlich, K. 2002. *Incorporating intelligent transportation systems into planning analysis: summary of key findings from a 2020 case study--improving travel time reliability with ITS.* United States: FHWA-OP-02-031. Joint Program Office for Intelligent Transportation Systems.

Yu, M., and W. D. Fan. 2017. "Calibration of microscopic traffic simulation models using metaheuristic algorithms." *International Journal of Transportation Science and Technology* 6 (1): 63-77.

U.S. Department of Transportation
Federal Highway Administration

U.S. Department of Transportation
Federal Highway Administration
1200 New Jersey Avenue, SE
Washington, DC 20590

November 2020

FHWA-HOP-20-057