



UNIVERSITY OF  
MARYLAND

---

## National Transportation Center

**Project ID: NTC2015-MU-R06**

### **GUIDANCE FOR FIELD AND SENSOR-BASED MEASUREMENT OF HCM AND SIMULATION PERFORMANCE MEASURES**

#### **Final Report**

by

Nagui M. Roupail, PhD  
Professor of Civil Engineering  
North Carolina State University in Raleigh  
roupail@ncsu.edu  
Tel: 919-515-1154

Kihyun Pyo and Thomas Chase  
Institute for Transportation Research and Education  
North Carolina State University

Mecit Cetin, PhD  
Old Dominion University

for

National Transportation Center at Maryland (NTC@Maryland)  
1124 Glenn Martin Hall  
University of Maryland  
College Park, MD 20742

**March 2018**



## **ACKNOWLEDGEMENTS**

This project was funded by the National Transportation Center @ Maryland (NTC@Maryland), one of the five National Centers that were selected in this nationwide competition, by the Office of the Assistant Secretary for Research and Technology (OST-R), U.S. Department of Transportation (US DOT). The authors are grateful for the financial support received by NTC@Maryland, which has yielded one journal publication in press, two conference presentations and the successful completion of a doctoral degree.

## **DISCLAIMER**

The contents of this report reflect the views of the authors, who are solely responsible for the facts and the accuracy of the material and information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation University Transportation Centers Program and the North Carolina Department of Transportation matching funds in the interest of information exchange. The U.S. Government and the NC Department of Transportation assume no liability for the contents or use thereof. The contents do not necessarily reflect the official views of the U.S. Government and the NC Department of Transportation. This report does not constitute a standard, specification, or regulation.

# TABLE OF CONTENTS

<b>EXECUTIVE SUMMARY .....</b>	<b>1</b>
<b>1.0 INTRODUCTION &amp; BACKGROUND.....</b>	<b>2</b>
1.1 BACKGROUND .....	2
1.2 RESEARCH OBJECTIVES .....	2
1.3 PROCESS FLOW CHART .....	3
1.4 REPORT ORGANIZATION.....	3
<b>2.0 LITERATURE REVIEW .....</b>	<b>4</b>
2.1 DATA COLLECTION TECHNOLOGIES AND APPLICATIONS.....	4
2.2 BLUETOOTH.....	4
2.2.1 Technology Description.....	4
2.2.2 Applications in Performance Measures .....	5
2.3 INDUCTIVE LOOP DETECTOR & WEIGH IN MOTION.....	7
2.3.1 Technology Description.....	7
2.3.2 Applications in Performance Measures .....	8
2.4 MICROWAVE RADAR .....	9
2.4.1 Technology Description.....	9
2.4.2 Applications in Performance Measures .....	10
2.5 VIDEO AND LICENSE PLATE RECOGNITION (LPR) .....	11
2.5.1 Technology Description.....	11
2.5.2 Applications in Performance Measures .....	12
2.6 AUTOMATED VEHICLE IDENTIFICATION .....	12
2.6.1 Technology Description.....	12
2.6.2 Applications in Performance Measures .....	13
2.7 PROBE VEHICLE FLEET DATA .....	13
2.7.1 Technology Description.....	13
2.7.2 Applications in Performance Measures .....	15
2.8 INSTRUMENTED VEHICLE DATA .....	16
2.8.1 Technology Description.....	16
2.8.2 Applications in Performance Measures .....	16
2.9 FREEWAY CAPACITY AND PRE-BREAKDOWN FLOW.....	17
2.9.1 Definition of Breakdown .....	17
2.9.2 Characteristics of Breakdown .....	18
2.9.2.1 Breakdown Probability Model.....	18
2.9.2.2 Identifying Breakdown.....	19
2.9.3 Breakdown Prediction Model in Planning Level.....	19
2.9.3.1 Traffic Prediction Techniques and Statistical Model .....	19
2.9.3.2 Breakdown Prediction Model .....	20
2.9.4 Freeway Capacity.....	20
<b>3.0 COMPARISON OF FREEWAY CAPACITY AND PRE-BREAKDOWN FLOW ESTIMATION METHODS.....</b>	<b>22</b>
3.1 METHODOLOGY COMPARISON .....	22
3.2 DATA SOURCES .....	22
3.3 ANALYSIS AND RESULTS.....	24

<b>4.0</b>	<b>PREDICTIVE MODEL OF BREAKDOWN FOR PLANNING LEVEL APPLICATIONS .....</b>	<b>31</b>
4.1	BACKGROUND .....	31
4.2	DATA COLLECTION .....	32
4.3	OVERALL FRAMEWORK.....	32
4.4	IDENTIFYING PEAK PERIOD BREAKDOWN .....	33
4.5	PLANNING LEVEL VARIABLES .....	35
4.6	MODEL DEVELOPMENT.....	36
4.7	CALIBRATION AND VALIDATION SETS.....	36
4.8	CALIBRATION RESULTS .....	37
4.9	VALIDATION COMPARISONS .....	40
<b>5.0</b>	<b>CONCLUSIONS AND FUTURE WORK .....</b>	<b>44</b>
5.1	SUMMARY OF RESEARCH AND FINDINGS.....	44
5.2	RECOMMENDATIONS FOR FUTURE RESEARCH.....	44
<b>6.0</b>	<b>REFERENCES.....</b>	<b>45</b>

## LIST OF TABLES

Table 1	Data Collection Technologies and Use Cases .....	4
Table 2	Breakdown and Capacity Definitions .....	22
Table 3	Capacity Comparison Site Descriptions .....	24
Table 4	ANOVA Results for Capacity and Bottleneck Methods .....	27
Table 5	Capacity Regression Parameters.....	29
TABLE 6	Summary of Site Distribution .....	32
TABLE 7	Calibration and Validation Dataset.....	37
Table 8	Calibrated Linear Regression Models.....	39
Table 9	Comparison of Site-Specific Fitted vs. Model Predicted Parameter Values .....	40
Table 10	Effects on Predicted Beta of Simultaneously Changing Free Flow Speed (FFS) and Inclement Weather .....	42

## LIST OF FIGURES

Figure 1–	The process flow of the research approach .....	3
Figure 2	The BluFax Concept & Device .....	5
Figure 3	BlueTOAD .....	5
Figure 4	Bluetooth validation one-step ahead forecasts (Click, 2014).....	6
Figure 5	The BlueMAC reports: Origin-Destination.....	7
Figure 6	Weigh in motion site for traffic data collection (KISTLER) .....	8
Figure 7	Kapsch Weigh in motion .....	8
Figure 8	EMX industries Loop Detect Devices .....	9
Figure 9	Microwave radar operation.....	9
Figure 10	Wavetronix .....	10
Figure 11	Vehicle-Based Detection & Dual Radar (Wavetronix) .....	10
Figure 12	IZ100series ALPR Camera System (INEXZAMIR).....	11

Figure 13 IZ500series ALPR Camera System (INEXZAMIR).....	12
Figure 14 LR-6XL reader (TagMaster) .....	13
Figure 15 RFID (CONFIDEX) .....	13
Figure 16 INRIX Connected Driver Network .....	14
Figure 17 Congestion Scan by INRIX .....	15
Figure 18 INRIX System Monitoring Dashboard.....	15
Figure 19. Speed comparison with GPS and OBD .....	17
Figure 20. GPS Location error.....	17
Figure 21 HCM-based Capacity by Breakdown Type.....	24
Figure 22 Florida-based Capacity by Breakdown Type .....	25
Figure 23 SFI Optimal Flow by Breakdown Type .....	25
Figure 24 Capacity Regression Fit Diagnostics.....	30
Figure 25 Overall Processing of the Methodology .....	33
FIGURE 26 Breakdown Activation Time Frequency for all sites.....	34
FIGURE 27 Theoretical Feature in Weibull Distribution.....	35
FIGURE 28 Modeled Breakdown Start Time Sensitivities.....	40
Figure 29 Validation Fitting CDF Plots Result.....	41
Figure 30 Sensitivity Analysis for Predicted Beta (Inclement weather in Excluded Set) .....	43





## EXECUTIVE SUMMARY

Performance measurement estimates in the HCM 6<sup>th</sup> Edition are heavily relied upon by practitioners and decision makers in the planning, design, and operations of our transportation systems. Despite their significance, limited guidance is available for properly collecting and applying field data to validate HCM predictions. In some cases it is difficult to obtain measurements (i.e., freeway density over a segment); in other cases there are inconsistent practices for determining measures (i.e., arterial travel speed). In addition to the HCM, microsimulation tools are frequently applied for traffic analysis, but guidance for field-measurement for calibration and validation of simulation-based metrics is similarly missing.

The cost and/or lack of availability of field data has historically been a barrier for analysts to verify their results, in which cases HCM and simulation users rely on a “reasonableness” check based on visual observations and judgment. However, the continued advancement of automated data from detectors and probes has remarkably reduced the barrier associated with obtaining field data to compute performance measures. The advancement of automated traffic data has the potential to add significant value to analysts. Crisp and clear definitions are needed to ensure that field data are properly collected, applied, and evenly compared against HCM or simulation predictions for all facility types.

Development of definitions and field-measurement techniques should be supported by real-world case studies that compare field-to-model results for key performance measures across multiple facility types. The definitions and methods should be vetted and approved by the TRB Highway Capacity and Quality of Service (HCQS) committee and ultimately incorporated into the HCM. Similarly, simulation guidance through, for example, the FHWA Traffic Analysis Toolbox should consider field-estimation guidance. With new emphasis on performance measurements in the MAP-21 legislation, clear guidance and consistency in performance estimation is more important than ever.

This research effort documents the field data collection technologies available for performance measurement as well as their use cases and limitations. The researchers identified freeway capacity as a performance measure with inconsistent measurement and application across the United States. Two applications of capacity were developed, the first addressing the difference in estimation methods for pre-breakdown flow and capacity used around the US and the second using regression analysis to identify geometric and land use impacts on the time of day of freeway congestion.

# 1.0 INTRODUCTION & BACKGROUND

## 1.1 BACKGROUND

Performance measurement estimates in the HCM 6<sup>th</sup> Edition (Transportation Research Board, 2016) are heavily relied upon by practitioners and decision makers in the planning, design, and operations of our transportation systems. Despite their significance, limited guidance is available for properly collecting and applying field data to validate HCM predictions. In some cases it is difficult to obtain measurements (i.e., freeway density over a segment); in other cases there are inconsistent practices for determining measures (i.e., arterial travel speed). In addition to the HCM, microsimulation tools are frequently applied for traffic analysis, but guidance for field-measurement for calibration and validation of simulation-based metrics is similarly missing.

The cost and/or lack of availability of field data has historically been a barrier for analysts to verify their results, in which cases HCM and simulation users rely on a “reasonableness” check based on visual observations and judgment. However, the continued advancement of automated data from detectors and probes has remarkably reduced the barrier associated with obtaining field data to compute performance measures. The advancement of automated traffic data has the potential to add significant value to analysts. Crisp and clear definitions are needed to ensure that field data are properly collected, applied, and evenly compared against HCM or simulation predictions for all facility types.

Development of definitions and field-measurement techniques should be supported by real-world case studies that compare field-to-model results for key performance measures across multiple facility types. The definitions and methods should be vetted and approved by the TRB Highway Capacity and Quality of Service (HCQS) committee and ultimately incorporated into the HCM. Similarly, simulation guidance through, for example, the FHWA Traffic Analysis Toolbox should consider field-estimation guidance. With new emphasis on performance measurements in the MAP-21 legislation, clear guidance and consistency in performance estimation is more important than ever.

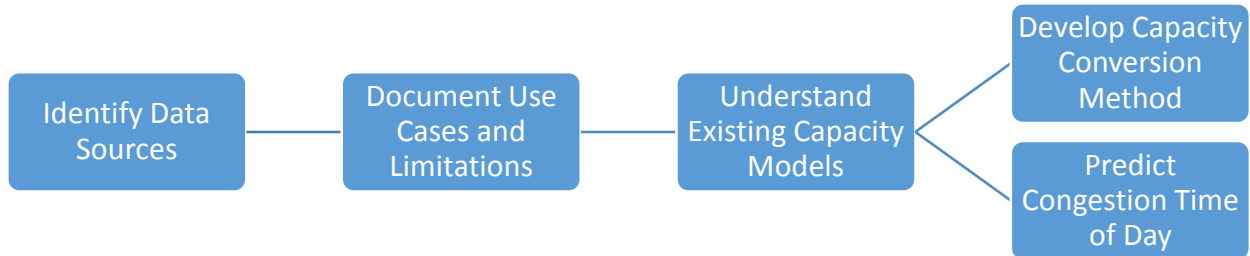
## 1.2 RESEARCH OBJECTIVES

The objectives of this research are to:

- 1) Identify technologies available to collect traffic data for performance measurement.
- 2) Document the use cases and limitations of the technology or collection methodology.
- 3) Develop new techniques to compare and use disparate methodologies used in the United States for capacity and pre-breakdown flow estimation:
  - a. Develop a comparison model to allow for agencies to compare output capacities from different capacity estimation methods.
  - b. Develop a regression analysis of the time of day of freeway congestion to identify the impacts of geometrics and land use.

### 1.3 PROCESS FLOW CHART

Figure 1 shows the process flow for the research project, with two major new methodologies based on the literature search and technology documentation.



**Figure 1– The process flow of the research approach**

### 1.4 REPORT ORGANIZATION

This report is organized in 5 Chapters. Following this introductory chapter, a review of the technology options and their use cases and limitations for performance measure data collection is given in Chapter 2. Chapter 3 documents a new comparison model developed to allow agencies to compare output capacities from different capacity estimation methods. Chapter 4 focuses on a new prediction model of the time of day of congestion based on geometric and land use factors. Chapter 5 gives overall conclusions and recommendations for future work.

## 2.0 LITERATURE REVIEW

### 2.1 DATA COLLECTION TECHNOLOGIES AND APPLICATIONS

A total of seven technologies were identified in use by agencies in collecting performance measures. Ten performance measures were identified from the HCM and simulation that could be collected from at least one of the technologies. Table 1 shows the combination of each technology and the potential for use in performance measurement. The following sections provide documentation and examples of each technology.

**Table 1 Data Collection Technologies and Use Cases**

Data Source		1. Bluetooth Probes	2. Inductive Loop and Weigh in Motion	3. Microwave Radar	4. Video and ALPR And ALP	5. Automated Vehicle Identification	6. Probe Vehicle Fleets	7. Instrumented Vehicles
1. Speed	Spot							
	Segment							
2. Travel Time								
3. Capacity								
4. Flow								
5. Occupancy								
6. Demand								
7. % Trucks								
8. Density								
9. Queue								
10. O-D								

<b>Primary Use</b>	<b>Secondary Use</b>	<b>Not appropriate</b>
--------------------	----------------------	------------------------

### 2.2 BLUETOOTH

#### 2.2.1 Technology Description

Bluetooth devices are a standard technology used for short-range, low power, wireless data transfer. It can make personal area networks, enabling communication with other fixed or mobile devices by using radio signals, specifically using short-wavelength UHF radio waves in the unlicensed industrial, scientific, and medical (ISM) band from 2.4 to 2.485 GHz (BlueMAC Analytics, 2015). It also involves technologies concerned with a spread spectrum, frequency hopping, full-duplex signal at a nominal rate of 1600 hops/sec. Currently, the 2.4 GHz ISM band is available and unlicensed in most countries (Bluetooth SIG, 2015).

## 2.2.2 Applications in Performance Measures

A Bluetooth data source has many benefits and uses, as it can monitor road speeds and travel times in real-time. From Traffax's BluFAX monitoring sensors, efficient and reliable real time logging, processing, display and reporting of data is able to be provided (Traffax, 2015). According to the BlueTOAD Company, the data is archived for robust analysis of not only travel time and speed but also speed trends, origin-destination, route patterns, trip length analysis and signal timing studies, MAC address detection counts and so forth (Trafficcast, 2015). **Figure 2** **Error! Reference source not found.** shows the BluFAX concept with device and the installation and operation of the Bluetooth device, BlueTOAD can be seen in Figure 3.

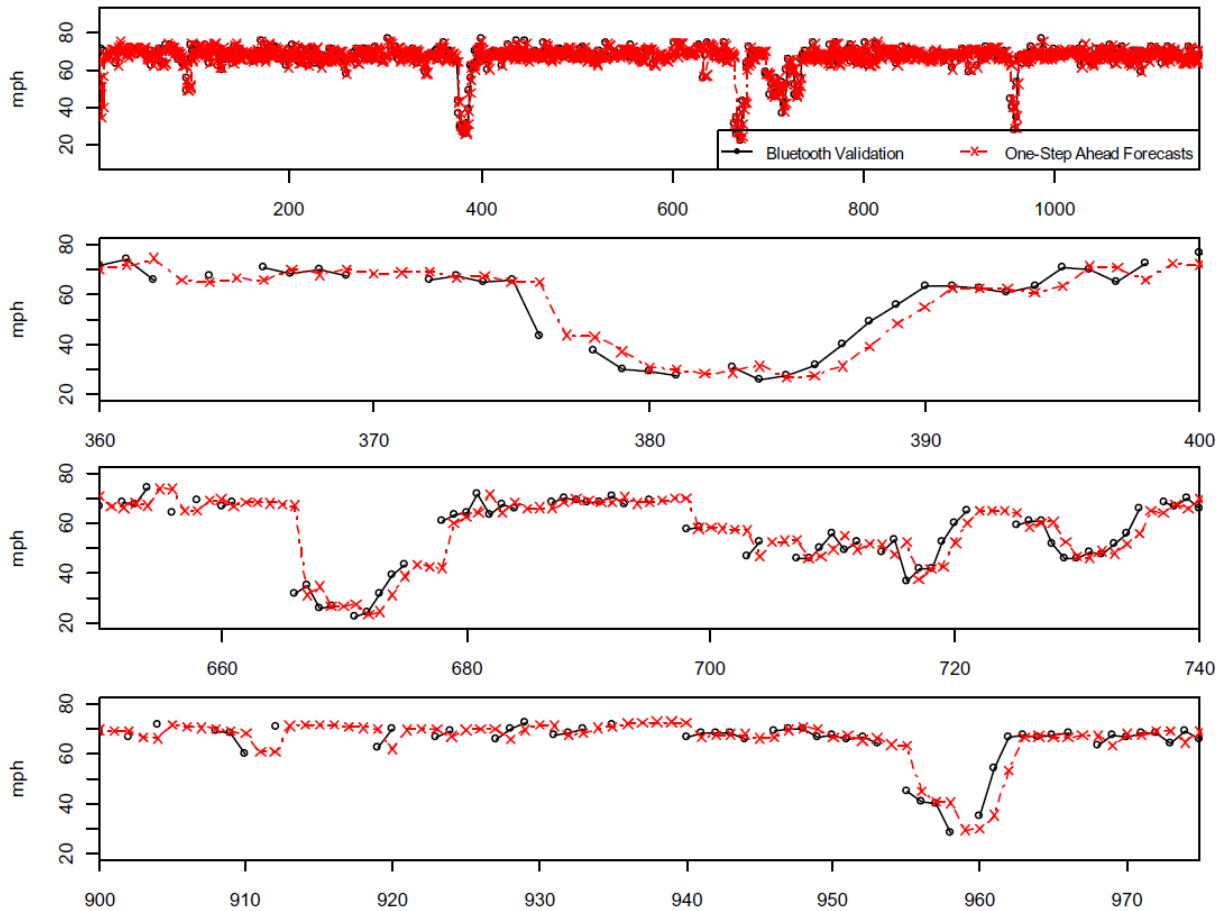


**Figure 2 The BluFAX Concept & Device**



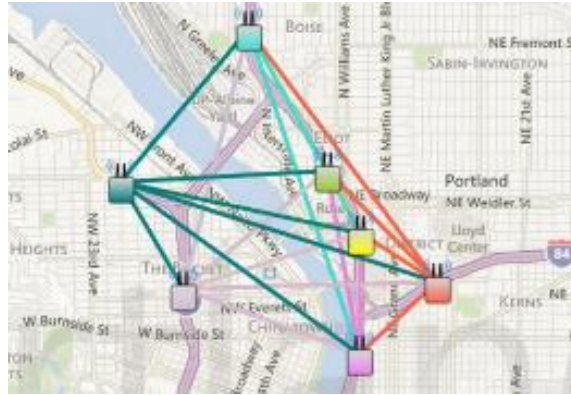
**Figure 3 BlueTOAD**

With respect to performance measure for speed, Brian Portugais performed adaptive traffic speed estimation research to validate collected data from Bluetooth based space-means speeds by comparing the results with data collected radar based sensors (Portugais, 2014). Figure 4 shows the results of Bluetooth validation in one case study. Steven focused on experimenting with Bluetooth-based speed data in order to check validity on the rural freeway. He mentioned that Bluetooth data collection is feasible on a rural freeway because it has confidence with a high enough sample size during daytime hours, but not overnight hours (Click, 2012).



**Figure 4 Bluetooth validation one-step ahead forecasts (Click, 2014)**

According to several company websites, travel time reporting is a key factor in using Bluetooth devices in the traffic environments. The BluFax portable and real-time systems are being used to collect data to drive travel time information all over the world. It provides cost-effective, accurate travel time data on freeways, as well as arterial roadways. There are several examples to apply Bluetooth. In the I-95 Corridor Validation Study, the BluFax data offered a baseline for validating GPS data acquired from INRIX and it was also used to evaluate signal timing changes along arterials (Traffax, 2015). BlueMAC created intelligent travel time reports automatically with detailed graphs of multiple variables, and tables with specific data points, and it also enable to access this data and save in a CSV format (BlueMAC Analytics, 2015). Qiao performed a study on stochastic freeway applications with short-time travel time prediction by using real time Bluetooth travel time data (Qiao, 2013). Porter tried to assess appropriateness with five different kinds of antennas in order to improve the Bluetooth-based travel time data collection system (Porter, 2013).



**Figure 5 The BlueMAC reports: Origin-Destination**

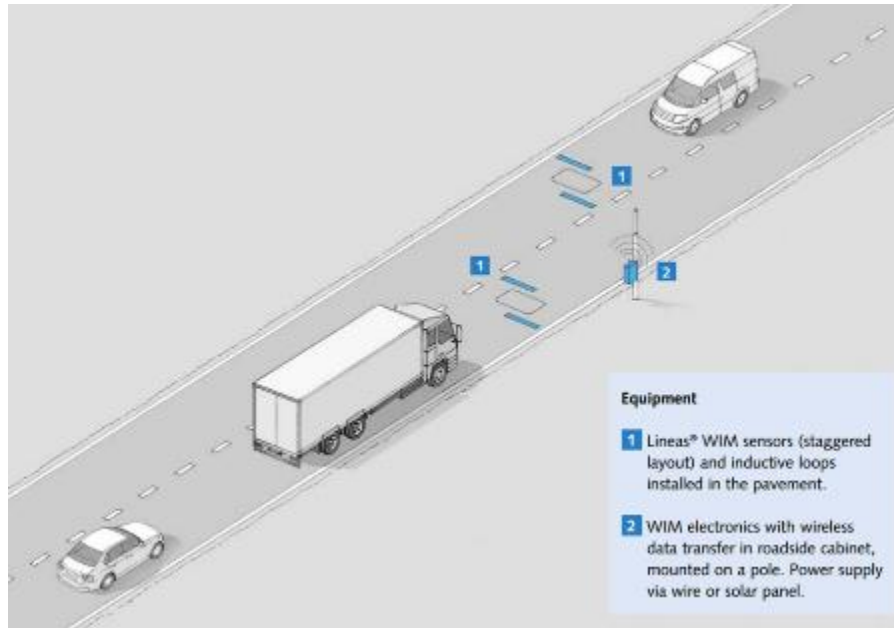
Traffax Company mentioned that while the concept of travel time and origin-destination applications is similar, the approach method for origin and destination applications has to fulfill requirements of each study. The BluFax device and BluSTAT software deal with various study cases like the following: Interchange studies, Freeway Corridor Origin-Destination Studies, Freeway-Arterial Corridor Studies, Area-wide, large-scale origin-destination studies, Cordon Studies, High-density urban environments, Pedestrian Origin-Destination Studies. The BlueMAC provides information like: how many trips were made to each location, the percentage of the traffic flow that went to each destination, and much more. As you can see from Figure 5, BlueMAC support data for Origin-Destination studies that can be reported in a variety of graphs and tables, it also provides accessible data in a CSV format like travel time.

## 2.3 INDUCTIVE LOOP DETECTOR & WEIGH IN MOTION

### 2.3.1 Technology Description

There are many studies and trials in terms of parameter measure based on loop detector. However, because loop detector has some deficiency like inaccuracy of data although it can cover almost type of vehicles, recently fusing data combined loop detector and probe data or Bluetooth have performed by many researchers and engineers.





**Figure 6 Weigh in motion site for traffic data collection (KISTLER)**

### 2.3.2 Applications in Performance Measures

Figure 6 shows the KISTLER weigh in motion technology. The Kistler loop detector company mentioned that loop detector can provide not only monitoring traffic real time but also key vehicle data like vehicle weight and imbalance, axle loads and distances, vehicle speed and driving behavior and much more through their brochure of web-site (Kistler, 2015). In order to improve freeway traffic speed estimation, Bachmann explored several techniques for loop detector with probe vehicle as a fusing data (Bachmann, 2013). Loop detectors cover almost all the vehicles with large amount of sample size but it is not imprecise which have some error data. Otherwise, the quality of probe vehicle data is excellent but it is only a small portion of the vehicles. This is because he used concept of fusing data to compensate their defects in order to improve assess performance measure for speed.



**Figure 7 Kapsch Weigh in motion**





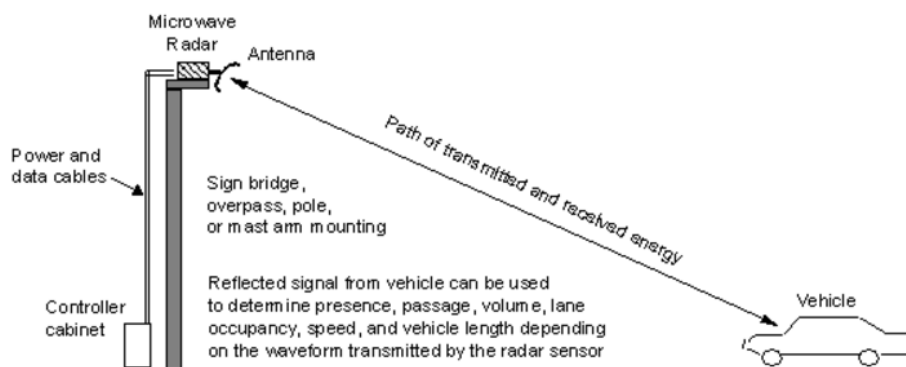
**Figure 8 EMX industries Loop Detect Devices**

Furthermore, some company website deal with both technology and devices with respect to loop detectors and weigh in motion. EMX Industries Company introduce many devices focusing on loop detector, their latest vehicle detector provides functions for automatic sensitivity boost (ASB), delay, fail-safe/fail-secure and infinite and normal (5 min.) presence (EMX, Inc., 2015). From Figure 7, Kapsch WIM Company concentrated on weigh in motion technology, they can measure and provide information not only various weight features (trailer, axle, unbalanced weight etc.) but also vehicle features such as vehicle speed, vehicle class, unique vehicle number and so on. They also supports additional applications like Automatic Number Plate Recognition (ANPR) cameras, Laser-scanner Vehicle Detection and Classification, surveillance cameras, section control and speed enforcement (Kapsch, 2015).

## 2.4 MICROWAVE RADAR

### 2.4.1 Technology Description

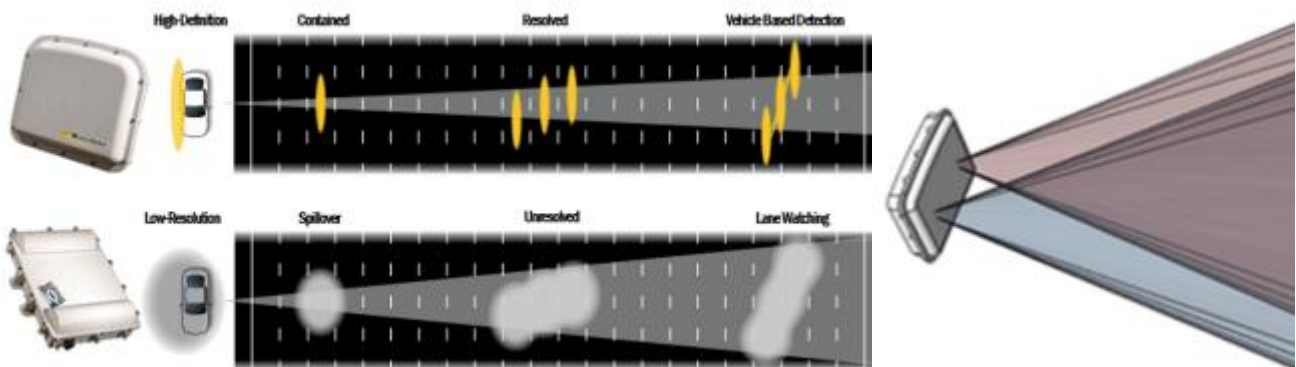
The Remote Traffic Microwave Sensor (RTMS) operates by a frequency modulated continuous wave radar unit, as it transmits electromagnetic energy at the X-band of 10.525 GHz. As shown in Figure 9, microwave radar sensors installed roadside transmit energy toward an area of the roadway from an overhead antenna (FHWA, 2007). When a vehicle enter in the area of the antenna beam, the part of the transmitted energy is reflected back towards the antenna and the receiver gets information like the vehicle's volume, lane occupancy, speed, and so on.



**Figure 9 Microwave radar operation**



**Figure 10 Wavetronix**



**Figure 11 Vehicle-Based Detection & Dual Radar (Wavetronix)**

#### 2.4.2 Applications in Performance Measures

The microwave radar device in the company of Wavetronix, SmartSensor HD, can detect performance measures such as average speed, count, occupancy, 85th percentile speed. Figure 10 and Figure 11 show operation of Wavetronix. One of their products is dual radar, which also provides a highly accurate measurement that is then used to calculate each individual vehicle's speed (Wavetronix, 2015). To be more specific, there are two categories in terms of measured quantities, per-lane interval data and per-vehicle data. The former includes volume, average speed, occupancy, classification counts, 85th percentile speed, average headway, average gap, speed bin counts, direction counts, and the latter includes speed, length, class, lane assignment, range. Furthermore, many companies including the IRD (international road dynamics) company provide applications such as following: actuated intersection control, stop-bar and mid-block detection, freeway traffic management and incident detection systems, ramp metering, queue detection and work zone safety systems, permanent and mobile traffic counting stations, traveler information and travel time prediction, enforcement of speed, and red-light violation (IRD, Inc., 2015).

In another research paper, Nale attempted to analyze traffic flow features experimenting with expressways in Beijing by using floating car data and remote traffic microwave sensors (RTMS) data. Through several tests, demonstrated and performed comparative analysis in terms of traffic flow characteristics are studied in the relationship between occupancy and density, relationship of flow-speed-density, and the relationship between FCD speed and RTMS speed (Zhao, 2009). Coifman also had much research related to RTMS measures, set out to evaluate the performance

of different kinds of four loop sensors and RTMS (Coifman, 2006). After that, he estimated aggregate data from freeway detectors between RTMS and loop detector with respect to both occupancy and flow (Coifman, 2005). It also proved that the speed measurement accuracy of RTMS can achieve up to 95%, which is higher than that of a single loop detector (Yu, 2013). This paper shows that the accuracy percentages for volume, speed, and classification measurements by using those sensing technologies with various conditions and parameters. Consequently, microwave radar sensors provided suitably accurate traffic volume and speed detection, but their performance in vehicle classification was not adequate (Yu, 2013).

## 2.5 VIDEO AND LICENSE PLATE RECOGNITION (LPR)

### 2.5.1 Technology Description

Video cameras and License Plate Recognition (LPR) are widely used in various real-life applications, such as automatic toll collection, traffic law enforcement, parking lot access control, and road traffic monitoring (FHWA, 2007; FHWA, 1998). An LPR algorithm identifies a vehicle's license plate number from a photo or photos taken by a camera. The algorithm includes a combination of a number of techniques, such as object detection, image processing, and pattern recognition. LPR is also known as automatic vehicle identification, car plate recognition, automatic number plate recognition, and optical character recognition (OCR) for cars.

Video detection and LPR technologies are provided by various vendors (Image Sensing, 2015). These technologies can provide real-time traffic measurement and data collection over a wide area. In typical applications, a number of zones are defined on the images or on the field of view of the camera. Depending on the configuration and purpose, video detection can provide per-lane presence as well as volume, occupancy, speed and classification information for the specified zones.



**Figure 12 IZ100series ALPR Camera System (INEXZAMIR)**



**Figure 13 IZ500series ALPR Camera System (INEXZAMIR)**

## 2.5.2 Applications in Performance Measures

Video image processing enables extracting useful information for traffic operations and management. On arterials, video detection is commonly used at signalized intersections instead of the traditional inductive loops to provide the needed sensing for actuated and adaptive signal operations. On freeways, surveillance cameras provide real-time feed to traffic operations centers to support incident response and other operational decisions.

For performance measurements, through video image processing, it is possible to classify vehicles, and estimate speeds, flow, and density. Accuracy of these estimates depends on the quality of the video (e.g., resolution, frame rate) and the complexity of the image processing algorithms. To obtain reliable results, image processing algorithms need to be calibrated for the particular field conditions and camera position and angle. If LPR is implemented, it is possible to estimate OD flows and travel times between two camera locations. This can be accomplished by matching the license plates observed at two locations. It is also possible to track individual vehicles as they progress within the field of view by tracking their color fingerprints or signatures (as it was done with the NGSIM datasets). Tracking individual vehicles allows estimating additional information such as lane changes, travel times, and turning movements.

Of course, the measurement of these traffic flow parameters are possible for the road segments within the field of view of the cameras. Obstructions, such as road geometry, vegetation and large trucks, may block the view and detection of traffic flow or vehicles. In addition, bad weather conditions may degrade the quality and reliability of the video data.

## 2.6 AUTOMATED VEHICLE IDENTIFICATION

### 2.6.1 Technology Description

AVI systems permit individual vehicles to be uniquely identified as they pass through a detection area. Although there are several different types of AVI systems, they all operate using the same general principles. A roadside communication unit broadcasts an interrogation signal from its antenna. When an AVI-equipped vehicle comes within range of antenna, a transponder (or tag) in the vehicle returns that vehicle's identification number to the roadside unit. The information is then transmitted to a central computer where it is processed. In most systems, the transponder and reader/antenna technology are independent of the computer system used to manage and process the vehicle identification information (IRD, Inc., 2015; Traffic Tech, 2015).

## 2.6.2 Applications in Performance Measures

Obvious applications of AVI include travel time and OD estimation between two locations where the (tag) readers are installed. Since not all vehicles carry tags or transponders, the measured attributes will be for a sample of the traffic. To the extent how well that sample represents the overall traffic is important to understand. For example, if only trucks carry transponders then, their travel times may not be generalizable to the rest of the traffic. In general, the higher the percentage of the vehicles equipped with transponders the better and more reliable the estimates become. Figure 14 and Figure 15 show widely used transponders for traffic and toll management which are produced by TagMaster and Confidex, respectively.

Eventually, faster and smoother traffic flow to reduce congestion is mainly benefit. There are some attempts in some research paper, Pérez et al. performed experiment in order to get proper speed with signals various options (Perez, 2010) and Lujaina et al. deal with how to prevent over speeding violations by using RFID and Global System for Mobile Communications (GSM) (Lujaina, 2014).



Figure 14 LR-6XL reader (TagMaster)



Figure 15 RFID (CONFIDEX)

## 2.7 PROBE VEHICLE FLEET DATA

### 2.7.1 Technology Description

Many companies carry out and provide probe vehicle data as a type of analysis, so consumers are available to get useful traffic data from named INRIX, Here.com, TOMTOM, RITIS and so



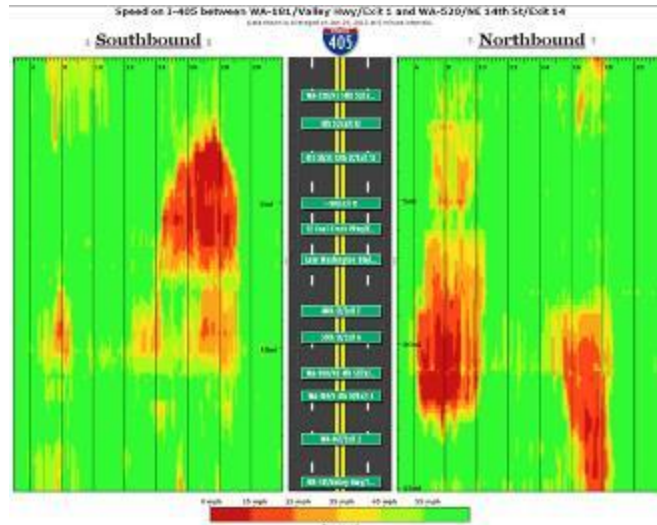
on. Additionally, there are some research fusing data between probe and loop detector data, mainly dealing with speed factor in terms of estimating and comparative analysis.

In the case of INRIX, it collects trillions of bytes of information about speed based on over 250 million real-time mobile phones and various vehicles including connected cars and fleet vehicles equipped with GPS locator devices. The data collected is processed in real-time, creating traffic speed information for major freeways, highways, and arterials across North America, as well as much of Europe, South America, and Africa. Figure 16 shows various types of probe vehicle sourcing with the information of events.



**Figure 16 INRIX Connected Driver Network**

Recently, the Here introduced new technologies like the traffic-jam warning system, called Traffic Safety Warning, and an interface, named Sensor Ingestion Interface Specification. The former systems alerts the driver with a proper time period by updating every minute to react to the event like a traffic jam ahead. Plus, the Here also works in the field of automated vehicles to determine HD Map data for manufacturers when testing their vehicles. The highly accurate mapped data of private test tracks provides these cars with a reliable navigation system to complement data collected from on-board sensors. HD Map data is also highly accurate, with accuracy levels of 10 to 20 centimeters, and it makes approximately 2.7 million changes to its global map database every day (HERE, 2015).



**Figure 17 Congestion Scan by INRIX**



**Figure 18 INRIX System Monitoring Dashboard**

## 2.7.2 Applications in Performance Measures

INRIX can archive and provide real-time traffic data based on one-minute intervals. This data can predict speed with the impact of special events before, during, and after an event as shown in Figure 17 and Figure 18. Speed positions can also be analyzed to determine the average speed for all significant roads to accurately map and quantify the extent of recurrent congestion (INRIX, 2015). By using this INRIX data, Kim carried out comparing INRIX with loop detector data in order to discover probe data features, the pros and cons by using speed performance measure (Kim, 2014). Bachmann also dealt with fusing data with loop detector data and tried to prove the developing accuracy of data to establish the true traffic speed. In order to this, the microsimulation model of a major freeway in the Greater Toronto Area was used. As a result, in order to get reliable data, we need to consider following this; the most data, depending on the technique, the number of probe vehicles and the traffic conditions (Bachmann, 2013). INRIX have descriptions that origin-destination analysis can enable the reduction of cost and complexity, if the accuracy of data

and reliability of information using GPS data are improved with a network of more than 250 million vehicles and devices (INRIX, 2015). INRIX also possible to provide how many vehicles pass a spot by time of day, day of week and actual vehicle counts on a specific day. Additionally, Abbas presents two models to calculate traffic control delay based on synchronized Bluetooth and global positioning system (GPS) probe vehicle data (Abbas, 2013).

## 2.8 INSTRUMENTED VEHICLE DATA

### 2.8.1 Technology Description

The Global Positioning System (GPS) originally developed by the Department of Defense for tracking of military ships, aircraft, and ground vehicles. However the usage of GPS has been expanded far beyond original expectations. There are civilian applications for GPS in almost every field, such as surveying, transportation, natural resource management, agriculture, etc. Probe vehicles are equipped with the GPS to monitor location, direction, and speed information (FHWA, 1998).

Many smartphones in the market integrated with a GPS. However GPS sensor has several limitations including low accuracy of GPS in urban areas with tall buildings because of the multi-path interference (Parkinson, 1996), low precision of GPS localization, and high-power consumption when the GPS is in use (Ballantyne, 2006; Raskovic, 2007; Simjee, 2005). Because of this high-power consumption, running an application that relies on the continuous use of GPS receiver depletes the phone battery quickly. However, the GPS can be turned on occasionally for a very short duration (e.g., one-two seconds) to locate the vehicle in a transportation network or a network link whereas the accelerometer data can be collected continuously to estimate when the vehicle comes to a stop.

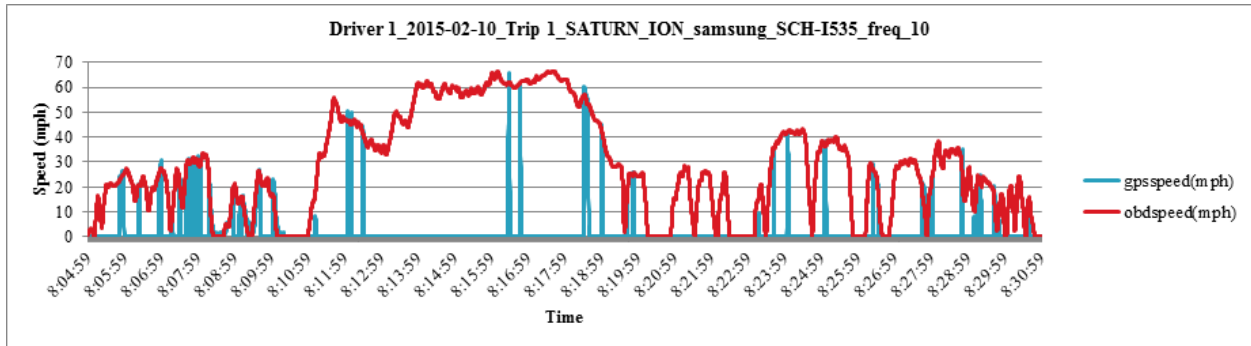
For the most part, the use of smartphones as a mobile platform for sensing traffic conditions on the roadways has been limited to the GPS data (White, 2011; Changepoint, 2015; Mohan, 2008). However, recently, data from accelerometer and other motion sensors have been utilized for the purpose of identifying the travel mode of the phone user. Researchers build models to predict various modes of travel such as driving a car, riding a bicycle, taking a bus, walking, running, riding a metro, riding on light rail, and riding a train (Manzoni, 2010; Lester, 2008; Jahangiri, 2014). Travel modes having very similar sensor data, such as driving a car and riding a bus, can increase the classification error (Jahangiri, 2014). Acceleration data are generally used together with GPS data in mode detection (Feng, 2013). Some recent studies also incorporate other sensors data such as gyroscope (Jahangiri, 2014). Using different sensor data in combination with GPS have proven to yield more accurate results in these studies.

### 2.8.2 Applications in Performance Measures

The advantages in using GPS in probe vehicles include relatively low operating costs after initial installation, and continuous collection of position and time data along the entire vehicle trajectory. The main disadvantage of GPS is occasional signals lost in urban areas due to large buildings, trees, tunnels, or parking garages. Therefore, vehicle speed can't be calculated. In Figure 19, sample raw GPS and OBD speed data for a 25-minutes trip is shown. As seen in the figure, at some points GPS speed suddenly drops to zero while vehicle is moving at high speed as indicated by the OBD speed. In Figure 20, small circles show the vehicle position for a short duration while



the vehicle waiting in a queue at a traffic signal. Although the vehicle is stationary there is variation in the raw GPS coordinates. In the picture, each small circle represents a raw GPS data point collected at 1 Hz and different colors represent different cars. Because of these potential inaccuracies in raw GPS data need to be filtered appropriate interpolation techniques should be applied to minimize the errors.



**Figure 19. Speed comparison with GPS and OBD**



**Figure 20. GPS Location error**

## 2.9 FREEWAY CAPACITY AND PRE-BREAKDOWN FLOW

### 2.9.1 Definition of Breakdown

Previous researchers have used several different definitions for the “breakdown” phenomenon. The term “breakdown” on a freeway is typically used to describe the transition event of traffic from a free flow speed (near or at the posted speed limit) to congestion. Freeway

breakdown is currently defined as the occurrence of one of the following: speed below a pre-specified threshold, variable speed threshold, or a sudden drop in speed (Aghdashi, 2015).

With respect to the first definition (speed below a pre-set threshold), the thresholds have also varied with the researcher. Graves et al. (1998) and Okamura et al. (2000) suggested that breakdown occurs when the speed is less than 30 mph during five consecutive one-minute intervals, when the speed is lower than 25 mph, or the queue length exceeds 1 km (0.62 miles) for a minimum period of 15 minute period. Brilon (2005) used a threshold speed of 70 km/h (43.5 mph) on German freeways.

With respect to the second definition (variable speed threshold), Aghdashi (2015) defines freeway breakdown as a speed reduction of more than 25% below the free flow speed for at least 15 minutes, i.e., 75% of free flow speed is variable speed threshold.

With respect to the third definition (sudden speed reduction), Persaud et al. (1998 and 2001) defined breakdown as a sharp decrease in flow and speed at a downstream ramp segment for the minimum period of 5 minutes. However, Kuhne et al. (2006) suggested that breakdown should be defined using three requirements: traffic flow remains more than 1000 veh/hr/ln, more than 10 mph speed reduction, and below the speed threshold of 46.5 mph. Recently, Elefteriadou et al. (2014) defined a breakdown as a sharp drop of minimum speed of 10 mph for at least 5 minutes.

Many researchers have also attempted to define breakdown using the concept of occupancy and features of flow rates. Elefteriadou (2003) proposed a concept of breakdown related to three different flow rates (the maximum pre-breakdown flow, the breakdown flow, and the maximum discharge flow). In addition, Elefteriadou concluded that breakdown flow is lower than the maximum flow on both pre-breakdown and discharge condition. To identify breakdown, Jia (2010) applied not only a single speed threshold, but also a density threshold based on the level of service. Amjad (2013) suggested the concept of mature or immature breakdown, formulating the change of flow rate as negative or positive, and increasing or decreasing. Recently, Song (2015) clearly defined the terms of breakdown (recurrent and non-recurrent), bottleneck, and congestion by probe vehicle data. However, the definition of breakdown is still not clear, due to the many methods used to study it, thus ongoing research into the concept of congestion in the transportation is crucial due to the fact that congestion is a serious issue.

## 2.9.2 Characteristics of Breakdown

### 2.9.2.1 Breakdown Probability Model

Research on breakdown is closely associated with freeway capacity and has contributed to the development of capacity. Recently, researchers have realized that breakdown and capacity are stochastic in nature. Elefteriadou (1995) developed a probabilistic model for breakdown at merging section, which measures the impact of ramp-vehicle clusters on the probability of breakdown activation. Lorenz and Elefteriadou (2001) analyzed speed and volume data collected at two freeway bottlenecks in Toronto to develop a breakdown probability model, and found that the highest volume appears immediately before a breakdown, and that the flow rate increases with the probability of breakdown. Kondyli (2009) suggested performance measures of lane changes and behavior, and showed that these have a significant impact on occurrence of breakdown when a ramp merging section occurs. Geistefeldt (2011) proposed that capacity design is influenced by a specific percentile of the breakdown probability, using a comparison of conventional capacity

and stochastic capacity as proof of the concept. Aghdashi (2016) suggested a parameter defined as the "Acceptable Breakdown Rate" through the use of a pre-breakdown flow rate distribution and sufficiently large breakdown samples. Shojaat (2016) provided a new stochastic performance measure called "Sustained Flow Index (SFI)" which accounts for the stochastic nature of capacity. Geistefeldt, Aghdashi, and Shojaat used the Weibull distribution to develop their breakdown probability model.

### *2.9.2.2 Identifying Breakdown*

Researchers have proposed several methods to identify breakdown and bottleneck: the plotting cumulative curve, wavelet transform method, and contour maps of the spatiotemporal status of congestion. Cassidy and Bertini (1999) demonstrated the use of transformed cumulative curves, including the vehicle arrival number and the cumulative occupancy with time, which described critical traffic features of bottleneck flows as measured by detectors and breakdown activation time information. Sarvi et al. (2007) studied the discharge flow rate under congestion and the average time in queueing status (the latter obtained by plotting the start and end times of breakdown cumulative curves with vehicle count versus time). Zheng et al. (2011), using the wavelet transform (WT) method, identified a variety of features such as the location of an active bottleneck, the activation time of congestion at an upstream sensor location, and the start and end of the transition from free flow to congestion. Song et al. (2015) suggested historic congestion contour map using INRIX data on link-based speed from vehicle probes, created the average historic congestion index (AHCI) for identifying a recurrent bottleneck, and provided useful information on congestion status via the spatiotemporal pattern.

## 2.9.3 Breakdown Prediction Model in Planning Level

### *2.9.3.1 Traffic Prediction Techniques and Statistical Model*

The two most-often used approaches for traffic prediction are: simulation models and data mining techniques. Clark (2003) used a multivariate extension of a non-parametric regression simulation model to predict the state of traffic based on the actual traffic data from London motorways. Ben's (1998) mesoscopic simulator (DynaMIT) provided traffic prediction and travel guidance using the trajectories of individual vehicles to simulate overall traffic data. Yuan et al. (2011) suggested a Cloud-based system for computing optimal driving routes and predicted future traffic conditions by using historical traffic patterns, including taxi trajectories.

Data mining techniques have been developed to predict traffic condition based on real-time datasets, often (due to the increase in accessibility of real-world data) in conjunction with statistical models. Box and Jenkins's (1970) ARIMA (Auto-Regressive Integrated Moving Average) time series model has been widely used in traffic prediction. Williams et al. (1998) developed a model which integrated ARIMA and exponential smoothing (ES) models. Ishak (2004), Van Lint (2002), and Park (1998) used a neural network (NNet) model to predict traffic parameters such as speed, travel time, and traffic flow. Pan et al. (2012) studied traffic prediction technique using the real-world data on the road network collected from LADOT (Los Angeles Department of Transportation). The use of ARIMA, NNet, ES, and HAM (Historical Average Model) models has improved (by about 70%) the accuracy of traditional short-term and long-term average speed.

Qi et al. (2013) suggested the use of a basic stochastic approach at the macroscopic level of the freeway for short-term traffic speed prediction during peak time. Qi and colleagues, using

speed transition probability data based on real world traffic data for a six-year period from a 38 mile segment of the I-4 in Orlando, showed that the transition probabilities could be categorized into cumulative negative/positive transition and expected values, and fitted using logistic and exponential models. Based on their work, Qi et al. suggested that future research should use probabilistic approaches rather than the conventional deterministic approaches.

Even though many attempts of traffic prediction are performed, these are limited to methodological approach without sufficient sample size. The conceptual traffic prediction models have studied, however, it would be important to develop a prediction model for breakdown and traffic congestion specifically, and these are as follows.

### *2.9.3.2 Breakdown Prediction Model*

Accurate prediction of breakdown is an increasingly significant research area in the field of transportation due to the numerous economic, environmental and social losses caused by congestion.

Graves, T. (1998) used CART (Classification and Regression Trees), a tree-based classifier of key factors such as speed and occupancy, to predict whether a breakdown will activate after five or ten minutes. Dong (2009) constructed a breakdown prediction model for travel time and reliability based on a transition probability matrix of Markov Chain data and real-world measurements. Dong's model predicted not only the probability of flow breakdown but also the resulting flow rate. Kondyli (2013) developed a breakdown probability model (BPM) using lifetime statistics data to identify breakdown, to assess the breakdown location, and to predict speed and flow. Kondyli used separately sections of freeway and ramp to predict breakdown probability. Kondyli recommended that any new traffic operator or system should be closely located in the merge area and suggested that speed drop is a better performance measure than other factors (e.g., occupancy, volume-occupancy correlation).

Given these circumstances, researchers have concentrated on aspects of characteristic and methodology of breakdown on prediction. However, the author has a motivation through review materials; it needs for additional prediction steps with wider perspective as a planning level based on methodologies from previous research and ITS (Intelligent Transportation System) archived data which is collected from RITIS, HERE, and PeMS.

### *2.9.4 Freeway Capacity*

Researchers have used breakdown events to develop freeway capacity estimation models with different definitions for capacity. The HCM 6<sup>th</sup> Edition (TRB, 2016) defines capacity as the flow rate at which the cumulative probability of breakdown is 15%. Further, the HCM recommends a binned Least Squares estimate of a Weibull distribution using observations of free flow volumes and pre-breakdown volumes to establish the breakdown probability distribution (Aghdashi, 2016). One major limitation to this methodology is the fitting of binned data leads to inconsistent distributions near the capacity estimate. This methodology uses speed dropping below 75% of the free flow speed to identify breakdown events.

The second methodology identified is used by the Florida DOT to measure capacity where fixed sensors are available (Kondyli, 2013). This method uses only flow observations from the pre-breakdown periods and estimates capacity as the 85<sup>th</sup> percentile of observed pre-breakdown flows. It is important to note that this percentile cannot be directly compared to the HCM

probability as the HCM distribution includes both uncongested and pre-breakdown observations. In this method, breakdown events are identified as a drop in speed of at least ten miles per hour within a ten minute period.

The final methodology was developed by Shojaat et al. (2017) based on initial research in Germany. The methodology defines the Sustained Flow Index (SFI) as the product of the breakdown probability distribution and the flow rate. The SFI was not initially developed for capacity analysis but rather for identification of optimal flow rates for planning or traffic management. In addition to the optimal value, it is also possible to identify the expected value which the authors indicate may be close to theoretical capacity. Breakdowns are identified as speed drops below 70 kilometers per hour, a threshold calibrated based on flow observations on German expressways.

### 3.0 COMPARISON OF FREEWAY CAPACITY AND PRE-BREAKDOWN FLOW ESTIMATION METHODS

#### 3.1 METHODOLOGY COMPARISON

A total of three breakdown definitions and three capacity estimation methods were identified in the literature for comparison. The methodologies are listed below based on the source in Table 2 as identified in the literature review in section 2.9.4.

**Table 2 Breakdown and Capacity Definitions**

<b>Source</b>	<b>Breakdown Definition</b>	<b>Capacity Definition</b>
Highway Capacity Manual	Speed drops below 75% of Free Flow Speed	15% Breakdown Probability from Weibull Distribution
Florida DOT (Kondyli, 2013)	Speed drops 10mph within 10 minutes	85 <sup>th</sup> Percentile Pre-breakdown Flow Rate
SFI (Shojaat, 2017)	Speed drops below 70 km/h	1) Maximum Sustained Flow Index 2) Expected Value of SFI

In order to test the differences in both breakdown definitions and capacity definitions and their impacts on the final capacity estimate, a full factorial study was performed using every combination of breakdown and capacity definition, resulting in nine final capacity estimates for each study location.

#### 3.2 DATA SOURCES

A total of 14 recurring bottlenecks were identified through speed analysis around the US. Data were collected for one year for each site as well as summary information on the bottleneck.

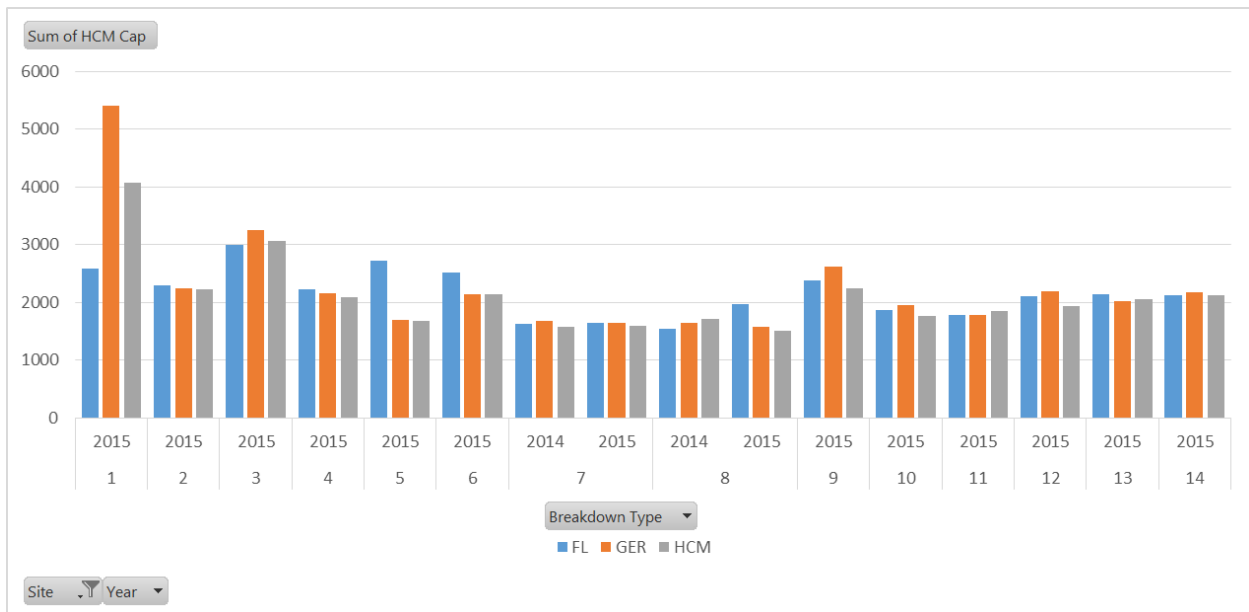
Table 3 shows the summary information for the 14 sites including the number of breakdown events observed and other geometric information. Speed and flow observations were collected from microwave radar and inductive loop sensors located upstream, within and downstream of the bottleneck to ensure the observed breakdowns were not due to spillback from other downstream events. Each 15 minute observation of speed and flow was classified into either 1) uncongested, 2) pre-breakdown, or 3) post-breakdown flow.

**Table 3 Capacity Comparison Site Descriptions**

No	State	City	Road Name	Source	Type	Data Duration (1 yr)	Breakdowns	AADT	Lanes
1	MD	Baltimore	I-695 WB	HERE	Weaving	2015.01.01 – 12.31	43	183901	4
2	MD	Baltimore	I-695 WB	HERE	Weaving	2015.01.01 – 12.31	178	191741	4
3	CA	Los Angeles	I-5 EB	PeMS	Lane Drop	2014.06.02 – 05.29	540	228000	4
4	CA	San Jose	I-880 NB	PeMS	Merge	2014.06.02 – 05.29	146	172000	3
5	FL	Miami	SR-826 NB	RITIS	Lane Drop	2014.06.02 – 05.29	309	200100	4
6	NC	Raleigh	I-440 EB	HERE	Merge	2014.06.02 – 05.29	137	80000	2
7	CA	San Diego	I-5 EB	PeMS	Merge	2015.01.01 – 12.31	282	206000	5
8	CA	Los Angeles	I-5 EB	PeMS	Merge	2015.01.01 – 12.31	962	218000	4
9	CA	Santa Ana	I-5 WB	PeMS	Merge	2015.01.01 – 12.31	549	359000	5
10	CA	Santa Ana	I-405 WB	PeMS	Merge	2015.01.01 – 12.31	446	252000	4
11	CA	San Diego	I-805 NB	PeMS	Merge	2015.01.01 – 12.31	294	181000	3
12	CA	Los Angeles	I-5 EB	PeMS	Merge	2015.01.01 – 12.31	615	218000	4
13	CA	Santa Ana	I-5 WB	PeMS	Weaving	2015.01.01 – 12.31	528	352000	5
14	CA	Santa Ana	I-405 EB	PeMS	Weaving	2015.01.01 – 12.31	283	229000	4

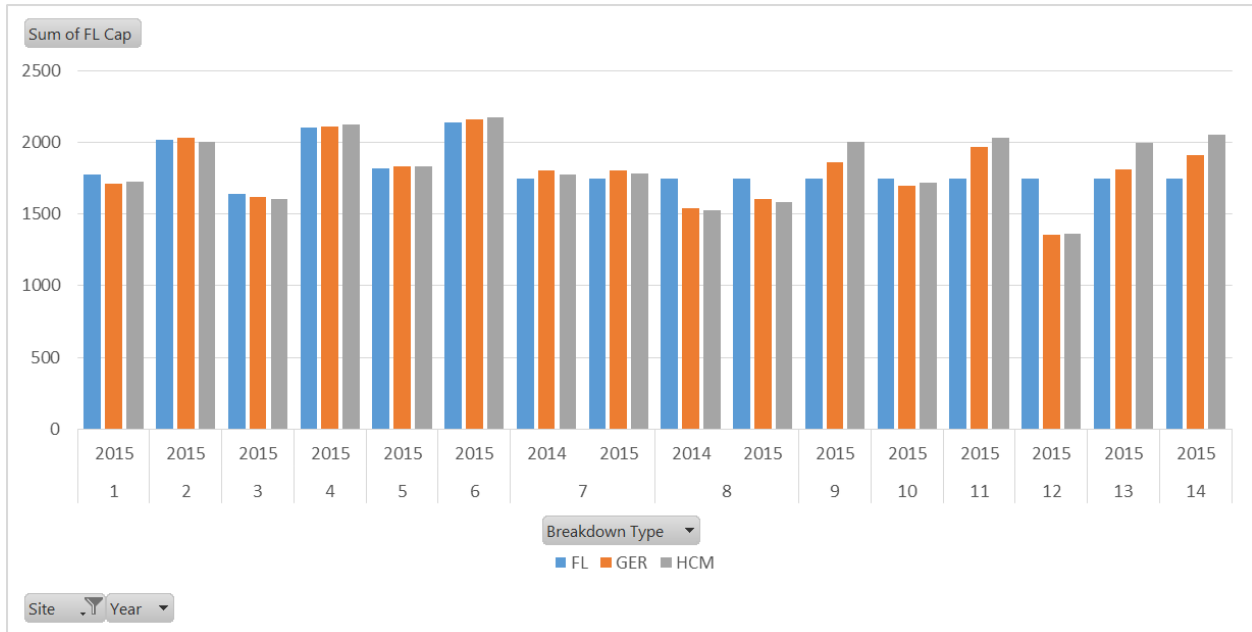
### 3.3 ANALYSIS AND RESULTS

After running the nine combinations of breakdown and capacity methods on each dataset, the output capacities were compared. Figure 21 shows the HCM capacities for the study sites based on each of the breakdown definitions with FL for the Florida definition, GER for the SFI definition adopted from German studies, and HCM. Comparing to Figure 22 for Florida method capacities and Figure 23 for SFI method capacities, it is clear that the HCM method is most prone to extreme capacity values.

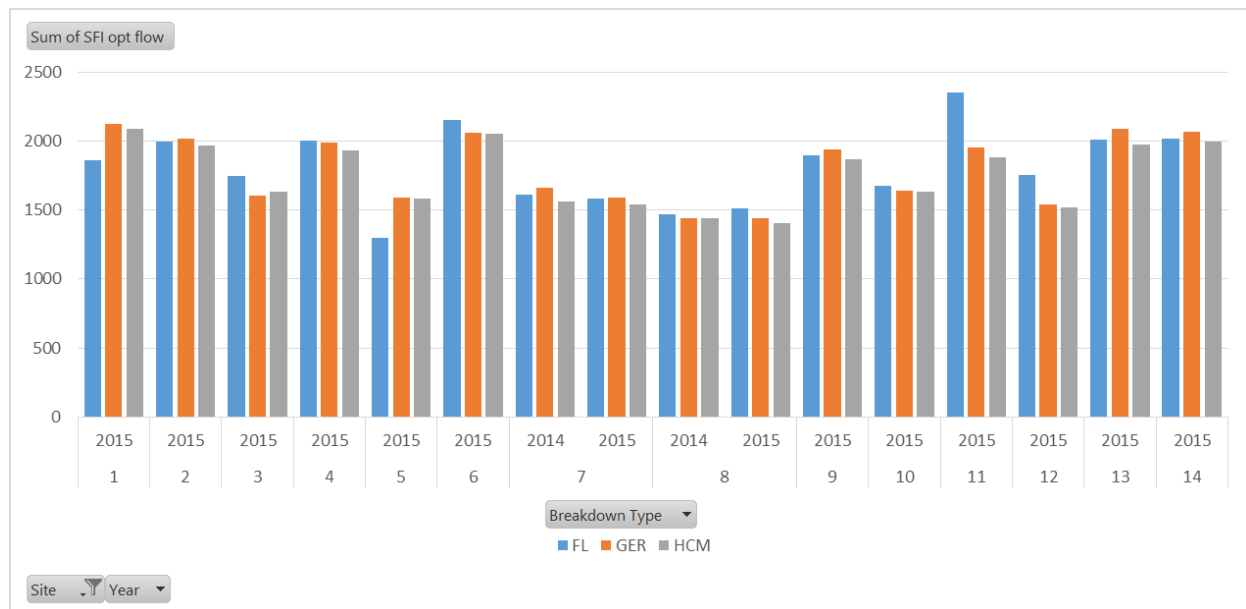


**Figure 21 HCM-based Capacity by Breakdown Type**





**Figure 22 Florida-based Capacity by Breakdown Type**



**Figure 23 SFI Optimal Flow by Breakdown Type**

Next, the resulting capacity values were run through an ANOVA to determine the significant factors out of 1) site differences, 2) breakdown method differences and 3) capacity method differences. Out of these three factors, only the site and capacity method were found to be significant as shown in

Table 4.

**Table 4 ANOVA Results for Capacity and Bottleneck Methods**

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
<b>Model</b>	19	18818310.86	990437.41	9.26	<.0001
<b>Error</b>	184	19672141.48	106913.81		
<b>Corrected Total</b>	203	38490452.35			

R-Square	Coeff Var	Root MSE	Capacity Mean
0.488909	16.58375	326.9768	1971.670

Source	DF	Anova SS	Mean Square	F Value	Pr > F
<b>Capacity_Type</b>	3	5634763.44	1878254.48	17.57	<.0001
<b>Site*Year</b>	16	13183547.43	823971.71	7.71	<.0001

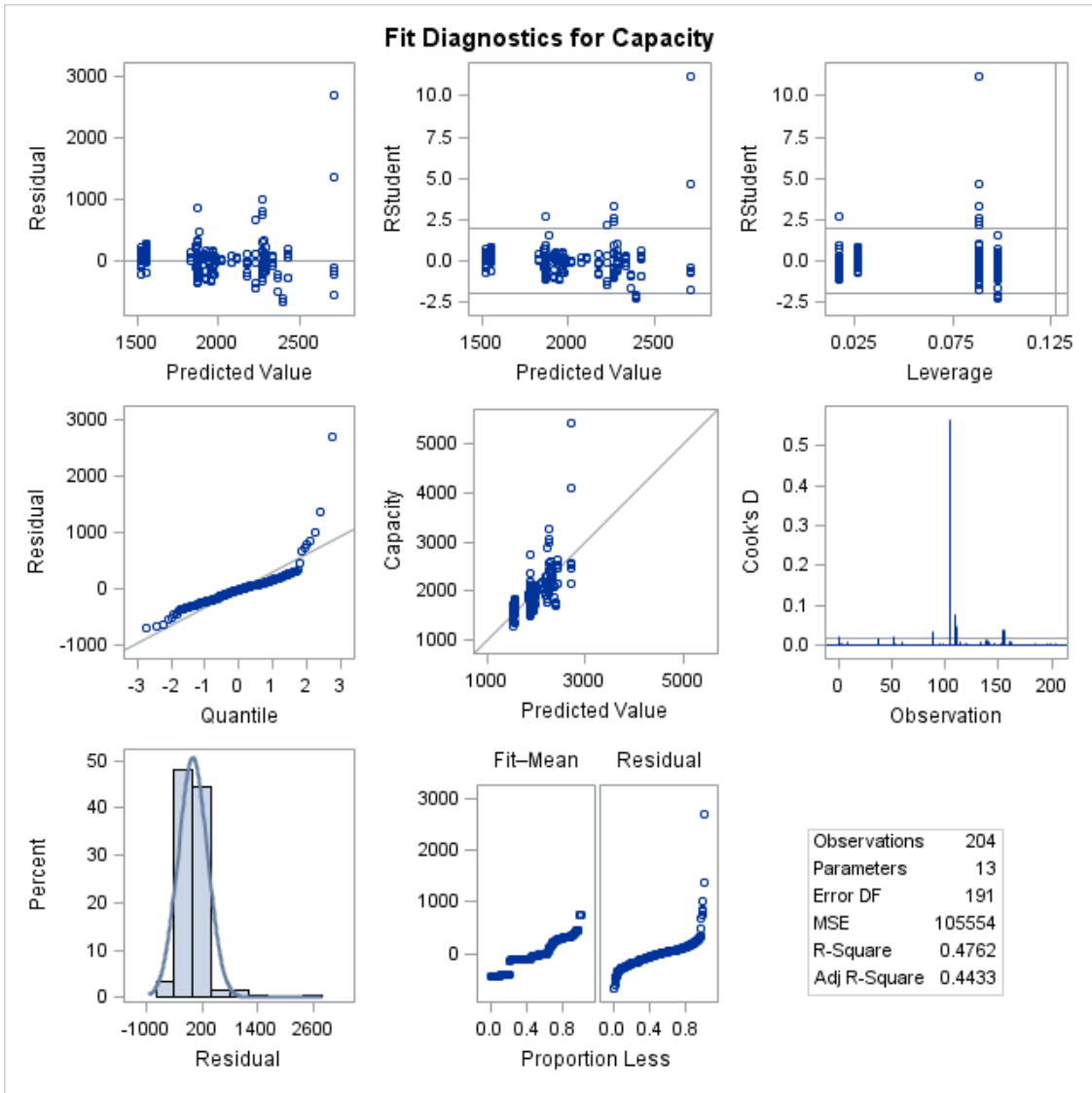
Based on the findings of the ANOVA analysis, a linear regression was fit using binary variables to represent each site, capacity method and breakdown identification method. Variables were removed using the backwards elimination method removing the least significant variable until all variables  $p < 0.1$  are removed.

Table 5 shows the resulting parameter estimates and remaining variables. In this model, sites 5, 7, 8, 10 and 12 are found to not be significantly different in capacity when accounting for the capacity estimation method. Additionally, the HCM and Expected Value of the Sustained Flow Index are found to not be significantly different. When applying this model to convert between capacity methods, we find that capacities from the Florida method are on average approximately 315 pcphpl lower than the HCM method, while the optimal SFI method capacities are on average approximately 345 pcphpl lower than the HCM method. It is important to note in the fit diagnostics shown in Figure 24 that there are a few outliers on the quantile plot indicating that more data may be needed to make a definitive comparison model.

**Table 5 Capacity Regression Parameters**

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	12	18329561	1527463	14.47	<.0001
Error	191	20160891	105554		
Corrected Total	203	38490452			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	1866.96452	42.11916	207390614	1964.77	<.0001
Site1	841.34205	100.26365	7432493	70.41	<.0001
Site2	462.63455	100.26365	2247323	21.29	<.0001
Site3	396.33847	100.26365	1649384	15.63	0.0001
Site4	414.88496	100.26365	1807360	17.12	<.0001
Site6	556.76552	100.26365	3254872	30.84	<.0001
Site9	415.71248	100.26365	1814577	17.19	<.0001
Site11	361.22141	100.26365	1370049	12.98	0.0004
Site13	401.43625	100.26365	1692086	16.03	<.0001
Site14	422.05746	100.26365	1870391	17.72	<.0001
Site15	307.61945	100.26365	993612	9.41	0.0025
C_FL	-314.49569	55.71841	3362856	31.86	<.0001
C_SFIOpt	-344.33353	55.71841	4031230	38.19	<.0001



**Figure 24 Capacity Regression Fit Diagnostics**

## **4.0 PREDICTIVE MODEL OF BREAKDOWN FOR PLANNING LEVEL APPLICATIONS**

### **4.1 BACKGROUND**

Traffic congestion leads to various societal problems not only in modern transportation systems but also with respect to economic, environmental and social issues. According to a recent Urban Mobility Report (Schrank et al. 2015), the cost of congestion, including delay and wasted fuel, is approximately \$160 billion for the 471 assessed urban areas in the United States and has been increasing since 1982 due to the fact that the number of residents and jobs are increasing commensurate with economic growth, increasing vehicle miles traveled. It is therefore becoming increasingly important to develop effective management plans and geometric improvements needed to improve freeway mobility.

Freeway congestion is usually classified into two types: recurring and non-recurring. The former is typically caused by excessive demand, or a physical constriction such as a lane drop, merge, diverge and weaving section whereas the latter is contributed by a crash or other unpredicted event, such as severe weather, and incidents (Skabardonis et al. 2003). According to data published by Cambridge Systematics (2005), recurrent congestion accounts for a large portion (approximately 40%) of all traffic congestion in the United States, and this can be improved through effective transportation management. Eventually, improved understanding and analysis of recurrent congestion is a focal point to alleviate congestion. This type of congestion typically activates at a fixed location called a “bottleneck” during peak travel periods because the demand of commuters regularly exceeds the capacity of the freeway. The bottleneck locations of a freeway are closely related to a change in the capacity or the demand of the section, or, in other words, a bottleneck has either a lower capacity than other sections or a higher demand than other sections. These conditions can be caused by a lane drop, weaving, merge or diverge segment or low posted speed limits (Darroudi, 2014). However, these are not enough to explain completely the characteristics of recurrent breakdown in the bottleneck, thus, additional influential factors must be identified. For example, several segments on the freeway may activate a recurrent breakdown due to the complex geometry near a downtown even though the section posing the bottleneck is a basic segment with low ramp density. In another case, recurrent breakdown may be caused by the correlation between the bearing of freeway and the position of the sun in peak periods, regardless of bottleneck type. Various explanatory variables on the planning level therefore may be applied to predict recurrent congestion.

Previous research in this area has focused on the issue of whether a bottleneck activates or not using short term prediction methods, with high resolution data requirements that are typically not available network-wide. Identifying the breakdown start time statistically is a more effective approach for analyzing recurrent congestion in peak time. Through the use of appropriate planning level variables and the statistical parameter values on breakdown activation time, a predictive model can provide ITS operators and freeway planners information for freeway segments where ITS data coverage is not available.

Traffic Management Center (TMC) operators are focused on the details of data and real-time information often at the scale of the individual vehicles. On the other hand, transportation planners have an interest in operations at larger temporal (daily) and spatial (facility or regional) scales.

From the planning level perspective, predictable recurrent congestion may have characteristics or regular patterns which can give rise to breakdown repeatedly. Possible planning level impacts on the start of breakdown may include the distance from downtown, type of city, bearing with respect to the rising or setting sun, speed limit and so forth. Many possible variables can support the prediction of recurrent breakdown location and activation time, while these models can be validated using massive traffic data available through Intelligent Transportation Systems (ITS), including the emerging probe data systems. Thus, in order to mitigate congestion through traffic management, it is useful to analyze variables associated with recurrent breakdown in order to deploy ITS strategies at the time when they are most needed.

The primary objective of this part for the final report is to estimate the distribution of freeway breakdown activation time using widely available planning level explanatory variables. The effort includes developing stochastic methods using the best fitting distribution and collecting massive ITS detector and probe vehicle data on fifty real world facilities in six states.

## 4.2 DATA COLLECTION

Data for model development were gathered from the freeway Performance Measurement System (PeMS), HERE (Traffic.com) and INRIX (26). A total of fifty locations were considered with 10, 37, and 3 sites extracted from PeMS, INRIX, and HERE, respectively. In TABLE 6, sensor sites comprise six states in CA, FL, NC, VA, PA, and MD. Various bottleneck types were identified, including 9 weaving sections, 26 merges, 7 diverges, and 8 lane drops. For this research, 15 minute aggregated data are collected for one year: calendar year 2015 at 46 sites and June, 2014 to May, 2015 for 4 sites. For reference, the average AADT per lane across all sites is 24,350 with a maximum value of 37,000 vehicles and a minimum of 11,000.

**TABLE 6 Summary of Site Distribution**

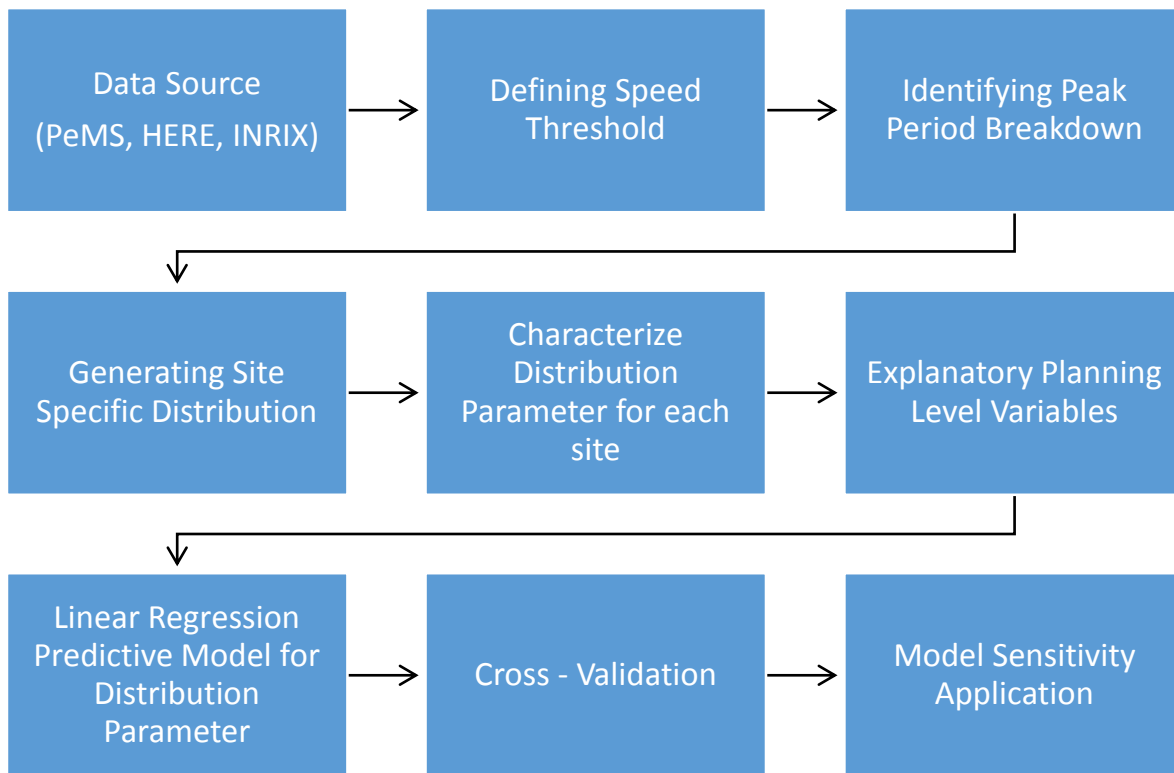
City	Number of Bottlenecks Analyzed	Facility / Direction
Los Angeles, CA	3	I-5 EB
San Jose, CA	1	I-880 NB
San Diego, CA	2	I-5 EB, I-805 NB
Santa Ana, CA	4	I-5 WB, I-405 WB, I-405 EB
Washington D.C.	9	I-395 NB, I-66 EB, I-95 NB, I-495 NB, I-270 SB, I-495 WB
Baltimore, MD	5	I-695 WB, I-695 SB, I-695 EB
Miami, FL	1	SR-826 NB
Philadelphia, PA	3	I-95 SB, I-276 WB, I-476 NB
Pittsburgh, PA	1	I-376 WB
Charlotte, NC	11	I-485 WB, I-85 EB, I-85 NB, I-77 EB, I-77 SB, I-77 WB, I-77 NB, I-485 NB, I-277 NB
Durham, NC	1	I-40 WB
Asheville, NC	1	I-240 EB
Raleigh, NC	8	I-440 EB, I-440 WB, I-540 EB, I-540 WB, I-40 WB, I-40 SB, I-40 EB
<b>Number of Sites</b>	<b>50</b>	

## 4.3 OVERALL FRAMEWORK

The overall framework is outlined in Figure 25. The first step was to collect the ITS data from three sources, i.e., PeMS, HERE, and INRIX. In the next step, the author defined the speed threshold for identifying breakdown by setting the free-flow speed. According to the data sources,



the free-flow speed can be determined using one of three approaches: 1) the average value that satisfies the low flow rate (less than 500 pcphpl) and high speed (more than 51 mph); 2) the speed limit plus 5 mph; or 3) the reference speed, then, 75% of the free flow speed is the speed threshold (HCM 2010). In step 3, the author restricted breakdown activation time as follows: a sustained breakdown period of at least 30 minutes; an activation time between 6:00 AM and 11:00 AM (morning peak time) or between 2:00 PM and 7:00 PM (evening peak time); and, only the first-occurring breakdown for each peak time. In the fourth step, the author created a graph of the breakdown activity time distribution. In the fifth step, the author used a distribution fitting software to get the best fit and characterized the parameters for each site based on the selected fitted distribution.



**Figure 25 Overall Processing of the Methodology**

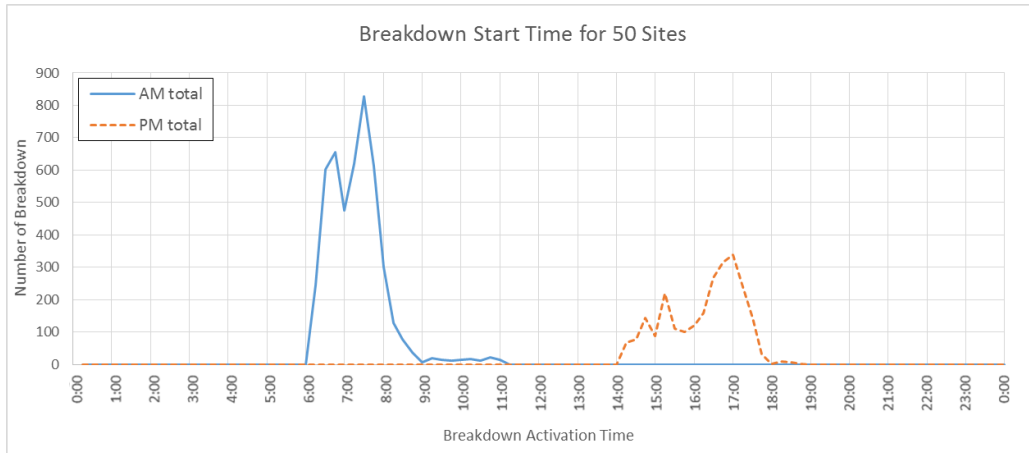
In the sixth step, the author began to construct the planning-level prediction model by considering which explanatory variables to include, such as ramp density, AADT, types of bottlenecks, bearing, weather, geometrical factor, and distance from downtown. In the seventh step, using linear regression, the author tested potential models and the different parameters using statistical analysis software. After that, the author validated and calibrated the final model to prove that the proposed explanatory variables are appropriate for planning-level predictions.

#### **4.4 IDENTIFYING PEAK PERIOD BREAKDOWN**

Defining when a breakdown occurs is key to establishing an effective breakdown prediction model. In order to identify the breakdown activation time, this paper uses variable speed threshold which is 75% of the free flow speed (4, 8, 13). The free-flow speed can be determined in three ways: the average value that is consistent with low flow rates (less than 500 pcphpl) and high speed

(more than 51 mph), 2); the speed limit plus 5 mph; or the reference speed established from the data provider.

For the purpose of identifying peak period breakdown in this study, the authors restricted the *modeling* of the breakdown activation time to the AM peak period, namely between 6 AM to 11 AM. This time boundary captured the highest frequency of breakdowns in our sites. Secondly, our focus was only on those breakdowns that lasted over 30 minutes. Finally, only the first occurring breakdown in each AM peak period is considered, signaling the start of the congestion. Therefore, breakdowns in the model are limited to one observation at each peak time for each day.

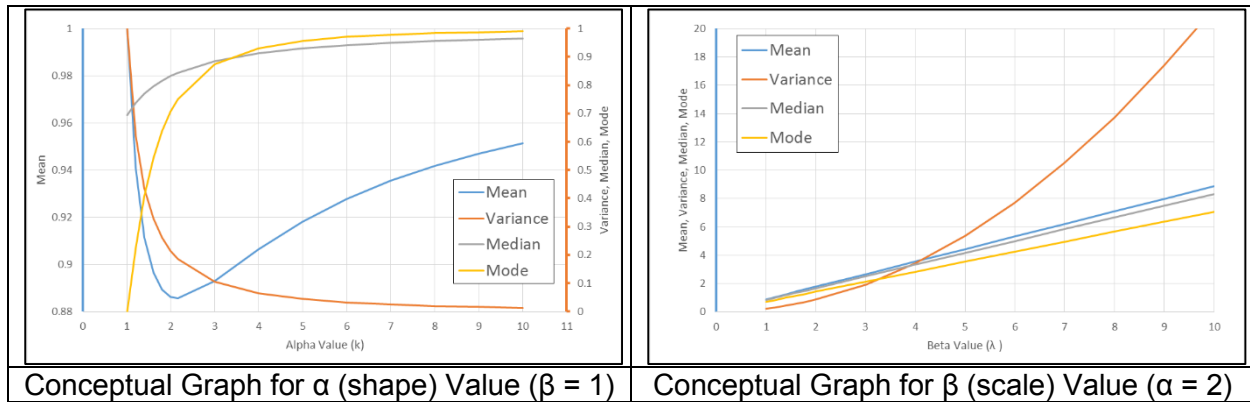


**FIGURE 26 Breakdown Activation Time Frequency for all sites**

FIGURE 26 shows the identified breakdown time distributions for the first AM and PM breakdown times at all sites (only AM data was used for modeling). It shows the overall frequency of breakdowns (total of 7,166) at each peak time and confirms that the number of AM peak time breakdowns (4,717) is greater than that of the PM peak time (2,449) in the sample. Examining similar graphs for all locations showed 29 sites with recurring congestion in the AM peak, 13 sites in the PM peak and 8 with fairly low sample size.

One of the main hypotheses in this research is that the distribution of recurring breakdown activation time can be modeled using a statistical distribution across different locations. The observations of breakdown activation time were entered into a distribution fitting software, EasyFit (27) to select the best fitting distribution. This process was scoped to distributions with only two parameters.

Goodness of fit was determined using the Kolmogorov-Smirnov (K-S) test. The Cauchy and Weibull distributions appeared to both be superior. The authors selected the Weibull distribution in part because it included only positive values, whereas the Cauchy distribution also includes negative values. In addition, when only the AM peak period is considered, the Weibull ranked higher than the Cauchy distribution.



**FIGURE 27 Theoretical Feature in Weibull Distribution**

In the Weibull distribution, alpha and beta represent the shape and scale parameters, respectively. Since our primary interest is in predicting the most likely activation time, the focus is on assessing how the value of the parameters affect the movement of the distribution mode. FIGURE 27 depicts various distribution statistics, and how they are sensitive to the shape and scale parameters. It is evident that the mode increases significantly with the shape parameter at low values, then stabilizes at high values. The mode also increases in a linear fashion with the scale parameter beta. This implies that any planning level variables that reduce either alpha or beta will tend to shift the mode of the distribution to the left, thus predicting an earlier most likely breakdown time.

#### 4.5 PLANNING LEVEL VARIABLES

The fifty recurrent breakdown activation locations identified have a wide variety of geometric and design characteristics. Features that may impact AM peak breakdown activation time include free flow speed, AADT, weather conditions, bottleneck type, ramp density, distance from downtown, bearing, sunlight, and unusual geometrics. All those variables were considered in the model development.

Free flow speed is an important variable that is used by the HCM to estimate freeway capacity (8). In this study, free flow speed values ranged from 49-70 mph, with an average of 64 mph across all fifty sites. Bottleneck type is classified by the HCM segment type, namely weaving, on-ramp, off-ramp, and lane drop. Many research efforts related to breakdown have focused on on-ramp bottlenecks, but rarely on weaving and lane drop sections. Ramp density can also contribute to recurring congestion due to the many lane changes around on and off ramps and due to the demand surges that may occur at merge locations. It is defined as the number of ramps located within 6 miles, 3 miles upstream and downstream from the sensor location.

AADT data for each site is collected from ArcGIS shape files available from each state. The value “AADT per lane” is used to account for traffic intensity. Distance from downtown may be an important factor to analyze and predict breakdown activation time because morning recurrent congestion is usually influenced by commuters traveling towards downtown. It is measured as the distance between the center point of downtown and the sensor location as the crow flies. Bearing represents the compass direction of travel on the freeway. It is a proxy for the effect of sun glare which may impact driver vision, resulting in excess braking and breakdown at lower flow rates.

Similarly, the presence of sunlight in a driver’s peripheral vision is estimated at each peak time using the bearing of the freeway segment. It is assumed that degree of the sunlight of sunrise and sunset are 90° and 270°, respectively. This factor is modeled as a binary variable indicating if due East or West are within a 60 degree arc centered on the bearing of the freeway segment (28). Principal geometric effects are also identified that may have a significant impact on breakdown activation time. Sharp horizontal curves and vertical grades exceeding 5%, are entered as binary variables.

Lastly, weather data is sourced from project SHRP2-L08 (29). That study categorized 11 weather event types identified as capacity impacting in the HCM 2010 (8). It includes weather event data for 91 cities in the United States for each hour of each month from 2001 to 2010 based on data reported by the National Weather Service (NWS). This study used the average (long term) value of the probability of inclement weather at each site, in the AM peak period analyzed (6 - 11 AM).

## 4.6 MODEL DEVELOPMENT

The team applied two linear regression models to predict the Weibull parameters that best fit the breakdown time distribution. Dependent variables ( $Y$ ) are the site-specific alpha and beta parameter values from a best fit Weibull distribution. The independent variables are characteristic ( $X_i$ ) at each site.

The statistical model is in the form:

$$y = \beta_0 + \beta_1x_1 + \dots + \beta_nx_n \quad \text{(Equation 1)}$$

Where:

$\beta_0$  = intercept term

$\beta_i$  = coefficient for explanatory variable (i)

n = the number of effective variables in each dataset

The parameters are characterized for each site based on the selected fitted distribution. The nine variables introduced earlier are included in the regression along with the squares of each numeric variable to account for any possible quadratic relationships. The Statistical Analysis System, or SAS (30) was used to perform the linear regression. Significant variables were identified using the backward selection algorithm (recursively removing insignificant variables) with a significance threshold of 0.05 required to remain in the model.

## 4.7 CALIBRATION AND VALIDATION SETS

The authors considered that the predictive model may perform well on sites where it is locally calibrated, but not be as robust when applied to cities where data were not used in the calibrated model. Thus, two different models were fit. The first model included sites in cities that has more than 3 sites in the overall dataset. The corresponding validation set included only sites that had cities represented in the calibrated model. The second model uses a validation set where all withheld sites are from cities not represented in the calibration dataset. This second approach tests the model transferability to sites outside the calibration domain. The dataset description is shown in TABLE 7.

**TABLE 7 Calibration and Validation Dataset**

Set Type	Name	Count	Description
Calibration Dataset	Repeated Site Removed	23	All AM sites(29) – Repeated Site Set(6) = 23
	Excluded Site Removed	23	All AM sites(29) – Excluded Site Set(6) = 23
Validation Dataset	Repeated Site	6	Choose randomly one site from repeated locations (Charlotte, Los Angeles, Philadelphia, Raleigh, Santa Ana, and Washington D.C.) and combine them.
	Excluded Site	6	All unique locations (Asheville, Baltimore(2), Pittsburgh, San Diego, San Jose)

## 4.8 CALIBRATION RESULTS

The results of the linear regression models are shown in

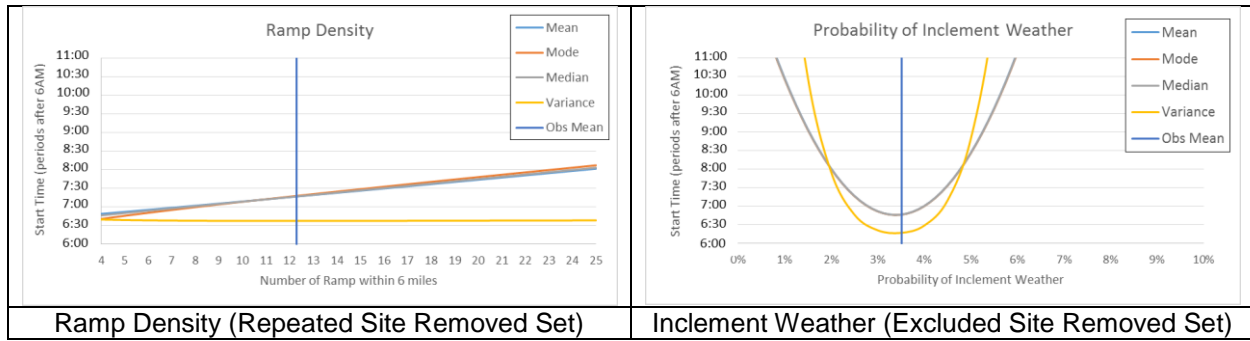
Table 8. Ramp density (the number of ramps within six miles) was significantly associated with breakdown activation time in the AM peak time. AADT per lane was also associated with congestion onset, except for the model run using the second calibration (unique sites held for validation). Both free flow speed and inclement weather have a significant impact on breakdown occurrence (beta model for both calibration datasets). Finally, the author also identified merging segments, weaving bottleneck type, and having “sun ahead” conditions as significant variables in each calibration dataset. Sensitivity analysis is applied to confirm the impact of varying a specific variable and find the range of feasible values. In each of the graphs that follow, trend curves of mean, mode, median, and variance for the breakdown start time are observed by changing variable value for each calibration set.

**Table 8 Calibrated Linear Regression Models**

Calibration Set type	$\alpha, \beta$ (adj. $R^2$ )	Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Repeated Site Removed Set	$\alpha$ , shape (0.3796)	Intercept	5.89323	1.33247	40.4012	19.56	0.0003
		Ramp Density	0.17975	0.06679	14.9594	7.24	0.0145
		AADT per lane*	-2.1162	5.9E-05	26.1997	12.69	0.0021
		Merge	1.81138	0.64698	16.1898	7.84	0.0114
	$\beta$ , scale (0.6400)	Intercept	144.298	39.17445	19.1086	13.57	0.002
		Free Flow Speed	-3.95856	1.23204	14.5391	10.32	0.0054
		Free Flow Speed squared	0.03125	0.00977	14.4075	10.23	0.0056
		Ramp Density	0.24145	0.0555	26.6572	18.93	0.0005
		AADT per lane*	-10.9	0.0003343	14.8394	10.54	0.0051
		AADT per lane squared*	1.7641	6.54E-09	10.2561	7.28	0.0158
Excluded Site Removed Set	$\alpha$ , shape (0.4798)	Inclement Weather squared**	-0.0761	224.23617	16.2286	11.52	0.0037
		Intercept	19.321	4.56112	23.03262	17.94	0.0006
		Ramp Density	0.11955	0.05093	7.07188	5.51	0.0321
		AADT per lane*	-8.9858	0.00032918	9.56446	7.45	0.0148
		AADT per lane squared*	1.3879	6.19E-09	6.46094	5.03	0.0394
		Degree from East***	-0.6533	0.02579	8.23531	6.42	0.0222
		Degree from East squared***	0.0249	0.00011129	6.4529	5.03	0.0395
	$\beta$ , scale (0.6575)	Sun Ahead	-2.20261	0.955	6.82764	5.32	0.0348
		Intercept	158.49263	34.34528	26.00423	21.3	0.0003
		Free Flow Speed	-3.74592	1.00607	16.92872	13.86	0.0018
		Free Flow Speed squared	0.02833	0.00798	15.38548	12.6	0.0027
		Ramp Density	0.13874	0.05337	8.25174	6.76	0.0194
		Inclement Weather**	-19.59306	451.43236	23.00272	18.84	0.0005
		Inclement Weather squared**	2.8861	6536.74428	23.80447	19.49	0.0004
Weaving	-2.3406	0.80896	10.22244	8.37	0.0106		

\*AADT in ten thousands, \*\*% Inclement Weather, \*\*\*Degree from East in tens of degrees

In the Repeated Site Removed, the breakdown start time increased with ramp density (FIGURE 28) but decreases (breakdown is earlier) with inclement weather, AADT per lane, and free flow speed. Start time is later in a merge section, but the difference between merge and non-merge sections is very small. In the Excluded Site Removed, all central tendencies of breakdown activation time occur later by increasing ramp density. For both AADT per lane and degree from East, all values except the variance have a zero slope (flat line), implying that change in these variables does not affect the central tendencies of start time. However, according to this model, breakdown is activated earlier by increasing free flow speed, and interruption of a driver's sight by sunlight or the presence of a weaving section. These effects result in decreased values of the mean, the mode, and the median, implying an earlier activation start time. Inclement weather plot (FIGURE 28) reveals an unexpected phenomenon, with all values have a U shape. A breakdown is activated much later compared to any actual segment breakdown start time (after 8:30 AM) when the inclement weather probability is low or high (less than 2% or more than 5%). This points to a core deficiency in the implementation of the second model for any practical purpose.



**FIGURE 28 Modeled Breakdown Start Time Sensitivities**

## 4.9 VALIDATION COMPARISONS

In order to test the goodness of fit for the two validation sets, the model-predicted Weibull parameters are estimated using the statistical model and compared to their fitted site-specific counterparts. The results are shown in Table 9. It is evident that the set selected from repeated cities has largely small differences between the best fit and predicted parameters. On the other hand, the difference for predicted Weibull parameters in the Excluded Set is excessive indicating poor prediction. Specifically, the predicted beta values in the Excluded Set tends to be significantly higher than the best fit parameters. Through these two validation results, it is clear that the predicted result from the Repeated Set fits much better than the Excluded Set.

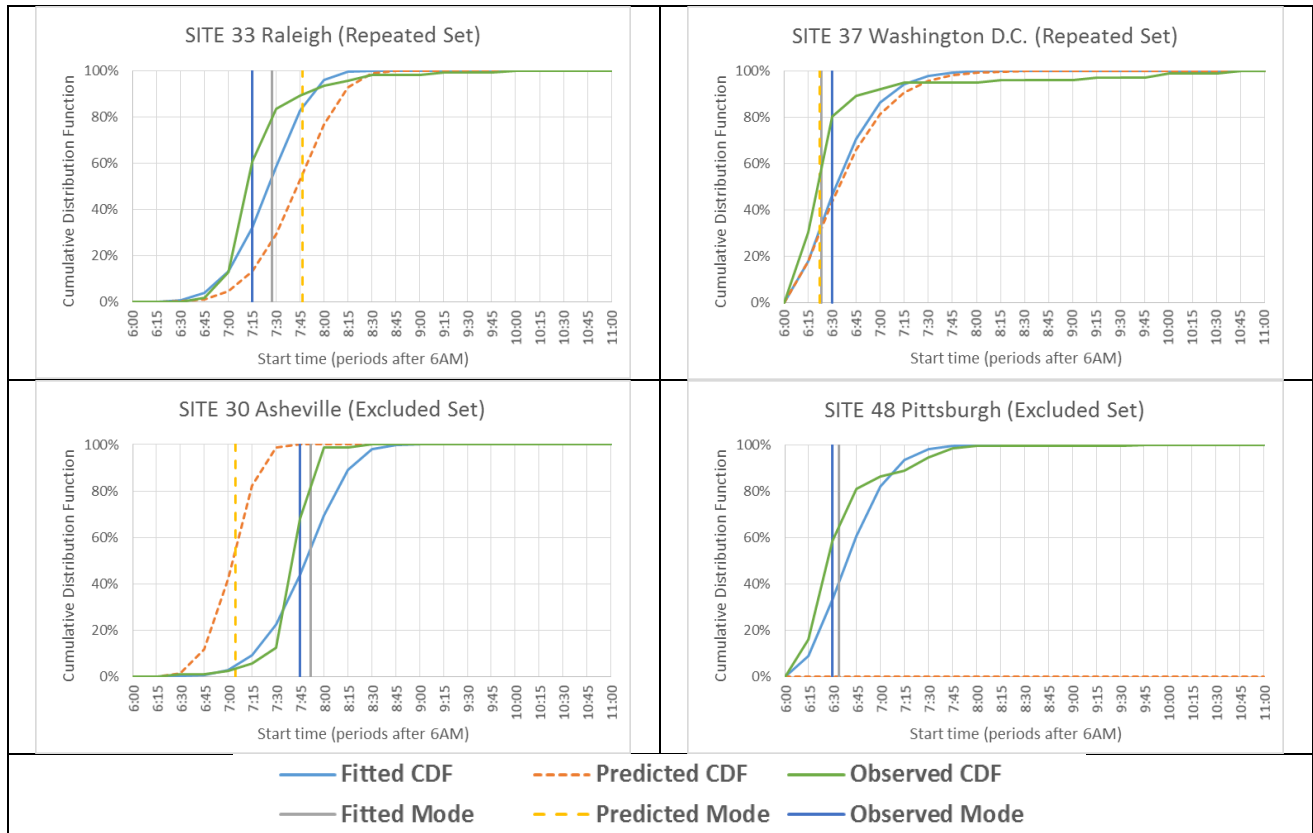
**Table 9 Comparison of Site-Specific Fitted vs. Model Predicted Parameter Values**

Validation Set type	Site No.	City	Site-Specific Fitted		Model Predicted		Absolute Difference	
			Alpha	Beta	Alpha	Beta	Alpha	Beta
Repeated Set	21	Charlotte	6.611	6.785	4.034	6.249	2.577	0.536
	8	Los Angeles	1.184	5.179	4.455	5.343	3.271	0.165
	49	Philadelphia	2.758	4.957	4.314	4.101	1.555	0.856
	33	Raleigh	4.520	6.178	4.960	7.425	0.440	1.248
	13	Santa Ana	1.943	3.233	1.860	7.587	0.083	4.354
	37	Washington D.C.	1.672	2.649	1.546	2.859	0.126	0.210
Excluded Set	30	Asheville	5.374	7.754	5.180	4.497	0.194	3.256
	42	Baltimore	2.598	5.641	3.418	10.281	0.820	4.641
	43	Baltimore	6.411	7.491	4.756	12.099	1.656	4.608
	48	Pittsburgh	2.092	3.097	3.745	176.868	1.653	173.771
	11	San Diego	3.012	3.930	2.277	17.173	0.734	13.242
	4	San Jose	8.448	8.370	3.267	18.244	5.181	9.874

In order to compare the best fit to predicted distribution of breakdown activation time, two sites are considered in each selection procedure from Figure 29. These are plotted using the fitted and predicted alpha and beta values from Table 9. The x-axis represents the time series of AM peak time (6 - 11AM); the y-axis shows the cumulative probability of a breakdown activation time based on the Weibull distribution. Additionally, the empirical distributions of activation times are



plotted in green. The vertical lines show the modes of the empirical, site-specific fitted and model predicted modes of the three distributions.



**Figure 29 Validation Fitting CDF Plots Result**

Firstly, the difference between fitted and predicted mode is about 15 minutes the Raleigh using the Repeated Set approach. The 80th percentile breakdown time is 8:00 AM for this facility. Both examples using the Repeated Site selection procedure tended to produce good matches between the fitted and predicted distributions, though they both deviated from the empirical distribution.

The two validation sites using the Excluded Site selection procedure had very poor predictions of breakdown activation time. The Asheville site shows that the predicted distribution shifted approximately one hour earlier than the observed and best fit Weibull distribution. This result is effected by the large difference in the beta parameter value between the best fit and predicted distribution (see Table 9). This low predicted beta value may be due to the fact that the bottleneck type at this site is a weaving section, which has a negative parameter estimate in the statistical model (see

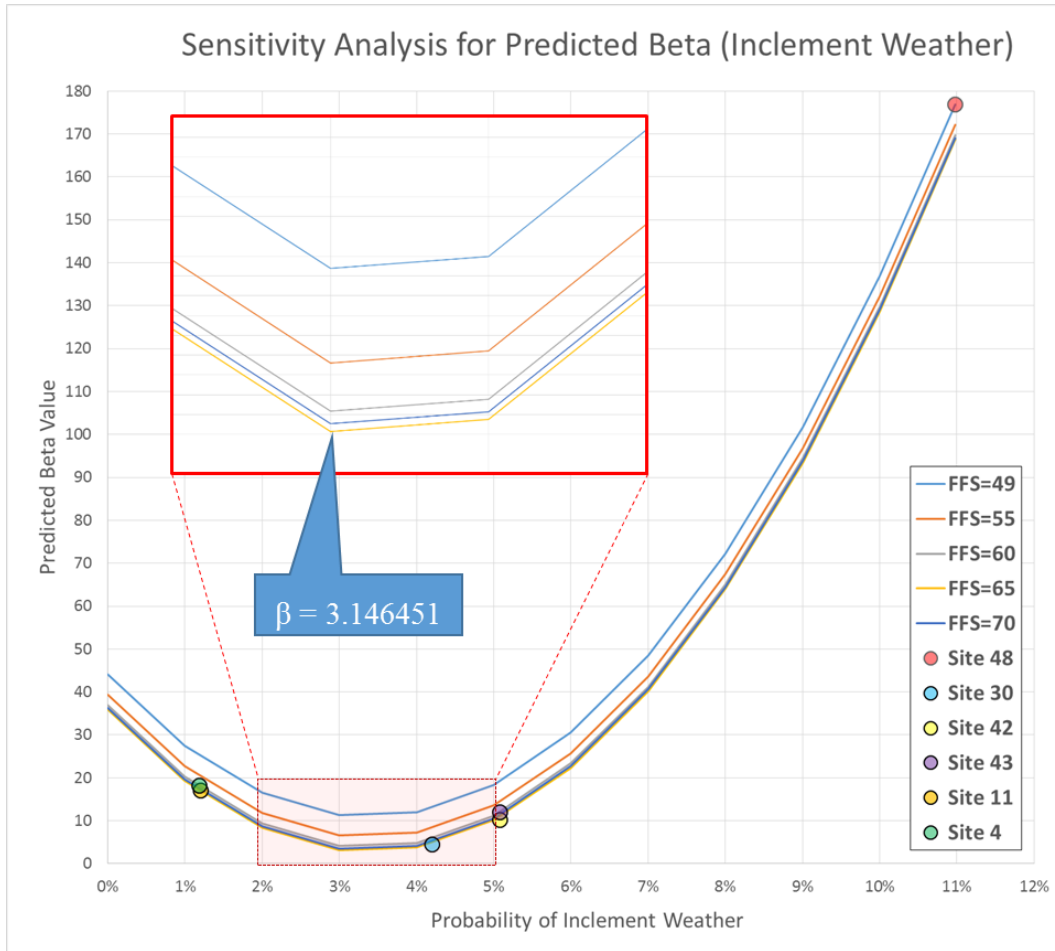
Table 8).

The Pittsburgh site plot shows another example of poor prediction caused by an excessively large predicted beta value. The poor beta prediction is a result of the Pittsburgh site having both free flow speed and inclement weather frequency that are beyond the range of values used in the calibration set. The free flow speed is 49 mph, the lowest value in the entire set of 29 total AM sites. Secondly, the average percentage of inclement weather in all studied site is only 3 percent while Pittsburgh value exceeds 10 percent. Inclement weather has a very high impact on the beta value as shown in the sensitivity analysis. The outlier inclement weather probability causes the predicted distribution to shift entirely outside of the time period of interest.

The author attempted to confirm what variables might be behind poor prediction results by conducting a sensitivity analysis in which free flow speed and inclement weather probability values were changed while other variables were held constant, in order to reduce the difference between the fitted and the predicted beta. Table 10 shows that the predicted beta value (3.146) closest to the fitted beta value (3.09) is achieved when inclement weather probability is 3% and free flow speed is 65 mph (similar to the average values of all 29 sites, 3.74% and 63.14 mph, respectively). The author concludes that the Pittsburgh site shows a poor predictive result due to low free flow speed and high inclement weather probability.

**Table 10 Effects on Predicted Beta of Simultaneously Changing Free Flow Speed (FFS) and Inclement Weather**

Inclement Weather Probability	FFS=49	FFS=55	FFS=60	<b>FFS=65</b>	FFS=70
11%	176.8677	172.0992	169.6593	168.636	169.0291
10%	136.8618	132.0933	129.6535	128.6301	129.0233
9%	101.619	96.85047	94.41062	93.38727	93.78042
8%	72.14837	67.37984	64.93999	63.91664	64.30979
7%	48.44994	43.6814	41.24155	40.2182	40.61135
6%	30.5237	25.75516	23.31531	22.29196	22.68511
5%	18.36966	13.60113	11.16128	10.13793	10.53108
4%	11.98783	7.219288	4.779438	3.756088	4.149238
<b>3%</b>	11.37819	6.609651	4.169801	<b>3.146451</b>	3.539601
2%	16.54075	11.77221	9.332364	8.309014	8.702164
1%	27.47552	22.70698	20.26713	19.24378	19.63693
0%	44.18248	39.41394	36.97409	35.95074	36.34389



**Figure 30 Sensitivity Analysis for Predicted Beta (Inclement weather in Excluded Set)**

Figure 30 shows the results of sensitivity analysis (Table 10). Each single dot represents the predicted beta value for a given site's probability of inclement weather. Outliers for inclement weather probability include two sites in California (Site 4 and 11), with an inclement weather probability of about 1% and a high predicted beta value (more than 15), and two sites in Baltimore (Site 42 and 43) with a high probability (about 5%) and a high predicted beta value (more than 10). To summarize, with the exception of site 30 (Asheville), it is difficult to predict congestion onset using the Excluded Set due to the inclement weather probability.

## **5.0 CONCLUSIONS AND FUTURE WORK**

### **5.1 SUMMARY OF RESEARCH AND FINDINGS**

The research team identified seven major technologies available for traffic data collection and which of the traditional HCM and simulation performance measures could be measured by each technology. Of these performance measures, the team found three primary methods for estimating freeway capacity across the country. A nationwide study of microwave radar and inductive loop sensor data found that the underlying breakdown method used by each of the three methods was not individually significant in differences in the final capacity estimate. Among the capacity estimation methods, the Florida DOT method using the 85<sup>th</sup> percentile pre-breakdown flow rate estimated capacities approximately 315 pcphpl lower than the HCM, while the optimal Sustained Flow Index method estimated capacities approximately 345 pcphpl lower than the HCM.

The research team also developed a congestion onset prediction model at the planning level that considers not only significant variables for breakdown activation time but also characterizing the breakdown activation time distribution. A set of locally calibrated models were shown to accurately predict breakdown activation time within 7 minutes. Interchange ramp density, inclement weather probability, AADT per lane, and free flow speed were found to significantly impact the breakdown activation time.

### **5.2 RECOMMENDATIONS FOR FUTURE RESEARCH**

The research team recommends that continued research is needed to address the collection or estimation methodologies for many HCM performance measures as there are many limitations due to the type of deployment or sampling of traffic. Ongoing research such as NCHRP 15-57 are not only developing new HCM methodology but are also focused on how emerging datasets can address these methods. The team also recommends that capacity estimation methods continue to be compared as the findings in this report are limited due to the sample size available for a nationwide study. More clarity on capacity differences are necessary as states begin to collect more data and need to compare to existing HCM or other methods. Finally, the researchers recommend further investigation into predicting breakdowns as the findings have implications for planning, operations and management of freeway facilities.

## 6.0 REFERENCES

- Abbas, Montasir, et al. "Microscopic modeling of control delay at signalized intersections based on Bluetooth data." *Journal of Intelligent Transportation Systems* 17.2 (2013): 110-122.
- Aghdashi, Seyedbehzad, et al. "Freeway Capacity: Theoretical Construct and Field Estimation Method." *Transportation Research Board 95th Annual Meeting*. No. 16-5813. 2016.
- Aghdashi, Seyedbehzad, et al. "WHITE PAPER ON FREEWAY SEGMENT CAPACITY FOR THE HCM 2010 MAJOR UPDATE" 2015.
- Anagnostopoulos, C.; Alexandropoulos, T.; Loumos, V.; Kayafas, E., "Intelligent traffic management through MPEG-7 vehicle flow surveillance," in *Modern Computing, 2006. JVA '06. IEEE John Vincent Atanasoff 2006 International Symposium on* , vol., no., pp.202-207, 3-6 Oct. 2006 doi: 10.1109/JVA.2006.30
- Ballantyne, W., G. B. Turetzky, G. Slimak and J. Shewfelt. Achieving Low Energy-Per-Fix in Cell Phone. *GPS World*, Vol. 17, No. 7, 2006, pp. 24-32.
- Bachmann, Chris, et al. "Fusing a bluetooth traffic monitoring system with loop detector data for improved freeway traffic speed estimation." *Journal of Intelligent Transportation Systems* 17.2 (2013): 152-164.
- Bachmann, Chris, et al. "A comparative assessment of multi-sensor data fusion techniques for freeway traffic speed estimation using microsimulation modeling." *Transportation research part C: emerging technologies* 26 (2013): 33-48.
- Ben-Akiva, Moshe, et al. "DynaMIT: a simulation-based system for traffic prediction." *DACCORS short term forecasting workshop*, The Netherlands. 1998.
- BlueMAC Analytics. BlueMAC Services. (2015) <http://www.mybluemac.com/?services=services>
- Bluetooth SIG. Bluetooth Fast Facts. (2015) <http://www.bluetooth.com/Pages/Fast-Facts.aspx>
- Brilon, Werner, Justin Geistefeldt, and Matthias Regler. "Reliability of freeway traffic flow: a stochastic concept of capacity." *Proceedings of the 16th International symposium on transportation and traffic theory*. Vol. 125143. Suite 1C, Joseph's Well, Hanover Walk, Leeds LS3 1AB, UK: Maney Publishing, 2005.
- Caltrans, CADOT website (2016), <http://www.dot.ca.gov/hq/traffops/mpr/source.html>
- Cambridge Systematics, Inc. "Traffic Congestion and Reliability" *Trends and Advanced Strategies for Congestion Mitigation*. Final Report. (2005)
- Cassidy, Michael J., and Robert L. Bertini. "Some traffic features at freeway bottlenecks." *Transportation Research Part B: Methodological* 33.1 (1999): 25-42.
- Changepoint: An R Package for Changepoint Analysis: R Package Version 2.0.1.
- Chen, Chao. "Freeway performance measurement system (PeMS)." *California Partners for Advanced Transit and Highways (PATH)* (2003).

Clark, Stephen. "Traffic prediction using multivariate nonparametric regression." *Journal of transportation engineering* 129.2 (2003): 161-168.

Click, Steven M., and Travis Lloyd. "Applicability of bluetooth data collection methods for collecting traffic operations data on rural freeways." *Transportation Research Board 91st Annual Meeting*. No. 12-1387. 2012.

Coifman, Benjamin. "Vehicle level evaluation of loop detectors and the remote traffic microwave sensor." *Journal of transportation engineering* 132.3 (2006): 213-226.

Coifman, Benjamin. "Freeway detector assessment: aggregate data from remote traffic microwave sensor." *Transportation Research Record: Journal of the Transportation Research Board* 1917 (2005): 149-163.

Coifman, B., & Cassidy, M. (2002). Vehicle reidentification and travel time measurement on congested freeways. *Transportation Research Part A: Policy and Practice*, 36(10), 899-917.

Confidex. Automatic Vehicle Identification. (2015)  
<http://www.confidex.com/markets/traffic-management/automatic-vehicle-identification>

Dehman, Amjad. "Breakdown Maturity Phenomenon at Wisconsin Freeway Bottlenecks." *Transportation Research Record: Journal of the Transportation Research Board* 2395 (2013): 1-11.

Dong, Jing, and Hani Mahmassani. "Flow breakdown and travel time reliability." *Transportation Research Record: Journal of the Transportation Research Board* 2124 (2009): 203-212.

Elefteriadou, Lily, Alexandra Kondyli, and Bryan St George. *Estimation of Capacities on Florida Freeways*. No. 3/11/2013-9/30/2014. 2014.

Elefteriadou, Lily, and Ponlathap Lertworawanich. "Defining, measuring and estimating freeway capacity." *Washington: Transportation Research Board Meeting*. 2003.

Elefteriadou, Lily, Roger P. Roess, and William R. McShane. "Probabilistic nature of breakdown at freeway merge junctions." *Transportation Research Record* 1484 (1995): 80-89.

EMX, Inc. Vehicle Loop Detector. (2015)  
[http://www.emxinc.com/vehicle\\_loop\\_detector\\_ultra\\_plg.html](http://www.emxinc.com/vehicle_loop_detector_ultra_plg.html)

Engmann, Sonja, and Denis Cousineau. "Comparing distributions: the two-sample Anderson-Darling test as an alternative to the Kolmogorov-Smirnoff test." *Journal of Applied Quantitative Methods* 6.3 (2011): 1-17.

Feng, T. and H. J. P. Timmermans. *Transportation Mode Recognition Using Gps and Accelerometer Data*. *Transportation Research Part C: Emerging Technologies*, Vol. 37, No. 0, 2013, pp. 118-130.

FHWA. *Travel time data collection handbook*, ITS Probe Vehicle Techniques. (1998)  
<http://trid.trb.org/view.aspx?id=497690>

FHWA *A Summary of Vehicle Detection and Surveillance Technologies use in Intelligent Transportation Systems*. (2007)  
<https://www.fhwa.dot.gov/policyinformation/pubs/vdstits2007/05.cfm>

G. Box and G. Jenkins. Time series analysis: Forecasting and control. San Francisco: Holden-Day, 1970.

Geistefeldt, J. Verkehrsablauf und Verkehrssicherheit auf Autobahnen mit vierstreifigen Richtungsfahrbahnen (Traffic Flow and Road Safety on Freeways with Four-Lane Carriageways). Doctoral Thesis, No. 30. Institute for Transportation and Traffic Engineering, Ruhr-University Bochum, 2007.

Geistefeldt, J., and W. Brilon, A comparative assessment of stochastic capacity estimation methods. In Transportation and Traffic Theory 2009: Golden Jubilee (pp. 583-602). 2009, Springer US.

Geistefeldt, Justin. "Empirical relation between stochastic capacities and capacities obtained from the speed-flow diagram." Transportation Research Circular E-C149, Greenshields Symp. Vol. 75. 2011.

GIS map info (2016) <http://www.igismap.com/formula-to-find-bearing-or-heading-angle-between-two-points-latitude-longitude/>

Graves, T., Karr, A., Roupail, N., and Takuriah, P. (1998). Real-Time Prediction of Incipient Congestion on Freeways from Detector Data. NISS Technical Report. <http://www.niss.org/sites/default/files/pdfs/technicalreports/tr79.pdf>, Accessed November 25, 2012.

Graves, T., Karr, A., Roupail, N., and Takuriah, P. (1998). Real-Time Prediction of Incipient Congestion on Freeways from Detector Data. NISS Technical Report. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.46.9033&rep=rep1&type=pdf>, Accessed May 7, 2016.

Grosvenor, Theodore; Grosvenor, Theodore P. (2007). Primary Care Optometry. Elsevier Health Sciences. p. 129. ISBN 0-7506-7575-6. Retrieved 29 November 2014.

Gupte, S.; Masoud, O.; Martin, R.F.K.; Papanikolopoulos, N.P., "Detection and classification of vehicles," in *Intelligent Transportation Systems, IEEE Transactions on*, vol.3, no.1, pp.37-47, Mar 2002 doi: 10.1109/6979.994794

HERE 360. "Slow down! HERE introduces its Traffic Safety Warning system". (2016) <http://360.here.com/2015/08/06/traffic-safety-warning/>

HERE 360. "HERE introduces HD maps for highly automated vehicle testing". (2016) <http://360.here.com/2015/07/20/here-introduces-hd-maps-for-highly-automated-vehicle-testing/>

HERE. *Autonomous cars can only understand the real world through a map*. (2015) <http://360.here.com/2015/04/16/autonomous-cars-can-understand-real-world-map/>

Image Sensing. ISS Product Brochure. (2015) <http://www.imagesensing.com/assets/documents/solutions/iss-product-brochure.pdf>

INRIX. INRIX Products. (2015) <http://inrix.com/products>

IRD, Inc. Microwave Radar. (2015) <http://www.irdinc.com/pcategory/counters-classifiers/microwave.html>

Ishak, Sherif, and Ciprian Alecsandru. "Optimizing traffic prediction performance of neural networks under various topological, input, and traffic condition settings." *Journal of Transportation Engineering* 130.4 (2004): 452-465.

ITE. *Manual of Transportation Engineering Studies*, 2<sup>nd</sup> Edition, Institute of Transportation Engineers (2010)

Jahangiri, A. and H. Rakha. Developing a Support Vector Machine (Svm) Classifier for Transportation Mode Identification Using Mobile Phone Sensor Data. *Annual Meeting of the Transportation Research Board*, 2014, pp.

Jia, Billy Williams, and Nagui Roupail. "Identification and calibration of site-specific stochastic freeway breakdown and queue discharge." *Transportation Research Record: Journal of the Transportation Research Board* 2188 (2010): 148-155.

Kapsch. Weigh In Motion Sensor. (2015)  
<https://www.kapsch.net/ktc/downloads/datasheets/video-sensor/Kapsch-KTC-DS-WIM?lang=en-US>

Kim, Seoungbum, and Benjamin Coifman. "Comparing INRIX speed data against concurrent loop detector stations over several months." *Transportation Research Part C: Emerging Technologies* 49 (2014): 59-72.

Kistler. Weigh In Motion Technology. (2015)  
<https://www.kistler.com/fileadmin/files/divisions/sensor-technology/weigh-in-motion/traffic-data-collection/200-746e-03.14.pdf>

Kondyli, A. (2009). Breakdown Probability Model at Freeway-Ramp Merges Based on Driver Behavior. Ph.D. Dissertation, University of Florida, Gainesville, Fla.

Kondyli, Alexandra, et al. "Development and evaluation of methods for constructing breakdown probability models." *Journal of Transportation Engineering* 139.9 (2013): 931-940.

Kuhne, Reinhart D., Reinhard Mahnke, and Julia Hinkel. "Modelling the Effects of Corridor Control Systems on Road Capacity." *5th International Symposium on Highway Capacity and Quality of Service*. 2006.

Lester, J., P. Hurvitz, R. Chaudhri, C. Hartung, G. Borriello and Mobilesense - Sensing Modes of Transportation in Studies of the Built Environment. *International Workshop on Urban, Community, and Social Applications of Networked Sensing Systems - UrbanSense08*, 2008,

Lorenz, Matt, and Lily Elefteriadou. "Defining freeway capacity as function of breakdown probability." *Transportation Research Record: Journal of the Transportation Research Board* 1776 (2001): 43-51.

Lujaina Al-Shabibi et al. Automobile Speed Violation Detection System using RFID and GSM Technologies. *International Journal of Applied Information Systems (IJ AIS)*, 2014.

Manzoni, V., D. Manilo, K. Kloeckl and C. Ratti. *Transportation Mode Identification and Real-Time Co2 Emission Estimation Using Smartphones*. 2010.

Mathwave (2016), <http://www.mathwave.com/easyfit-distribution-fitting.html>

May, Adolf. *Traffic flow fundamentals* (1990)



Mohan, P., V. N. Padmanabhan and R. Ramjee. Nericell: Rich Monitoring of Road and Traffic Conditions Using Mobile Smartphones. *6th ACM Conference on Embedded Networked Sensor Systems, SenSys 2008, November 5, 2008 - November 7, 2008*, 2008, pp. 323-336.

Moses, Ren, and Enock Mtoi. Evaluation of Free Flow Speeds on Interrupted Flow Facilities. Technical Report. Department of Civil Engineering, FAMU-FSU College of Engineering. Tallahassee, United States, 2013.

Movable Type Scripts (2016) <http://www.movable-type.co.uk/scripts/latlong.html>

Okamura, H., Watanabe, S., and Watanabe, T. (2000). An Empirical Study on the Capacity of Bottlenecks on the Basic Suburban Expressway Sections in Japan. In *Proceedings of the 4th International Symposium on Highway Capacity*, TRB Circular E-C018, Transportation Research Board, Washington D.C., pp. 120-129.

Paek, J., et al. (2010). Energy-efficient rate-adaptive GPS-based positioning for smartphones. *Proceedings of the 8th international conference on Mobile systems, applications, and services*. San Francisco, California, USA, ACM: 299-314.

Pan, Bei, Ugur Demiryurek, and Cyrus Shahabi. "Utilizing real-world transportation data for accurate traffic prediction." *Data Mining (ICDM), 2012 IEEE 12th International Conference on*. IEEE, 2012.

Park, Byungkyu, Carroll Messer, and Thomas Urbanik II. "Short-term freeway traffic volume forecasting using radial basis function neural network." *Transportation Research Record: Journal of the Transportation Research Board* 1651 (1998): 39-47.

Parkinson, B. W. Gps Error Analysis. In *Global Positioning System: Theory and Applications*, 1996.

Pérez, Joshué, et al. "An RFID-based intelligent vehicle speed controller using active traffic signals." *Sensors* 10.6 (2010): 5872-5887.

Persaud, Bhagwant, Sam Yagar, and Russel Brownlee. "Exploration of the breakdown phenomenon in freeway traffic." *Transportation Research Record: Journal of the Transportation Research Board* 1634 (1998): 64-69.

Persaud, Bhagwant, et al. "Breakdown-related capacity for freeway with ramp metering." *Transportation Research Record: Journal of the Transportation Research Board* 1748 (2001): 110-115.

Porter, J. David, et al. "Antenna characterization for Bluetooth-based travel time data collection." *Journal of Intelligent Transportation Systems* 17.2 (2013): 142-151.

Portugais, Brian, and Mandar Khanal. "Adaptive traffic speed estimation." *Procedia Computer Science* 32 (2014): 356-363.

Qi, Yaoyao, and Sherif Ishak. "Stochastic approach for short-term freeway traffic prediction during peak periods." *Intelligent Transportation Systems, IEEE Transactions on* 14.2 (2013): 660-672.

Qiao, Wenxin, Ali Haghani, and Masoud Hamedi. "A nonparametric model for short-term travel time prediction using bluetooth data." *Journal of Intelligent Transportation Systems* 17.2 (2013): 165-175.

Ramirez, Sergio Varela, and Kristian Kovacic. "Origin-Destination Matrix Estimation of Traffic Flow on Highway Network." (2013).

Raskovic, D. and D. Giessel. Battery-Aware Embedded Gps Receiver Node. *Mobile and Ubiquitous Systems: Networking & Services, 2007. MobiQuitous 2007. Fourth Annual International Conference on, 2007*, pp. 1-6.

RITIS Help (2016), <https://vpp.ritis.org/suite/help/#data-types>

Road Traffic Technology. Tagmaster. (2015) [http://www.roadtraffic-technology.com/contractors/access\\_control/tagmaster/](http://www.roadtraffic-technology.com/contractors/access_control/tagmaster/)

Sarvi, Majid, Masao Kuwahara, and Avishai Ceder. "Observing freeway ramp merging phenomena in congested traffic." *Journal of advanced transportation* 41.2 (2007): 145-170.

Schrank, et al. "2015 URBAN MOBILITY SCORECARD" Published jointly by The Texas A&M Transportation Institute and INRIX. (2015)

Sheldon, M. R. Introduction to probability models. Elsevier, 2010.

Shojaat, Siavash, et al. "Sustained Flow Index: A Stochastic Measure of Freeway Performance." Transportation Research Board 95th Annual Meeting. No. 16-5907. 2016.

Shojaat, S., Geistefeldt, J., Parr, S. A., Escobar, L., & Wolshon, B. (2017). *Applying the Sustained Flow Index to Estimate Freeway Capacity* (No. 17-06788).

Simjee, F. and P. H. Chou. Accurate Battery Lifetime Estimation Using High-Frequency Power Profile Emulation. *Low Power Electronics and Design, 2005. ISLPED '05. Proceedings of the 2005 International Symposium on, 2005*, pp. 307-310.

Skabardonis, Alexander, Pravin Varaiya, and Karl Petty. "Measuring recurrent and nonrecurrent traffic congestion." *Transportation Research Record: Journal of the Transportation Research Board* 1856 (2003): 118-124.

Song, T.-J., Kim, S., Hajbabaie, A., Williams, B.M., Roupail, N.M., List, G.F., 2015. A Novel Collision Classification Methodology Based on Temporal Link Speed Data and Congestion Thresholds, 94th Annual Meeting of Transportation Research Board. Transportation Research Board, Washington DC.

Traffax, Inc. Traffic and Travel Time Studies. (2015) <http://www.traffaxinc.com/content/traffic-and-travel-time-studies>

Traffic Tech. RTMS. (2015) [http://www.traffic-tech.com/remote\\_traffic\\_microwave\\_sensor.php](http://www.traffic-tech.com/remote_traffic_microwave_sensor.php)

Traffic.com. ITIP Stakeholder Applications. (2016) [www.stakeholder.traffic.com/#/sensor-manager/map](http://www.stakeholder.traffic.com/#/sensor-manager/map)

Trafficcast. BlueToad. (2015) <http://trafficcast.com/products/view/blue-toad/TRB>. Highway Capacity Manual 2010. Transportation Research Board of the National Academies, Washington, D.C.

Transportation Research Board. "Highway Capacity Manual." Washington, DC. (2016)

Van Lint, J., S. Hoogendoorn, and H. Van Zuylen. "Freeway travel time prediction with state-space neural networks: modeling state-space dynamics with recurrent neural networks."

Transportation Research Record: Journal of the Transportation Research Board 1811 (2002): 30-39.

Wavetronix. Smart Sensor HD. (2015)  
<http://www.wavetronix.com/en/products/smartsensor/hd/features>

White, J., C. Thompson, H. Turner, B. Dougherty and D. Schmidt. Wreckwatch: Automatic Traffic Accident Detection and Notification with Smartphones. *Mobile Networks and Applications*, Vol. 16, No. 3, 2011, pp. 285-303.

Williams, Billy, Priya Durvasula, and Donald Brown. "Urban freeway traffic flow prediction: application of seasonal autoregressive integrated moving average and exponential smoothing models." *Transportation Research Record: Journal of the Transportation Research Board* 1644 (1998): 132-141.

Yu, Xin, and Panos D. Prevedouros. "Performance and Challenges in Utilizing Non-Intrusive Sensors for Traffic Data Collection." (2013).

Yuan, Jing, et al. "Driving with knowledge from the physical world." *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2011.

Zegeer, John, et al. *Incorporating Travel Time Reliability into the Highway Capacity Manual*. No. SHRP 2 Report S2-L08-RW-1. 2014.

Zhan, Xianyuan, Ruimin Li, and Satish V. Ukkusuri. "Lane-based real-time queue length estimation using license plate recognition data." *Transportation Research Part C: Emerging Technologies* 57 (2015): 85-102.

Zhang, L, and D. Levinson 2004a. "Some properties of Flows at Freeway Bottlenecks." *Transportation Research Record* 1883, pp. 122-131.

Zhao, Nale, et al. "Analysis of traffic flow characteristics on ring road expressways in Beijing: using floating car data and remote traffic microwave sensor data." *Transportation Research Record: Journal of the Transportation Research Board* 2124 (2009): 178-185.

Zheng, Zuduo, et al. "Applications of wavelet transform for analysis of freeway traffic: Bottlenecks, transient traffic, and traffic oscillations." *Transportation Research Part B: Methodological* 45.2 (2011): 372-384.