Final Report

Development of a New Methodology to Characterize Truck Body Types along California Freeways

Contract Number: 11-316

Submitted By:

Stephen G. Ritchie, Ph.D. (Principal Investigator)

Yeow Chern A. Tok (Co-Investigator)

Institute of Transportation Studies

University of California, Irvine, 92697-3600

Prepared for:

Nesamani Kalandiyur

Air Quality Planning and Science Division

California Air Resources Board

1001 I street

Sacramento, CA 95814

January 2016

Disclaimer

The statements and conclusions in this report are those of the contractor and not necessarily those of the California Air Resources Board. The mention of commercial products, their source, or their use in connection with material reported herein is not to be construed as actual or implied endorsement of such products.

Acknowledgement

The project team expresses their appreciation to the following people for their support and assistance during the course of this research study:

Stan Norikane of Caltrans Division of Traffic Operations, for granting access to Irvine and Saigon Weigh-In-Motion sites for prototype testing, facilitating data collection at the Weigh-In-Motion sites in Willows and Redding, and granting access to WIM sites in the Central Valley for system deployment.

Vahid Nowshiravan of Caltrans Division of Planning, for assisting with coordinating and scheduling deployment efforts with Caltrans District 6 and 10 offices.

John Slonaker of Caltrans DRISI, for facilitating testing of data collection equipment at traffic cabinets, and actual data collection at the Irvine data collection site.

Rod Klashinsky, Neil Klassen, Ben Lutjens, Richard Maynard, Curtis Uyemoto and Daniel Leyva of IRD – for providing technical assistance during data collection efforts as well as test and actual deployments of WIM sites.

Eric Nelson, Fredrico Hormozi and Alex Kalbasi of Caltrans District 12 for providing technical support and access to the ILD Test Deployment Site.

John Castro, Arlene Cordero and Richard Montoya of Caltrans District 6, and Norma Gallegos, Scott Lockwood, Ihab Elzaanoun and David Arias of Caltrans District 10, for providing technical assistance with setting up equipment for deployment at ILD sites.

Lianyu Chu of CLR Analytics, for providing advanced detector card samples for prototype testing.

This Report was submitted in fulfillment of 11-316 Development of a New Methodology to Characterize Truck Body Types along California Freeways by the University of California, Irvine under the sponsorship of the California Air Resources Board. Work was completed as of August 25, 2015.

Table of Contents

Disclaimer i
Acknowledgement ii
Table of Contents iii
List of Figures vi
List of Tables viii
Abstract ix
Glossary of Terms, Abbreviations, and Symbols x
Executive Summary xii
1 Introduction 1
2 Background 3
2.1 Overview 3
2.2 Current Sources of Commercial Vehicle Data 3
2.3 Defining Truck Body Class 4
2.4 Weigh-In-Motion Systems 8
2.5 Inductive Signature Technology 11
2.6 Integrating Technologies 13
3 Data Collection and Processing 14
3.1 Overview 14
3.2 Data Sources 14
3.3 Data Collection Sites 16
3.4 Data Processing 19
3.5 Vehicle Classification Scheme 21
3.5.1 Single Unit Trucks 21
3.5.2 Combination Trucks 24
3.6 Data Summary 27
4 Model Development 30
4.1 Overview 30
4.2 Body Classification Models 30
4.2.1 Selecting Base Classifiers 31
4.2.2 Combining Classifiers 31
4.2.3 Model Inputs 32

4.2.4 Inductive Signature Body Classification Model 35
4.2.5 Inductive Signature and WIM Body Classification Model 43
4.2.6 Spatial Transferability Analysis 56
4.3 Methods for Backcasting Body Classification Data 57
4.3.1 Modeling Framework 58
4.3.2 Data 60
4.3.3 Results 61
4.3.4 Sensitivity Analysis 63
4.3.5 GVW Distribution Estimation 65
4.3.6 Conclusion 66
4.4 Optimal Site Selection for System Deployment 67
4.4.1 Background 67
4.4.2 Sensor Location Data 69
4.4.3 Sensor Location Problem 70
4.4.4 Solutions 77
4.4.5 Conclusion 83
4.5 Propagation of Gross Vehicle Weight to Inductive Loop Detector Sites 84
4.5.1 Methods 84
4.5.2 Data 87
4.5.3 Results 90
4.5.4 Conclusions 91
4.6 Summary and Conclusions 92
5 System Hardware and Data Architecture 93
5.1 Overview 93
5.2 Design of WIM and Inductive Signature Hardware 93
5.2.1 Integration Design for 1060-series WIM Controllers 93
5.2.2 Comparison of Inductive Signatures between 1060 and iSinc WIM Controllers 96
5.3 Hardware Configuration for Deployment 99
5.3.1 Inductive Loop Detector Sites 99
5.3.2 WIM Sites 100
5.4 Data Communications Architecture 102
5.4.1 Overview 102
5.4.2 Supporting Legacy IST Hardware 103
5.4.3 Transitioning the Middleware 104
5.4.4 Using the Data Downstream 106

5.5 Data Processing 108
5.5.1 Field Unit Data Processing 108
5.5.2 Server Data Processing 109
6 System Deployment 113
6.1 Overview 113
6.2 Test Deployments 113
6.2.1 Lab Development and Testing 113
6.2.2 WIM Test Deployment 113
6.2.3 ILD Test Deployment 114
6.3 Deployment in the California San Joaquin Valley 114
6.3.1 WIM Site Deployments 115
6.3.2 ILD Site Deployments 116
6.4 Truck Activity Monitoring System 118
6.4.1 System Design 118
6.4.2 Web Interface Design 118
6.5 Summary of Findings 124
6.5.1 Truck Travel Pattern Case Study 124
6.5.2 Problems Encountered 128
7 Conclusions 133
7.1 Implementation Effort and Cost 133
7.2 Limitations 133
7.3 Follow-up Studies and Future Recommendations 134
8 References 135
Appendix A – FHWA-CA Classification Scheme 138
Appendix B – Data Collection Site Configurations 139
Appendix C – Database Structure for Data Groundtruth System 140
Appendix D – Modified Decision Tree Model Architecture for Predicting Truck Volume by Body Configuration using WIM data 141
Appendix E – Pin-out Cross Assignment between 1060 WIM and IST-222 Detector Cards 142
Appendix F – Vehicle Classification Scheme for Inductive Loop Signature Only Classification Model 143
Appendix G – Vehicle Classification Scheme for Combined Weigh-In-Motion and Inductive Loop Signature Classification Model 150

List of Figures

Figure 1-1 Project Phases and Tasks Flow Chart 2
Figure 2-1 FHWA 13 Class Axle Based Classification Scheme 5
Figure 2-2 Samples of vehicle body configurations 6
Figure 2-3 Examples of spatial and temporal trends in truck body class. 7
Figure 2-4 WIM and TMS Sites in California 8
Figure 2-5 Types of WIM controllers deployed in California: (a) 1060 series and (b) iSINC Lite 9
Figure 2-6 WIM Site Configuration 10
Figure 2-7 Examples of Inductive Signatures 12
Figure 2-8 WIM Site Configuration (left) with data outputs from Advanced ILDs and WIM (right). 13
Figure 3-1 Data Collection System Architecture with SLR Still Image Camera 15
Figure 3-2 Data collection setup at WIM site 15
Figure 3-3 Examples of still images collected for various truck types 16
Figure 3-4 Data Collection Sites 17
Figure 3-5 Customized User Interfaces for Data Groundtruth 20
Figure 3-6 Examples of Single Unit Trucks 22
Figure 3-7 Examples of Combination Truck Trailer Body Classes 24
Figure 3-8 Examples of Tractor Units 25
Figure 3-9 Number of semi-trailer truck samples by body type across all sites 29
Figure 4-1 WIM system measurement and derived features for five axle semi tractor-trailer trucks 33
Figure 4-2 Inductive signature feature extraction procedure 34
Figure 4-3 Inductive Signature Parsing using WIM axle spacing measurements 35
Figure 4-4 Model Structure for Inductive Signature Model 36
Figure 4-5 Model Structure for WIM and Inductive Signature Model 44
Figure 4-6 FHWA Class 5 Model Framework 45
Figure 4-7 Examples of Axle Groups with FHWA Class 8 46
Figure 4-8 Box plots of spacing between the 1st and 2nd axles, vehicle length, and vehicle length interacted with axle spacing by tractor body class 50
Figure 4-9 Deriving GVW distribution for vans 60
Figure 4-10 Descriptive statistics of WIM variables; (a) Length (b) Axle spacing between 3rd and 4th axle (c) Overhang. 61
Figure 4-11 Baseline approaches and APE; (a) Proportion of body configuration from observed and VIUS (2002) data (b) APE from baseline proportion approach. 63
Figure 4-12 Results of transferability analysis 64
Figure 4-13 Hourly volume estimation using MDT model at Irvine (March 20, 2013) 65
Figure 4-14 Weight histogram and GMM; (a) van (b) platform (c) tank (d) 40ft box container. 66
Figure 4-15 Location of Data WIM sites in the San Joaquin Air Basin 68
Figure 4-16. Heavy Truck GPS Trajectory Coverage and WIM Station Locations in California 69
Figure 4-17. Re-id Performance Decay Assumptions: (a) Distance-based, (b) Location-specific. 76
Figure 4-18. Comparison of FSP site selection criteria results for P = 5 79
Figure 4-19. FSP results with truck body types 80
Figure 4-20. Comparison of RSP location selection results for $\boldsymbol{g0}$, $\boldsymbol{g1}$, $\boldsymbol{\tau}$ =100 and $\boldsymbol{g2}$, with (P = 10) 82
Figure 4-21. RSP Results with Decay Performance Assumptions 83

Figure 4-22 Example of GVW distribution by Body Class Group Contributing to Overall Site GVW Distribution 86
Figure 4-23 GVW Distributions along Southbound I-5 and SR-97 in Northern California 87
Figure 4-24 ATRI Truck Trip Trajectories 89
Figure 4-25 Shared Truck Trip Trajectories with the Lodi Northbound WIM Site as an origin 90
Figure 4-26 Spatial Interpolation of GVW Distributions Results 91
Figure 5-1 1060 WIM Controller Loop Sensor Module (Left) and IST-222 Loop Sensor Module Adapter (right) 94
Figure 5-2 A modified 222 Input File fabricated for the IST-222 LSM Adapter 95
Figure 5-3 Comparison of Hardware Setup for Standalone 1060 WIM Controller (top) and 1060 WIM Controller Integrated with IST-222 Detector Cards for Inductive Signature Data Logging 95
Figure 5-4 Experimental Setup: I-5 Yale location referred to as IST - IST (a), and I-405 Westminster location referred to as IST - iSinc (b) 97
Figure 5-5 Signature Transformation: Normalization (a), Shift (b), and Stretch Step (c) 98
Figure 5-6 Deployments at Inductive Loop Detector Sites 100
Figure 5-7 Deployment Hardware from NB 1-405 at Saigon 102
Figure 5-8 Overview of deployed field processing units shown on the ITS Signature Data Viewer 107
Figure 5-9 Graphical representation of signatures streaming live from SR-205 at Mountain House Parkway 108
Figure 5-10 Flowchart of hourly data processes 110
Figure 5-11 Flowchart of the WIM and signature record alignment algorithm 111
Figure 6-1 Overview of site deployments 114
Figure 6-2 Locations of deployed WIM sites 115
Figure 6-3 Locations of deployed ILD sites in Caltrans District 6 117
Figure 6-4 Locations of deployed ILD sites in Caltrans District 10 118
Figure 6-5 TAMS welcome page 119
Figure 6-6 Daily volumes by lane and aggregated FHWA classes at WIM site 121
Figure 6-7 Daily volumes by lane and aggregated vehicle classes at ILD site 121
Figure 6-8 Detailed hourly classification volumes at WIM site 122
Figure 6-9 Detailed hourly classification volumes at ILD site 123
Figure 6-10 Location of detector sites SR205_MH, I5_HAMMER and SR120_I5 for analysis of intermodal freight movement 125
Figure 6-11 TAMS heat map of hourly directional semi-tractor truck volumes along the SR-205 freeway at Mountain House Parkway 126
Figure 6-12 TAMS heat map of hourly directional semi-tractor truck volumes along the I-5 freeway at Valley Forge 127
Figure 6-13 TAMS heat map of hourly eastbound semi-tractor truck volumes along the SR-120 freeway east of the I-5 freeway 128
Figure 6-14 Classification count errors due to low signature magnitudes 129
Figure 6-15 Classification count errors due to high signal noise 130
Figure 6-16 Coverage maps for deployment sites affected by data communications issues (obtained from Sensorly) 131

List of Tables

Table 2-1 Data Gaps, Definitions, and Existing Sources for Commercial Vehicle Data 4
Table 3-1 Summary of Data Collection Sites 18
Table 3-2 Summary of Data Collection Time Periods 19
Table 3-3 VIUS and Model Single Unit Trucks Body Classification 23
Table 3-4 VIUS and Model Drive Unit Body Classification 26
Table 3-5 VIUS and Model Body Classification for Semi-Trailers for Existing VIUS Classes 27
Table 3-6 Volume by site and FHWA class 28
Table 4-1 Inductive Signature Only Model Tier 1 and 2 Cross Classification Table and Volume Accuracy 38
Table 4-2 Single Unit Truck without Trailer MCS Summary 39
Table 4-3 Inductive Signature Only Model Tier 3 Single Unit Truck with Trailer MCS Results 40
Table 4-4 Inductive Signature Only Model Tractor Results 40
Table 4-5 Semi Tractor Trailers MCS Summary 41
Table 4-6 Multiple Semi Tractor Trailer Combination Trucks MCS Summary 42
Table 4-7 Inductive Signature Only Model Tier 3 Results Summary 43
Table 4-8 FHWA 4 Body Classification Model Results 47
Table 4-9 FHWA Class 5 without Trailer MCS Summary 48
Table 4-10 FHWA 6 MCS Summary 48
Table 4-11 FHWA 7 MCS Summary 49
Table 4-12 Results of the WIM-Only Tractor Classification Model 51
Table 4-13 Spatial Transferability Results of the WIM-Only Tractor Classification Model 51
Table 4-14 FHWA Class 8 MCS Summary 52
Table 4-15 FHWA Class 9 MCS Results Summary 53
Table 4-16 FHWA 14 MCS Summary 53
Table 4-17 FHWA 11 and 12 MCS Summary 54
Table 4-18 WIM-Signature Model Summary 55
Table 4-19 Spatial Transferability Analysis Cross Classification Table for Irvine Data 57
Table 4-20 Summary of Data for Backcasting Model Development 60
Table 4-21. Results of Decision Tree Model 62
Table 4-22 Results of Sensitivity Analysis on Backcasting Model 64
Table 4-23. VIUS Body Types in CSFFM 73
Table 4-24. Body Type Distribution by Commodity Group (FHWA Class 9) 73
Table 4-25. Comparison of FSP location selection results for different truck body type priorities 78
Table 4-26. RSP location selection results 81
Table 4-27 Sample of the Directional Spatial Weight Matrix from GPS Truck Trip Trajectories 89

Abstract

The purpose of this project was to develop a new methodology to characterize truck body types along California Freeways. With new information on truck activity by body types, results from this study are expected to improve heavy duty vehicle classification in the Emission Factors (EMFAC) model and the California Vehicle Activity Database (CalVAD), and provide critical data that is required for the analysis of freight movement that will benefit the California Statewide Freight Forecasting Model (CSFFM) and other freight- or truck-related studies.

This study sought to develop two types of classification models: the first from the combination of inductive loop signature and weigh-in-motion (WIM) data, and the second from standalone inductive loop signature data. The key benefit of these models is their readiness for implementation at existing traffic detector infrastructure such as inductive loop detector (ILD) and WIM sites. It was demonstrated through this study that the modifications to existing inductive loop detector and WIM sites were minimal, and did not compromise existing operations. The standalone inductive signature classification model (designed for implementation an existing ILD sites) demonstrated the ability to distinguish over 40 truck configurations, while the combined inductive loop signature and WIM classification model was able to identify over 60 truck types. These models were subsequently deployed at sixteen selected sites in the California San Joaquin Valley. A prototype web interface called the Truck Activity Monitoring System (TAMS, http://freight.its.uci.edu/tams) was designed to generate dynamic reports of the results via an interactive web-based user interface.

Other models developed in this study include a method for estimating truck volumes by a reduced number of body types from standalone WIM data, an optimal site selection model for determining the optimal sites for deployment of the advanced classification system developed in this study, and a method for estimating gross vehicle weight distributions at inductive loop detector sites instrumented with inductive signature technology by using data obtained from affiliated WIM sites.

The project was separated into three phases: proof-of-concept truck body classification models were developed in Phase 1; model enhancement was performed in Phase 2; and system deployment took place as Phase 3.

Keywords
Truck body classification, inductive loop signature, Weigh-in-motion (WIM), truck activity monitoring system

Glossary of Terms, Abbreviations, and Symbols

Abbreviation	Definition
1060	1060 Weigh-in-motion controller unit
3S2	Five-axle semi-tractor trailer with 3 axle tractor pulling a tandem axle semi-trailer
APE	Absolute Percent Error
ASTM	American Society of Testing and Materials
ATRI	American Transportation Research Institute
Caltrans	California Department of Transportation
CalVAD	California Vehicle Activity Database
CART	Classification and Regression Tree
CCR	Correct Classification Rate
CEC	California Energy Commission
CFS	Commodity Flow Survey
CSFFM	California Statewide Freight Forecasting Model
DT	Decision Tree
FAF	Freight Analysis Framework
FAZ	Freight Analysis Zones defined in the California Statewide Freight Forecasting Model
FHWA	Federal Highway Administration
FSP	Flow-interception Sensor Placement
GA	Genetic Algorithm
GMM	Gaussian Mixture Model
GPS	Global Positioning System
GVW	Gross Vehicle Weight
HHDT	Heavy-Heavy Duty Trucks
HPMS	Highway Performance Measurement System
ILD	Inductive Loop Detector
IRD	International Road Dynamics
iSinc/iSinc Lite	iSinc Weigh-in-motion controller unit
IST-222/IST	Inductive Signature Technologies 222 detector card
LSM	Loop Signature Module of the Weigh-in-motion controller
MAPE	Mean Absolute Percent Error
MCS	Multiple Classifier Systems
MDT	Modified Decision Tree
MLFF	Multilayer Feed Forward Neural Network
MT	Multi Trailer Axle Configuration
NB	Northbound
NBC	Naïve Bayes Combination
NCFRP	National Cooperative Freight Research Program

NCHRP	National Cooperative Highway Research Program
NT	Normalized Tree
OD	Origin-Destination
PC	Passenger Cars
PeMS	Caltrans Performance Measurement System
PNN	Probabilistic Neural Network
RV	Recreational Vehicle
SB	Southbound
ST	Semi-Trailer Axle Configuration
SU	Singe Unit Trucks without Trailers
SVM	Support Vector Machine
TAMS	UCI-ITS Truck Activity Monitoring System
TOD	Time of day
TMS	Traffic Monitoring Site
UCI	University of California, Irvine
UCI-ITS	University of California, Irvine Institute of Transportation Studies
USB	Universal Serial Bus
VIUS	Vehicle Inventory and Use Survey
WCU	Weigh-in-motion Controller Unit
WIM	Weigh-in-motion

Executive Summary

Background

The purpose of this project was to develop a new methodology to characterize truck body types along California Freeways. With new information on truck activity by body types, results from this study are expected to improve heavy duty vehicle classification in the Emission Factors (EMFAC) model and the California Vehicle Activity Database (CalVAD), and provide critical data that is required for the analysis of freight movement that will benefit the California Statewide Freight Forecasting Model (CSFFM) and other freight- or truck-related studies.

In this project, inductive signature technology was used to develop and deploy an advanced vehicle classification system comprising two distinct classification models at sixteen selected locations for two types of facilities: existing inductive loop detector (ILD) and weigh-in-motion (WIM) sites. ILD sites currently provide only vehicle volume counts. And although WIM sites provide axle-based truck classification and axle weight data, they cannot provide information on truck configuration that can provide further insight on industry and freight activity. Through this study, it is now possible to obtain higher resolution truck data which can enable more accurate estimates of GHG and other truck emissions, allow for decision makers to make more informed decisions for pavement management across a wider range of locations, and provide insight into the spatial distribution of body types for freight forecasting applications. The project will enhance CalVAD by incorporating a higher level of detail of commercial vehicle body classes, thus expanding the estimation of heavy-heavy duty truck (HHDT) activity by CalVAD. It will also be used to improve the EMFAC model by increasing the number of truck types in the model and will improve on the understanding of truck activity in the State of California.

The project was divided into three phases with the following accomplishments: Phase 1 developed proof-of-concept body classification models; Phase 2 enhanced the proof-of-concept models and created techniques for obtaining body classification predictions for historical WIM data; and Phase 3 deployed the developed classification models to selected WIM and ILD sites located in California's Central Valley.

Objectives and Methods

The research carried out under this project represents a completely new method for obtaining high resolution truck data. The objectives of Phase 1 included the (1.1) development of body type classification models using inductive signature data, (1.2) development of body type classification models using inductive signature data fused with WIM data, and (1.3) investigation of hardware interface configurations between WIM controllers and inductive signature technologies. Phase 2 objectives included (2.1) data collection, (2.2) model enhancement, (2.3) development of a methodology to propagate weight data to ILD locations, (2.4) development of a method to generate body class estimates from historical WIM data, and (2.5) development of an optimal facility location model. Lastly, Phase 3 was comprised of (3.1) equipment and model deployment to selected ILD and WIM sites and (3.2) system shakedown efforts for deployed sites.

Phases 1 and 2 required extensive data collection for model development and enhancement. Data was collected at seven WIM sites and one ILD site across California. Collected data – comprising inductive loop signatures, WIM controller outputs, and still images – were pre-processed, loaded into a relational database, and processed using a specially developed software user interface designed to assist with associating each collected sample's data (signature, WIM data, and photo) and classifying the record according to a preliminary body configuration classification scheme. The truck body configuration classification scheme was originally derived from the body classes defined in the 2002 Vehicle Inventory and Use Survey (VIUS), and was further refined to reflect the variety of body configurations observed in the collected data. The proof-of-concept model for ILD-based classification included data from a single ILD site while the proof-of-concept model for loop and WIM focused only on the two most common truck classes: the FHWA class 5 single-unit two-axle trucks and FHWA class 9 five-axle semi-tractor-trailers.

Unlike the proof of concept models developed during Phase 1 which employed a Feed Forward Neural Network model architecture, the fully enhanced models developed in Phase 2 adopted a multiple classifier systems approach with probabilistic model combination to produce even more accurate and detailed truck body class predictions for all vehicle types ranging from two axle pickup trucks to six or more axle semi-tractor trailers.

For the propagation of weight data to ILD sites (Task 2.3), a Gaussian Mixture Model (GMM) approach was employed in conjunction with Global Positioning Data (GPS) of truck trajectories to model the gross vehicle weight distribution of trucks at ILD sites by body type.

For the backcasting task (Task 2.4), historical estimates of body class volumes of five axle semi-trailers were estimated from WIM axle weight, spacing, and length data using a modified decision tree framework while historical estimates of gross vehicle weight distributions were estimated by body class using GMMs in combination with the modified decision tree approach.

Lastly, a method to select the deployment locations of the body classification models utilized an optimization model that selects the optimal site locations that capture the most number of unique truck trajectories in the GPS data set.

Results

Two main classification models were developed in this study: the truck body classification models for WIM sites, and ILD sites.

The enhanced body classification models produced for WIM sites are subdivided by FHWA axle class such that each axle class has a unique set of body classes and resulting correct classification rates (CCR). Overall, the nine axle stratified models produce a total of 63 body classes with individual models ranging from four to 16 body classes. The CCRs were in the approximate range of 75% to 96%.

The ILD body classification model was stratified into three tiers with the first tier predicting the general vehicle configuration of the vehicle as a single unit or multi-unit truck, the second tier predicting the body configuration, and the final tier predicting the body class. In summary, the ILD model consisted of 47 body classes across four body configuration groups with CCR ranging from 72% to 94%.

The backcasting approach was applied to five axle semi-tractor trailers to predict volumes by trailer body type across five trailer classes: vans, tanks, platforms, intermodal containers, and others. The model produces body class volumes estimates with low error in the range of approximately 5% absolute percent error in volume. Finally, for the propagation of gross vehicle weight (GVW) to ILD site locations, accuracy of the model was determined by applying a hold-one-out method in which the GVW was estimated for an individual WIM site and compared to the observed GVW distribution at that site. This was repeated for all WIM sites. Approximately 65% of the predicted GVW distributions were found to be statistically significant representations of the observed distributions at the 95% confidence level.

The deliverables and results related to hardware development were demonstrated as part of the Phase 3 deployments to the sixteen selected sites (4 WIM and 12 ILD site locations) in the California San Joaquin Valley Air Basin. The deployment locations are shown in Figure ES-1. The deployment was performed on a variety of facility types, including freeways, highways and arterials.

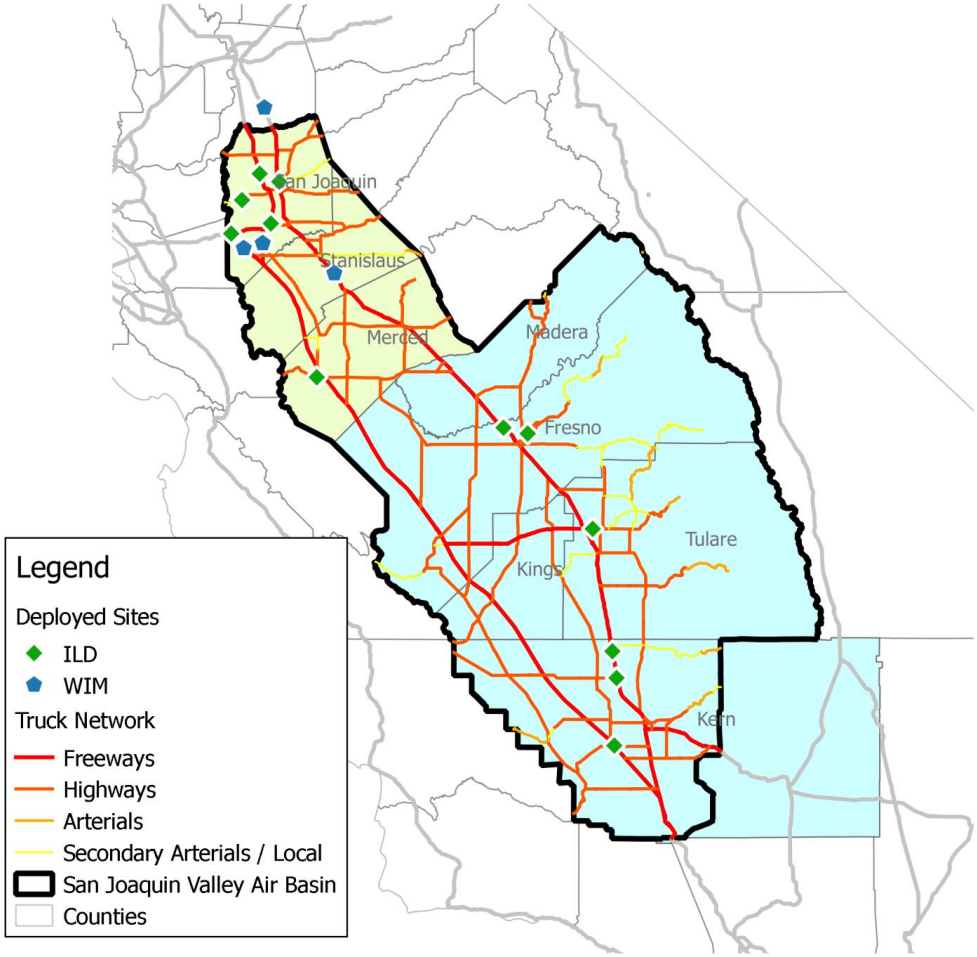


Figure ES-1 Overview of site deployments

Data from the deployed sites were live-streamed, processed and archived into a relational database during the deployment period. The prototype Truck Activity Monitoring System (TAMS) web-interface[1]

[1] http://freight.its.uci.edu/tams

was developed to provide access to on-demand detailed hourly summary reports of the body classification models at the deployed sites (sample queries results from TAMS are shown in Figure ES-2).

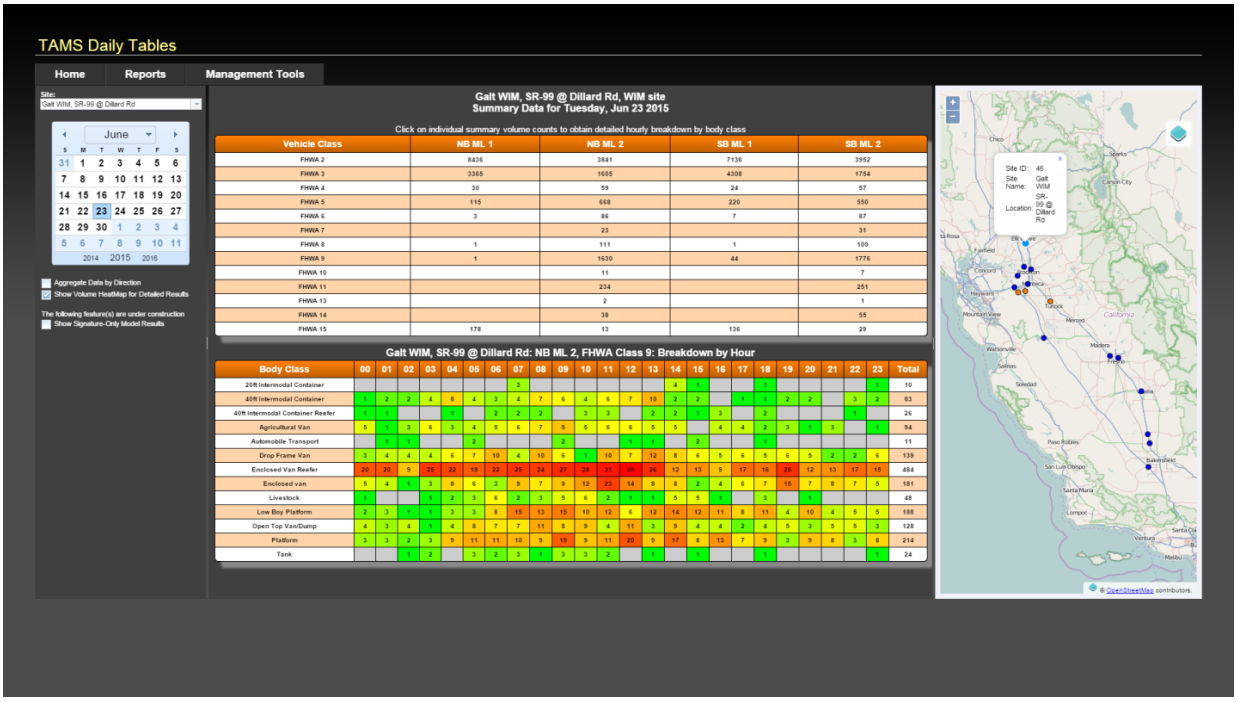


Figure ES-2 Detailed hourly volumes of FHWA class 9 trucks by body configuration at the Galt WIM site along the SR-99 freeway

Implementation Effort and Cost

Site implementation hardware at existing ILD sites primarily comprises a field processing unit, signature capable ILD cards, and wireless modems. The number of ILD cards required varies by site, depending on the loop configuration (single or double inductive loop sensors per lane) and the number of lanes to be monitored. They typically range been two and four for the monitoring of truck traveled lanes. For compatible WIM sites with 1060 series controllers, some additional hardware is required to interface the ILD cards with the controller as well as to receive WIM data records from the WIM controller. This includes a customized ILD adapter and a serial null modem. Hence, depending on the number of ILD cards required, the hardware implementation costs for an existing loop detector site may range from approximately $2,500 and $4,000. The implementation costs at WIM sites are expected to be about $500 higher than ILD sites. In addition, a wireless data communications monthly subscription cost of about $40 is required for data transmission from the field units to the server. This monthly recurring costs may be eliminated if existing high speed communications infrastructure can be utilized.

Limitations

Some limitations were observed during the development and deployment of the classification systems in this study:

- The traffic sensors used in this study assume that most drivers observe good lane discipline. Vehicles straddling between lanes or changing lanes while over the sensors will generate

erroneous data. Hence, it is important to identify locations where there are no significant weaving movements.

- Sensor data quality is typically affected under severe congestion. Although the quality of inductive signature data is not affected by a vehicle's speed, it is sensitive to significant accelerations and decelerations relative to a vehicle's speed because they introduce a skew to the signature waveform. This may affect the accuracy of the classification models developed in this study. Hence, the system performance may be compromised during congested traffic conditions. It should be noted that this weakness is not confined within inductive signature technology; congested conditions are known to affect other conventional traffic measures such as axle spacing measurements at WIM sites and volume count accuracy at conventional ILD sites.
- Inductive signature magnitudes were observed to have a high signal-to-noise ratio for some lanes at certain sites. This may be attributed to the physical quality of the loop installations, as poor splices or deteriorating insulation may cause such problems. There are two concerns with a high signal-to-noise ratio. It may lead to false detections of noise signals as vehicles, and result in misclassification errors due to the noise in signature data being misidentified as vehicle signature features. This problem may be addressed by verifying the condition of inductive loop sensors and leads, and their connections.
- The amplitudes of inductive signatures were found to be low at some sites. As a consequence, trucks pulling trailers with a high ground clearance may be detected as multiple vehicle signatures. These fragmented signatures are typically misclassified as passenger vehicles or single unit trucks. Hence, affected sites will typically undercount multiple unit trucks, and provide an overestimate of passenger car and light truck volumes. A quick assessment of the raw inductive signature stream will identify sites affected by this problem.

Future Recommendations

This study has demonstrated the potential to identify truck body configurations at an unprecedented level of detail. With the emergence of alternative drivetrain technologies in trucks, it will be a worthwhile effort to investigate the potential of inductive signature technology in identifying trucks by drivetrain as well. This will be particularly useful in providing truck activity emissions estimates through monitoring truck activity by drive train technologies and provide information on the best policies to maximize the benefit of technology investments.

Conclusions

This project has demonstrated and applied inductive signature technology to detect truck body class at inductive loop detector (ILD) and weigh-in-motion (WIM) sites in California. Extensive data collection efforts across the state resulted in an exceptional inventory of truck body types that were used to develop two truck body classification models: a standalone inductive signature only model and an integrated WIM and inductive signature model, which were designed for implementation at ILD and WIM sites, respectively.

The standalone inductive signature only classification model is capable of distinguishing five main vehicle configurations (Passenger Vehicle, Single Unit Truck, Truck with Single Trailer, Tractor with Semi-Trailer, and Tractor with Multiple Trailers) with an accuracy of 93.4% and volume error of 1.3%. The model further expands the truck classifications to over 40 detailed body configurations. The model

predicted 34 of these configurations with an accuracy exceeding 70%, with 19 configurations reporting a volume error of within 10%, based on independent test data.

The classification model combining WIM and inductive signature data is capable of predicting 63 body configurations, comprising 26 single unit truck, 25 semi-trailer body configurations, and 12 multi-unit truck body configurations. Independent test data results showed that 51 body configurations achieved classification accuracies above 70%, and 33 body configurations reported lower than 10% in volume errors.

In addition to the body class models developed around inductive signature technology, the extensive dataset was leveraged to develop a methodology to predict body class volumes from WIM data for sites not equipped with inductive signature technology which is useful for backcasting tasks related to the validation and calibration of the CSFFM. Furthermore, the procedure designed to estimate gross vehicle weight distributions at ILD sites showed promising results.

The deployment of inductive signature technology and corresponding body classification models to 16 sites in the California San Joaquin Valley gives practitioners and researchers a valuable tool to assess detailed truck activity, freight movements and impacts. The results obtained at selected sites have been shown to corroborate strongly with existing freight facilities in the region.

With new information on truck activity by body types, results from this study are expected to improve heavy duty vehicle classification in the EMFAC model and the CalVAD, and provide critical data that is required for the analysis of freight movement that will benefit the CSFFM and other freight- or truck-related studies.

As a follow-up effort, Caltrans has sponsored a $1M study to further enhance the classification models and expand the number of deployed sites to over 90 across the State of California. These future deployments will be located along major truck corridors within metropolitan areas, at regional cordon lines, and near state boundaries. This follow-up study will also further enhance the TAMS web interface through which users can examine and download individual body class predictions in addition to hourly summaries by location.

1 Introduction

A significant proportion of goods movement is transported by trucks, and the value and tonnage of goods are expected to grow over time. Trucks have a significant impact on pavement infrastructure, traffic congestion, pollution and "quality of life". To provide a better understanding of the behavior of freight-related truck movements, it is necessary to obtain detailed and comprehensive truck data. However, current traffic detection infrastructure in the state of California is not equipped to collect sufficiently detailed truck data to address these concerns. The prevailing inductive loop detector (ILD) equipped traffic detection infrastructure managed by the California Department of Transportation (Caltrans) comprise Weigh-In-Motion (WIM) and ILD sites. There are about 160 directional locations within the state monitored with WIM technology that are used primarily for continuous collection of axle weight measures and axle-based classification of trucks[2]. These stations are located along freeway and expressway corridors that experience heavy truck usage. Although there are nearly 8,000 Traffic Monitoring Sites (TMS) currently being deployed along the mainline of major truck corridors in California, most of which are equipped with ILD technology for vehicle detection. However, these were not originally designed to provide detailed truck data.

The purpose of this project was to develop a new methodology to characterize truck body types along California Freeways. To do so, selected existing ILD and WIM data collection sites were updated with inductive signature technology. This approach facilitates the acquisition of advanced truck data without compromising existing traffic operations at each site. The higher resolution truck data can enable more accurate estimates of GHG and other truck emissions, allow for decision makers to make informed decisions for pavement management, and provide insight into the spatial distribution of body types for freight analysis applications. The project was divided into three phases with the following general goals: Phase 1 developed proof-of-concept body classification models; Phase 2 enhanced the proof-of-concept models and created techniques for propagating WIM classifications to ILD locations as well as techniques for estimating truck volumes by body class using only WIM data; Phase 3 deployed inductive signature capabilities to selected WIM and ILD sites located in California's Central Valley. The flow chart shown in Figure 1-1 depicts the overall scope of the project.

[2] There are approximately 50 additional WIM sites in California which are designated as PrePass™ stations used to screen overweight trucks for further assessment at Commercial Vehicle Enforcement Facilities, and not typically used for WIM data collection.

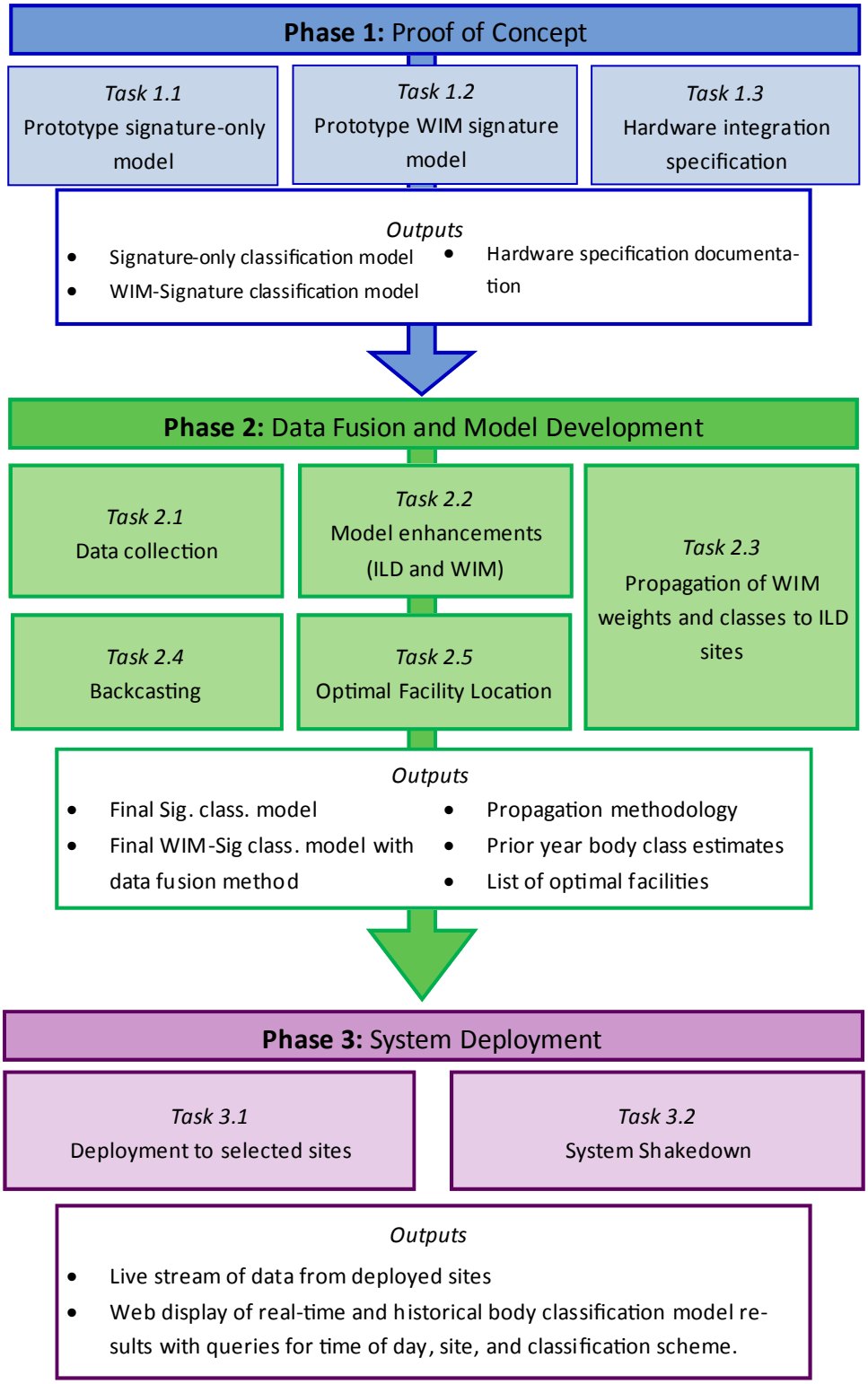


Figure 1-1 Project Phases and Tasks Flow Chart

2 Background

2.1 Overview

Commercial vehicle impacts on infrastructure, safety, emissions and the economy far exceed their modest proportion in daily traffic volumes. At the national level, trucks account for around 10% of the annual vehicle distance traveled (Highway Statistics, 2010). Although this represents a small portion of the total travel, the impacts of trucks on the environment, the economy, and infrastructure, are much more substantial than that of passenger vehicles. In fact, according to CARB's Mobile Source Emissions Inventory, heavy-duty diesel trucks are the "single largest source of nitrogen oxide emissions in California" as well as the "largest source of diesel particulate matter" (CARB, 2010). Further, the economic impacts of trucks in regard to freight transport are considerable. The Bureau of Transportation Statistics reported that "trucking as a single mode (including for-hire and private use) was the most frequently used mode in freight movement, hauling an estimated 70 percent of the total value, 60 percent of the weight, and 34 percent of the overall ton-miles" (BTS, 2006). Given the vast health, safety, and economic impacts of commercial vehicles, many agencies are interested in increasing the amount of and improving the quality of publically available commercial vehicle activity data.

2.2 Current Sources of Commercial Vehicle Data

While many states have wide ranging resources for passenger traffic, such as California's Performance Measurement System (PeMS), a significantly smaller set of states collect or measure commercial vehicle traffic at the same level as passenger vehicles. Conventional data sets, although commendable, have various limitations and lack detail, especially regarding truck travel patterns and characteristics beyond basic volume measurement. Due to privacy concerns in the trucking industry, commercial vehicle activity data can be difficult and expensive to obtain, and many times is incomplete due to small sample sizes. To further exacerbate the lack of data, a much used national data set, the Vehicle Inventory and Use Survey, was discontinued over ten years ago, and has not found a replacement since. If policies are to be formed which reduce the negative impacts of truck traffic such as temporal shifts or route restrictions, there needs to be a way to assess whether the policy has had any affect, which points to the need for route specific, temporally continuous, up-to-date, and representative truck data.

The National Cooperative Freight Research Program (NCFRP) Repot 39 reviewed the current state of truck activity data and determined critical gaps in freight data (NCFRP, 2014). Critical information gaps, their basic definition, and the best publically available data sources for each variable are summarized in Table 2-1.

Table 2-1 Data Gaps, Definitions, and Existing Sources for Commercial Vehicle Data

Variable	Definition	Best Publically Available Sources
Vehicle Miles Traveled (VMT)	Measure of the extent of motor vehicle operation within a specific geographic area	Highway Performance Measurement System (HPMS)
Tons/Ton-Miles	Total weight of the entire shipment multiplied by the mileage traveled by the shipment	Commodity Flow Survey (CFS)
Value/Value-Miles	Market value of goods shipped multiplied by the mileage traveled by the shipment	Commodity Flow Survey (CFS)
Origin-Destination (OD) Flows	The start and end points for a particular truck trip	Commodity Flow Survey (CFS) Freight Analysis Framework (FAF)
Vehicle Speed	Speed of vehicle	Roadside traffic counters Weigh-in-motion (WIM) GPS traces

The major source for several of these data gaps include the Commodity Flow Survey (CFS) which is a comprehensive survey of businesses, warehouses, and freight managing offices conducted at the national level every five years. In addition to the CFS and HPMS, several other existing data sources were identified in NCFRP Report 39 as sources for commercial vehicle data including the Vehicle Inventory and Use Survey (VIUS), Weigh-in-Motion (WIM) systems, privately owned truck GPS data, and state and federal truck registration records.

Many of these sources lack in their ability to segment each of these variables by: (1) commodity type, (2) vehicle type, (3) vehicle characteristics, and (4) spatial coverage. Unfortunately, no single source addresses each and every data gap with the desired level of detailed segmentation. According to NCFRP Report 39, the CFS and VIUS possess the best ability to cover each of the data gaps at some level of segmentation. Other research highlighted VIUS, GPS, and WIM data as three core data sets which provide truck data stratified across the segmentation categories (ITS, 2010).

2.3 Defining Truck Body Class

One major limitation is that data sources which provide commodity information (e.g. CFS, SAS, and VIUS) come from surveys so cannot be linked to link or route level while data sources that provide observed volumes, weights, and vehicle types (e.g. WIM and GPS) do not provide commodity information. This means there is a significant advantage in connecting observed vehicle data to commodity information. Since body configuration is closely linked to commodity carried and other operating characteristics, body class data can provide the link to the desired commodity information currently provided only in CFS or VIUS.

Vehicle configuration generally refers to axle or length based groupings of vehicles and takes different forms depending on the agency using the data. A commonly used scheme, FHWA's Scheme F, shown in Figure 2-1 defines 13 axle based classes. Within each axle based category, vehicles can be further distinguished by body configuration. Examples of body configurations are shown in Figure 2-2 . Semi-tractor trailer body types are classified by their drive unit and trailer unit. Similarly, multi-trailer trucks are defined by drive and trailer units. For example, the body type of a commonly observed five axle semi-tractor trailer listed by the FHWA scheme Figure 2-1 as 'Class 9 single trailer' might have a body configuration of a van, intermodal container, a tank, or a platform. In California, the California

Department of Transportation (Caltrans), the California Air Resources Board (CARB), and the California Energy Commission (CEC) have developed their own classification schemes which represent different levels of detail regarding truck characteristics.

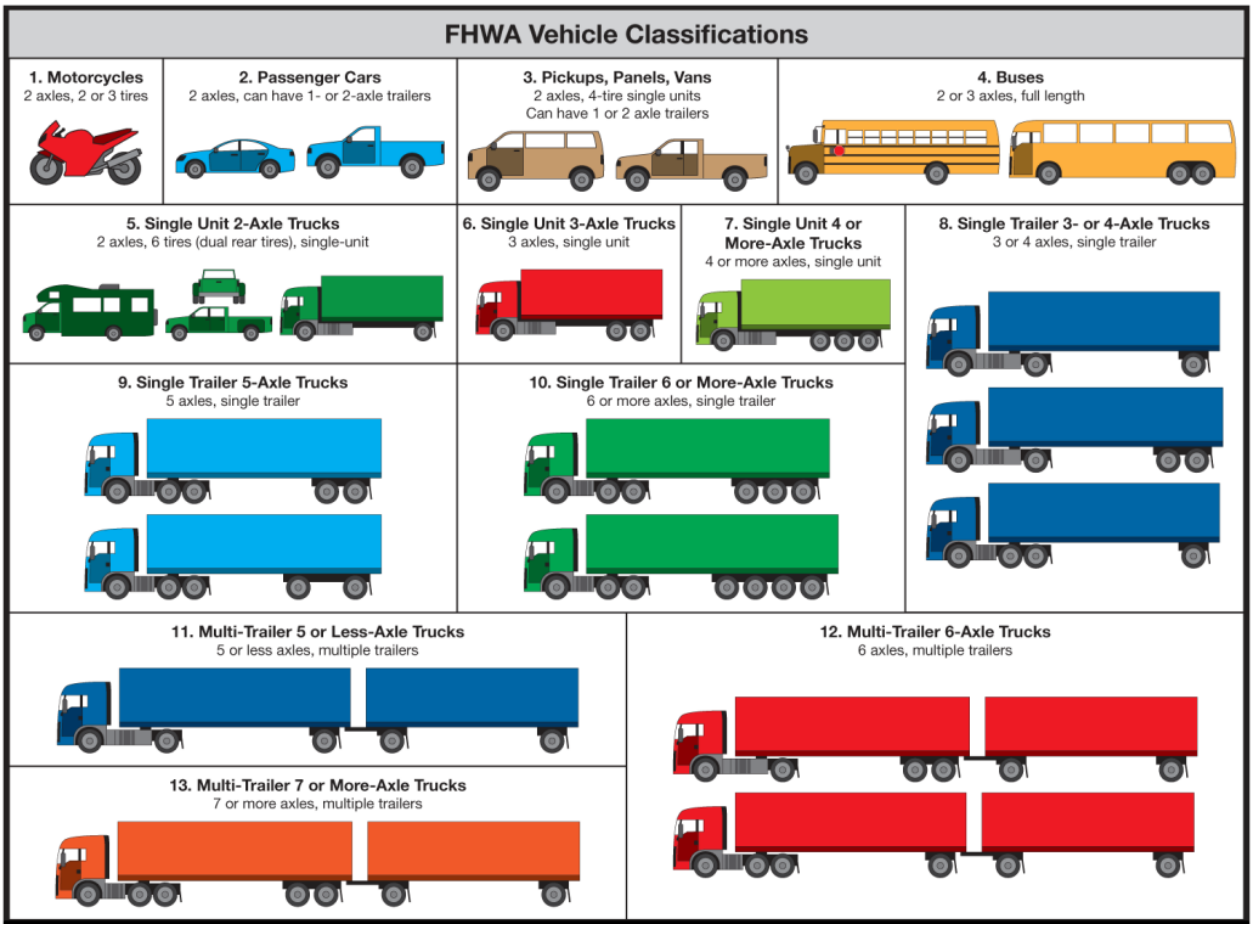


Figure 2-1 FHWA 13 Class Axle Based Classification Scheme

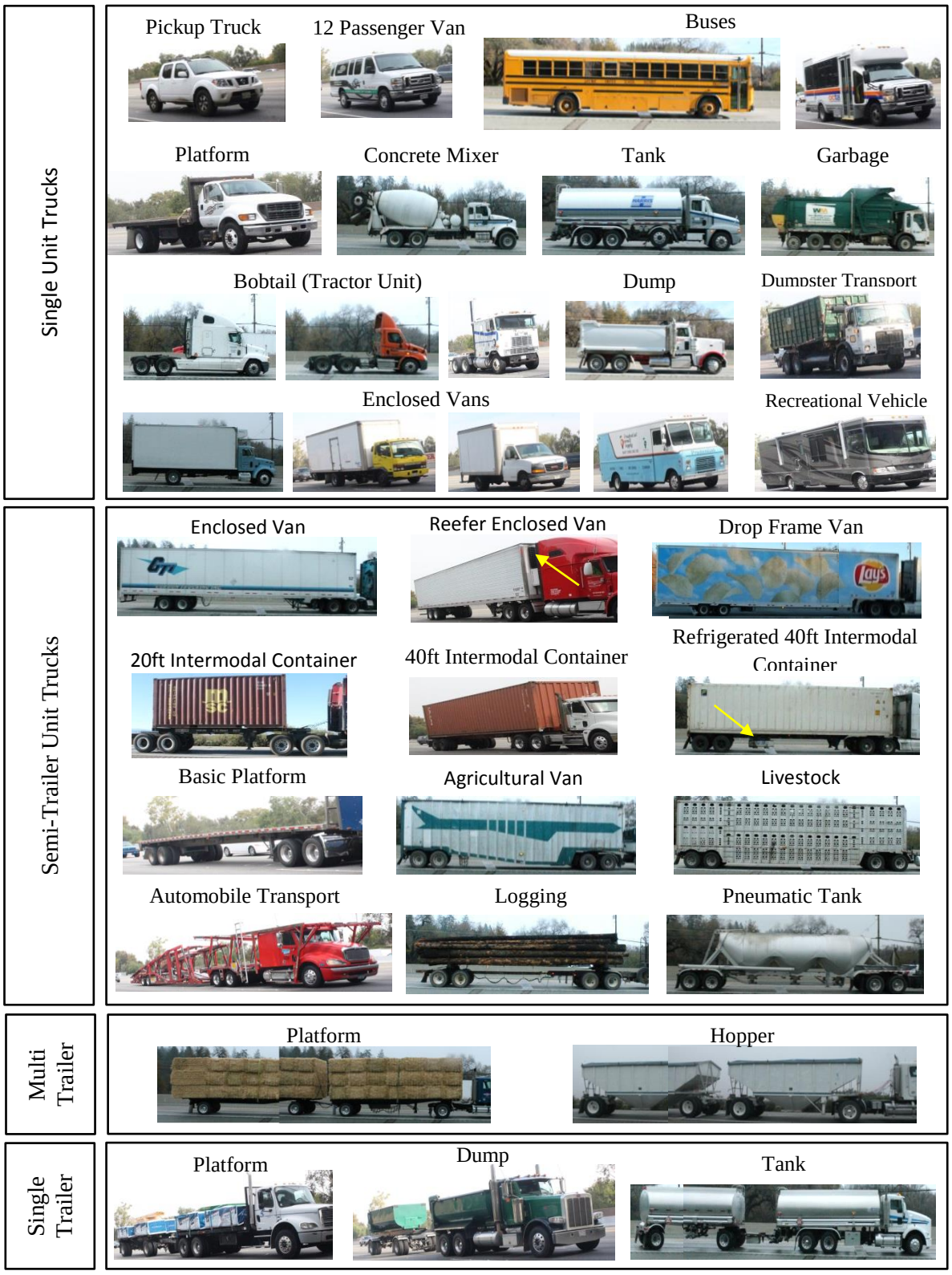
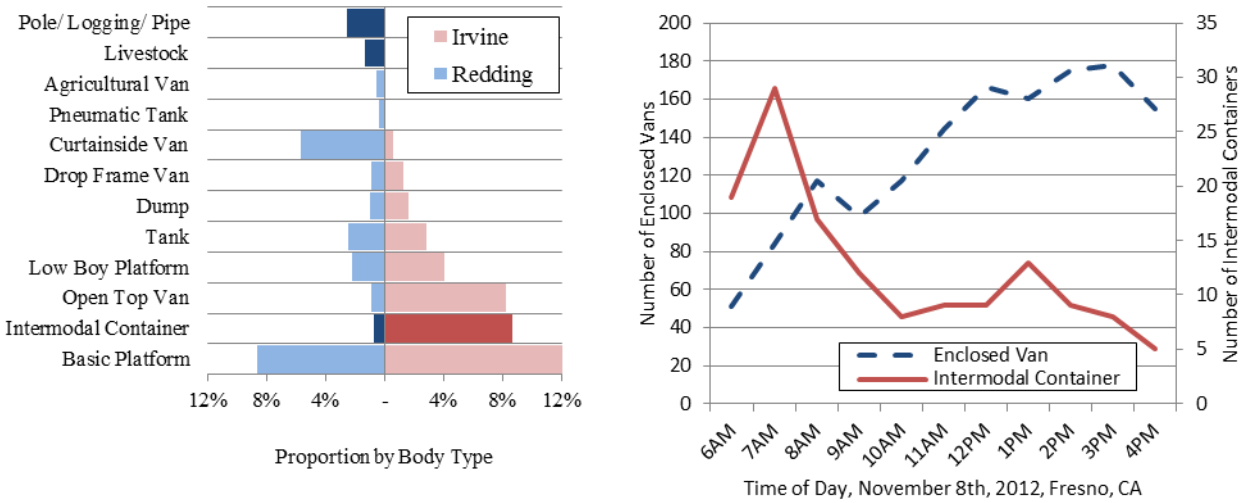


Figure 2-2 Samples of vehicle body configurations

Spatial and temporal trends in truck, trailer, and drive unit body types signal differing activity patterns, industry-specific operating characteristics, regional land uses, and seasonal commodity flow patterns. Figure 2-3(left) shows the variation in minority trailer body classes at two WIM sites in California- an urban location (Irvine) near the Ports of Los Angeles and Long Beach and a rural site (Redding) 120 miles from the Oregon-California border. Although vans represent over 60% of truck traffic, the proportions of industry specific, minority classes contrast significantly between these two sites. In Figure 2-3(right) the volume of enclosed vans observed at a WIM site located in central California (Fresno) peaks in the late afternoon while intermodal container traffic decreases over this same period. Existing data collection methods do not possess the necessary sophistication required to capture the dynamic behavior of commercial vehicle operations illustrated in these figures.



(a) Spatial trends in truck body class **(b) Temporal trends in truck body class**

Figure 2-3 Examples of spatial and temporal trends in truck body class.

The needs for body class data of commercial vehicles are twofold. First, body class data is needed to fill critical gaps for existing transportation programs. Freight transportation planning programs rely heavily on the results of the Vehicle Inventory Use Survey (VIUS), but with the discontinuation of that resource, a critical gap has been opened and a replacement data source is desperately needed. Second, body class data collected at the link and route level presents an increased level of detail that has yet to be captured by any other data source. The previous figures clearly demonstrate that body class varies significantly by location and time of day, so ignoring this level of detail can lead to significant modeling errors.

Body class data can allow agencies to further develop existing models for emissions estimation or freight forecasting by replacing existing sources that are either inaccurate or lacking in necessary detail or to create new models designed to make full use this new data source. Consequently, better models will lead to more effective management of transportation facilities, and improved confidence in estimates of emissions and air quality. Although emissions models are currently not designed to harness truck body classification data as input, their inclusion in future models will likely yield significant improvements in emissions estimations due to the improved fidelity of truck characterization and estimates of their activity. Since trucks can been related to industry through body classification, agencies can design programs to reduce emissions more effectively by designing specific policies that will yield maximum

benefits by targeting specific industries that generate high-emitting truck movements. Additionally, in line with freight transportation planning data needs, body class information will help to distinguish between long and short haul movements. Through integration with WIM data, it will also be possible to associate empty movements by industry and/or commodity. This will lead to improvement in emissions inventory models.

2.4 Weigh-In-Motion Systems

Weigh-in-Motion (WIM) devices have been used since the 1980s to collect data for truck routing, pavement management and design, weight enforcement, traffic safety, and transportation policy (Nichols and Bullock, 2004). There are currently 106 WIM sites in California that perform continuous data collection, as shown on the map in Figure 2-4. Two main types of WIM controllers are currently deployed in the State of California: the earlier DOS-based 1060 series controllers (Figure 2-5a) and the current Linux based iSinc family of controllers (Figure 2-5b), which include the iSinc WCU-II and iSinc WCU-3 Lite.

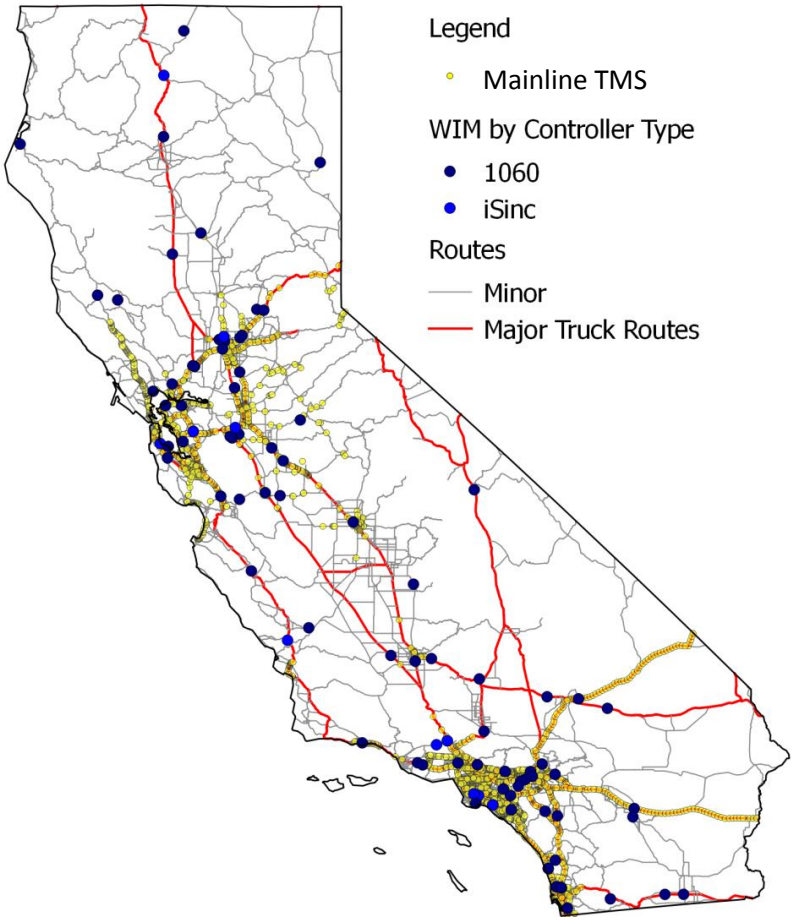


Figure 2-4 WIM and TMS Sites in California

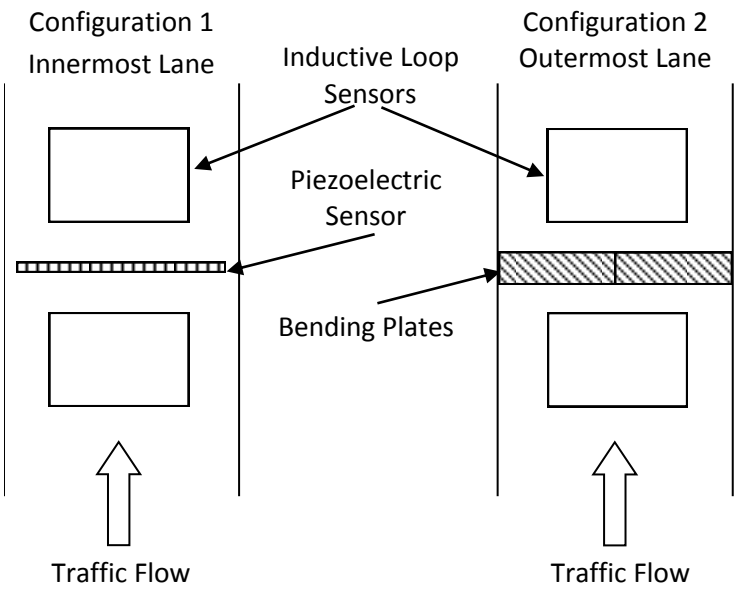
(a) 1060 series WIM controller

(b) iSinc Lite WIM controller

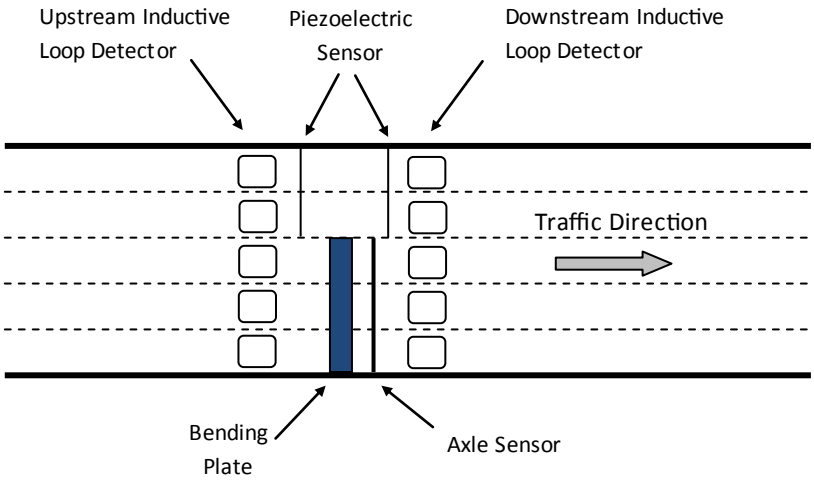
Figure 2-5 Types of WIM controllers deployed in California: (a) 1060 series and (b) iSINC Lite

The main distinction between the controllers for the purpose of this work is in their built-in ability to process inductive signature data. The loop sensor module (LSM) of the 1060 WIM controllers is designed only to obtain conventional bivalent inductive loop data. On the other hand, the LSM of the iSINC Lite controller has the ability to obtain inductive signature data. The caveat for the iSINC controller however, is that inductive signature data is currently designed only for diagnostic and troubleshooting purposes. Hence, the inductive signature data can only be manually logged when the system is in diagnostic mode, and is not currently available as an operational feature within the system. As WIM controllers become damaged and require replacement, older controllers are replaced with the new iSinc models. This is a great advantage for the work developed in this project because this essentially means that all WIM controllers will be capable of producing signature outputs.

A typical WIM station, as depicted in Figure 2-6, includes bending plates or pressure sensors straddled by square inductive loop detectors in the outermost lanes and piezoelectric sensors straddled by inductive loops in the innermost lanes. Figure 2-6b depicts a five lane highway sensor site in which the outer three lanes are equipped with bending plates while the inner two lanes, used mostly by passenger vehicles, are equipped with piezoelectric sensors.



(a) Typical inner and outer lane WIM configurations



(b) Example of site configuration for five lane highway

Figure 2-6 WIM Site Configuration

WIM stations collect vehicle arrival time and date, axle weights and gross weight, axle spacing, and speed (Lu et al., 2002). Vehicle classification is determined from the number of axles, axle spacing, and weight according to the classification sieve for FHWA Scheme F which includes 13 axle-based classes, or for California, the 14 class modified axle-based scheme. A basic decision tree approach divides vehicles into FHWA classes based on number of axles and inter-axle distances. This is approach can lead to classification error for certain vehicle classes since many of the axle counts and distances overlap. Reported errors in classification range have been shown to be as high as 9.5% (Kwigizile et al., 2005). In Chapter 5, improvements to the standard FHWA classification decision tree are introduced by addition variables such as length and axle spacing ratios to further define certain classes.

Agencies using WIM data are aware that WIM data is prone to accuracy errors in speed, spacing, and weight measurements (FHWA, 2001). The inaccuracies are the result of several possible factors: (1)

vehicle dynamics such as speed, acceleration, tire condition, load, and body type; (2) site conditions such as pavement smoothness; (3) environmental factors such as temperature and precipitation (Lee, 1998; NCHRP, 2008). Prozzi et al. (2007) modeled the load errors as systematic and random where random error is a result of statistical fluctuations in estimation which can be over- or under- estimations of the true value. On the other hand, systematic errors are persistent inaccuracies in which the true value is either consistently over- or under-estimated. Through proper calibration procedures such as those outlined in American Society of Testing and Materials (ASTM) Standard E1318-02 (ASTM, 2009) and National Cooperative Highway Research Program Synthesis 386 (NCHRP, 2008), systematic error can be addressed, but random disturbances in the data will persist regardless of calibration.

2.5 Inductive Signature Technology

Inductive loop detector technology has been used since the 1960s. An inductive loop detector consists of several coils of electrified wire embedded beneath the pavement and connected to a roadside control unit in which loop detector cards process the inductive magnitude changes to measure vehicle presence. The use of inductive loop signature technology for classification was introduced by Pursula and Pikkarainen in 1994. There are approximately 25,000 inductive loop detectors in California at around 8,000 mainline Traffic Monitoring Sites (TMS) as shown in Figure 2-4. TMSs are generally located in or near metropolitan areas.

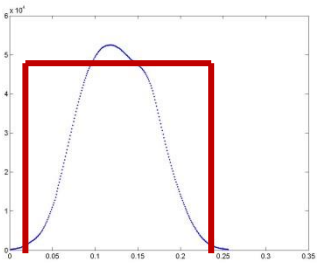
Conventional inductive loop detectors (ILDs) measure bivalent signals from inductive loops embedded in the pavement and are capable of measuring aggregated volumes and occupancies. The red lines in Figure 2-7 depict the bivalent outputs, e.g. [0,1], of a conventional loop detector. Unlike many other detector systems such as imaging or acoustic sensors, loop detectors are inherently accurate, achieving the best volume count accuracy compared with other common detection technologies, providing a good technology platform to develop the proposed system. Finally, ILDs are generally robust because magnetic inductance is not affected by changes in temperature, lighting, visibility and humidity.

Advanced inductive loop detectors measure the inductance change in an inductive loop sensor at rates of up to 1200 samples per second (IST, 2006), producing analog waveform outputs, referred to as inductive signatures, for each traversing vehicle. A significant advantage of advanced inductive loop detectors is that they can replace conventional detectors without altering the system's intended functions (e.g. occupancy and volume measures).

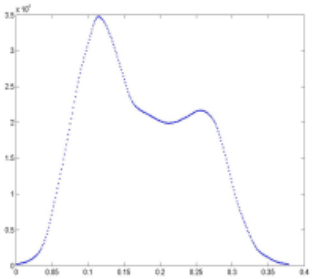
In addition to vehicle classification, inductive signature technology has been used to develop other advanced traffic monitoring applications. These include the accurate estimation of vehicle speeds from only single inductive loop sensors (Tok et al, 2009), as well as anonymous tracking of vehicles between adjacent ILD sites to obtain traffic measures such as section and corridor travel time, and section-based density (Jeng et al, 2010, Hernandez et al, 2013).

Samples of inductive signatures of various vehicle types are presented in Figure 2-7. The shape of the signature is the result of the ferrous components of the vehicle with the overall duration of the signature correlated with the length of the vehicle. Comparing passenger car and semi-tractor trailer signatures, it is easy to distinguish the general shape difference between the signatures. The differences in signature shapes of vehicle within the same axle class are more subtle where spikes and valleys generally relate to the undercarriage or chassis of a trailer but can also be the result of refrigeration tanks or axle placement, for example.

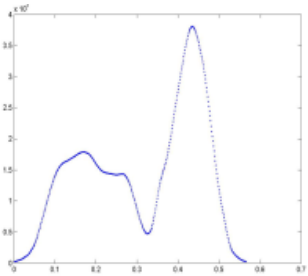
Passenger Car



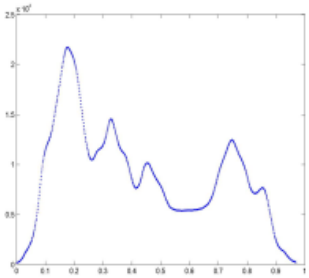
Single Unit Truck



Single Unit Truck with Trailer



Semi-Tractor Trailer (Enclosed Van)



Semi-Tractor Trailer (Livestock Van)

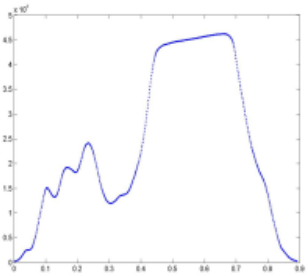


Figure 2-7 Examples of Inductive Signatures

2.6 Integrating Technologies

As shown on the left side of Figure 2-8, the outermost lanes at a typical WIM site are equipped with bending plates or pressure sensors to measure axle weights while inductive loop detectors straddle the weight sensors to detect vehicle presence. WIM sites measure speed, volume, truck weight, axle spacing, and length, however, axle-based information cannot be directly associated with a truck's function or body configuration. Advanced ILDs measure inductance change in an ILD up to 1,200 samples per second producing analog waveform outputs, called inductive signatures, which strongly correlate with vehicle body type. A significant benefit of advanced ILD technology is that it requires no in-pavement infrastructure upgrades thus, implementation costs are minimal. The WIM system can be equipped with advanced ILD technology by simply swapping out detector cards in the WIM controller with advanced signature capable ILDs. Test deployments show that the modification does not alter the WIM site functionality, so regular data reporting requirements from WIM site are not affected. Further details on the hardware integration are described in Chapter 5.

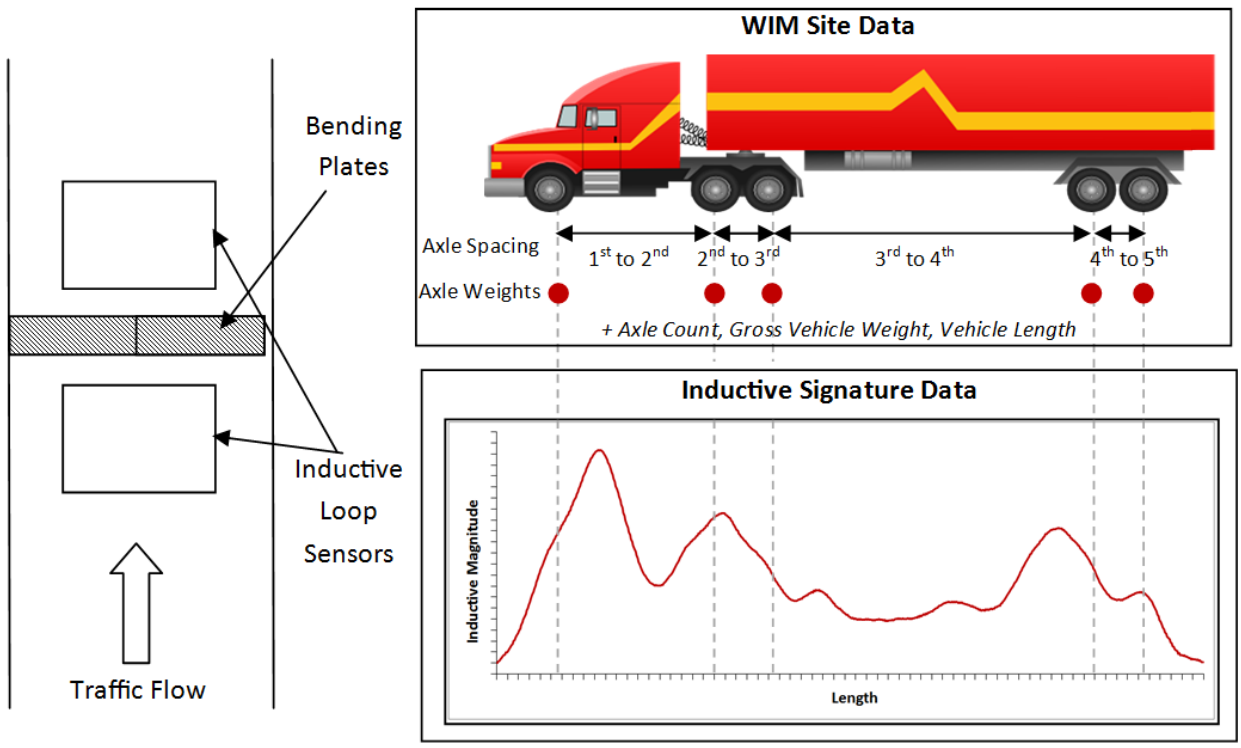


Figure 2-8 WIM Site Configuration (left) with data outputs from Advanced ILDs and WIM (right).

3 Data Collection and Processing

3.1 Overview

The purpose of this effort was to collect and process a large dataset of truck data. This addressed a major drawback of previous vehicle classification modeling efforts due to the lack of significant data for model training and testing. This is especially true for commercial vehicles which on average account for only around 10% of the total highway vehicle population (Caltrans, 2012) thus requiring extended data collection periods or selection of specific truck routes with high volumes of truck traffic to obtain sufficiently large samples. An additional complication is that unlike passenger vehicles, commercial vehicle body types vary by location and are influenced by local industry and land uses. For example, intermodal container trailers are more prevalent nearer to port areas while trucks with logging trailers appear in regions tied to the logging industries. Smaller commercial vehicles such as single unit trucks may also vary in body type by location. For instance, service related body types like garbage trucks or firetrucks might be observed in higher volumes near busier urban areas compared to remote rural sites. Significant efforts were made in this project to collect and process an abundant sample of commercial vehicles for model training and testing. This, in part, helped to facilitate more advanced modeling efforts and expand the predictive capabilities of the developed models. Care was taken to select data collection locations and seasonal time periods that would capture the extreme diversity of commercial vehicle body types. In total, around 35,000 vehicle records were captured and processed from disparate Weigh in Motion (WIM) and inductive loop detector (ILD) site locations in California over a time period spanning the fall, winter, and spring seasons.

3.2 Data Sources

In addition to inductive loop signature and WIM data, still image data was collected for each passing vehicle by connecting a digital camera with a remote trigger to the ILD detector card. Figure 3-1 depicts the hardware configuration including the still image camera connected to the inductive signature cards. With this configuration, the loop activation triggered the camera and a series of still images were captured for each passing vehicle at a rate of three frames per second while the vehicle was over the loop. Figure 3-2 shows a typical data collection hardware setup at WIM site. The traffic cabinet on the right contains the WIM controller and ILD signature cards. The still image camera shown to the left of the cabinet connects to the ILD signature detector cards within the cabinet. The field PC clock which set the timestamps for the ILD signature and photo data was synchronized against the WIM controller clock in order to ease the data groundtruth process that required the processor to link the WIM record, inductive signature, and photo together. Figure 3-3 shows examples of still images captured for longer combination trucks and a passenger vehicle.

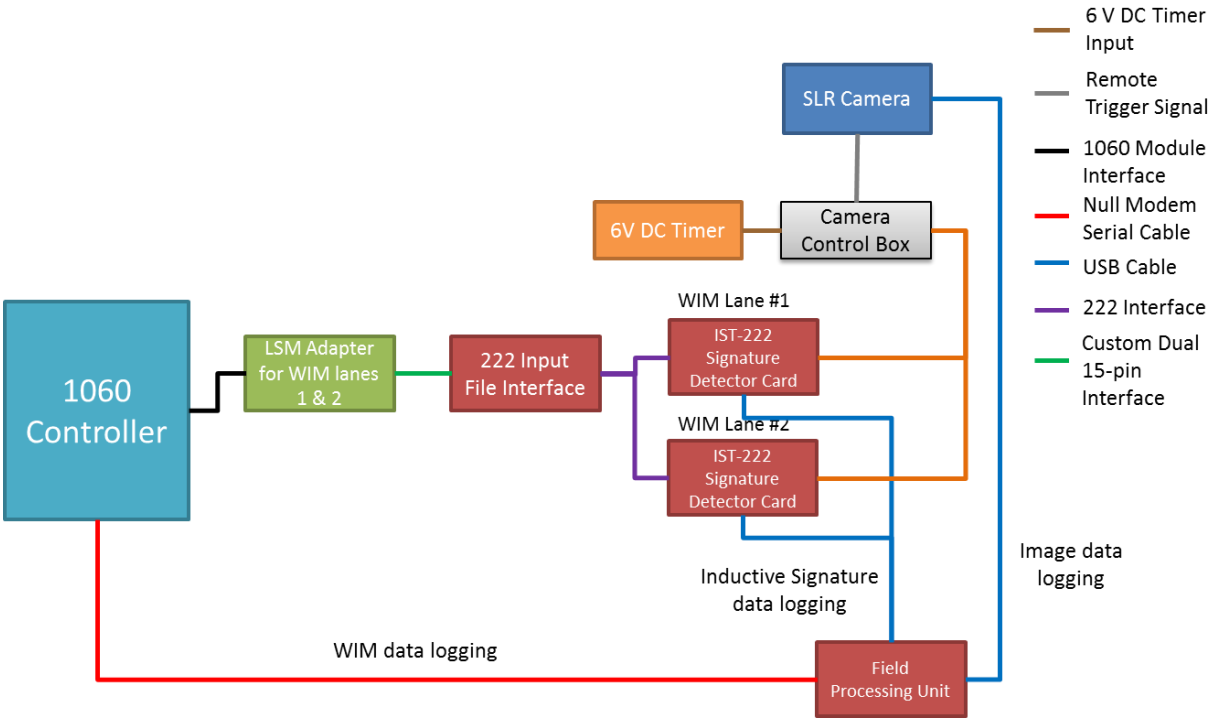


Figure 3-1 Data Collection System Architecture with SLR Still Image Camera

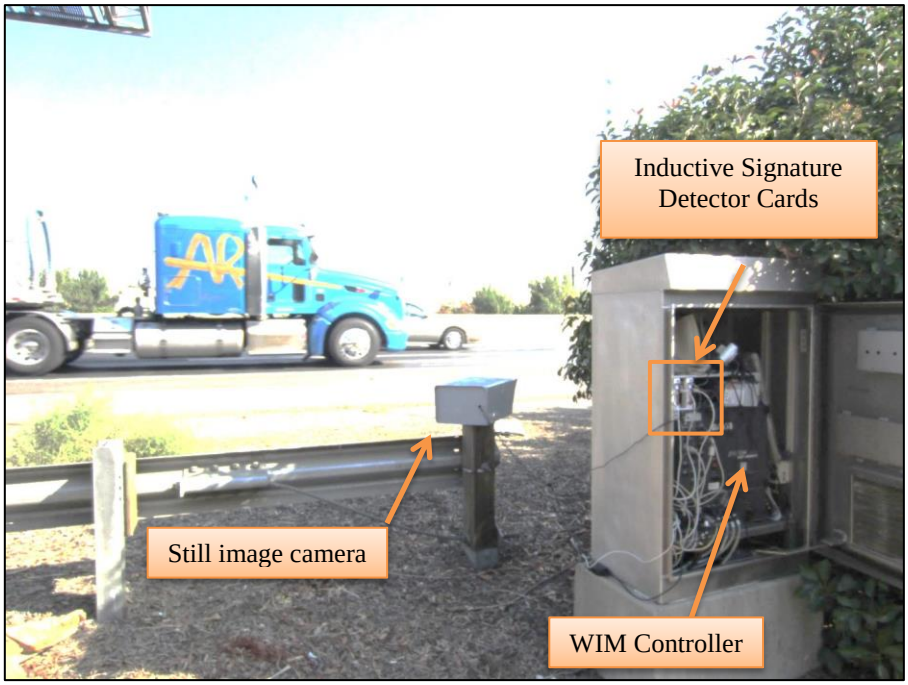


Figure 3-2 Data collection setup at WIM site

(a) Tractor-Trailer Combination truck in two still images

(b) Tractor-Trailer Combination truck in three still images

(c) Pickup trucks in single still images

(d) Single unit truck in two still images

Figure 3-3 Examples of still images collected for various truck types

3.3 Data Collection Sites

Data collection was performed at seven sites across the State of California. Figure 3-4 shows the locations of the seven data collection sites. The site name, data type, date of collection, and number of collected truck samples are summarized in Table 3-1. The WIM site along southbound (SB) I-5 in Irvine, was used for testing and development of the hardware interface configuration between the WIM controller and the inductive signature detector cards. Sites at Redding, Willows, and Fresno were selected to increase the diversity of the truck body types for model development. For example, sites in

northern and central California were expected to have a greater presence of logging trucks and agricultural trucks that would not typically be found at WIM stations closer to UC Irvine such as Irvine and San Onofre. Sites at San Onofre and Leucadia were selected for use in Phase 2 to develop the propagation methodology by which detailed classification based on WIM data can be propagated to ILD sites equipped with inductive loop signature capabilities. All sites other than the Saigon site contained 1060 series WIM controller units which require use of the configured hardware interface. The Saigon site along I-405 uses a newer WIM controller system called iSinc Lite which has the capability to collect inductive signature data without additional detector hardware. Data was collected at this station to establish a comparison between inductive loop signatures produced by various controller types, i.e. inductive signature detector cards at 1060 controllers and iSinc controllers. Diagrams of each site including the lane configurations and detector positions can be found in Appendix B.

Figure 3-4 Data Collection Sites

Table 3-1 Summary of Data Collection Sites

Site Name	Irvine	Fresno	Willows	Redding	Leucadia	San Onofre	Saigon*
Location	SB I-5 Southern California	SB SR-99 Central California	NB I-5 Northern California	SB I-5 Northern California	SB I-5 Southern California	SB I-5 Southern California	NB I-405 Southern California
Description	Urban, Approx. 45mi from San Pedro Bay Ports	Semi-Urban, Agricultural	Rural	Rural, Approx. 120mi from OR-CA border	Semi-Urban	Semi-Urban	Urban
Collection Dates	Sept. 21, 2012 Oct. 2-3, 2012 March 20-25th, 2013	Nov. 7-8, 2012	Dec. 10-12, 2012	Dec. 10-12, 2012	Jan. 9-10, 2013	Jan. 9-10, 2013	Oct. 9, 2012
WIM Site Number	15	10	108	2	84	810	112
Controller Type	1060 WIM / TMS	1060 WIM	1060 WIM	1060 WIM	1060 WIM	PrePass 1060 WIM	iSinc WIM and TMS
California Postmile	R25.8	25	R10.9	R24.9	R42.2	R68.4	18.6
Total Lanes	5 SB	3 NB, 3 SB	2 NB, 2 SB	2 NB, 2 SB	2 SB	2 SB	5 NB
No. Lanes for data	2 SB	2 SB	2 NB	2 SB	2 SB	2 SB	2 NB
Approximate Total Truck Percentage[1]	5%	22%	25%	25%	5%	7%	4%
Total Samples Collected	6,963	9,718	6,908	5,110	4,017	1,906	97

[1] Percent of total traffic, Source: Caltrans Traffic Counts for AADTT (Caltrans, 2012)

* For signature comparison between different controller systems

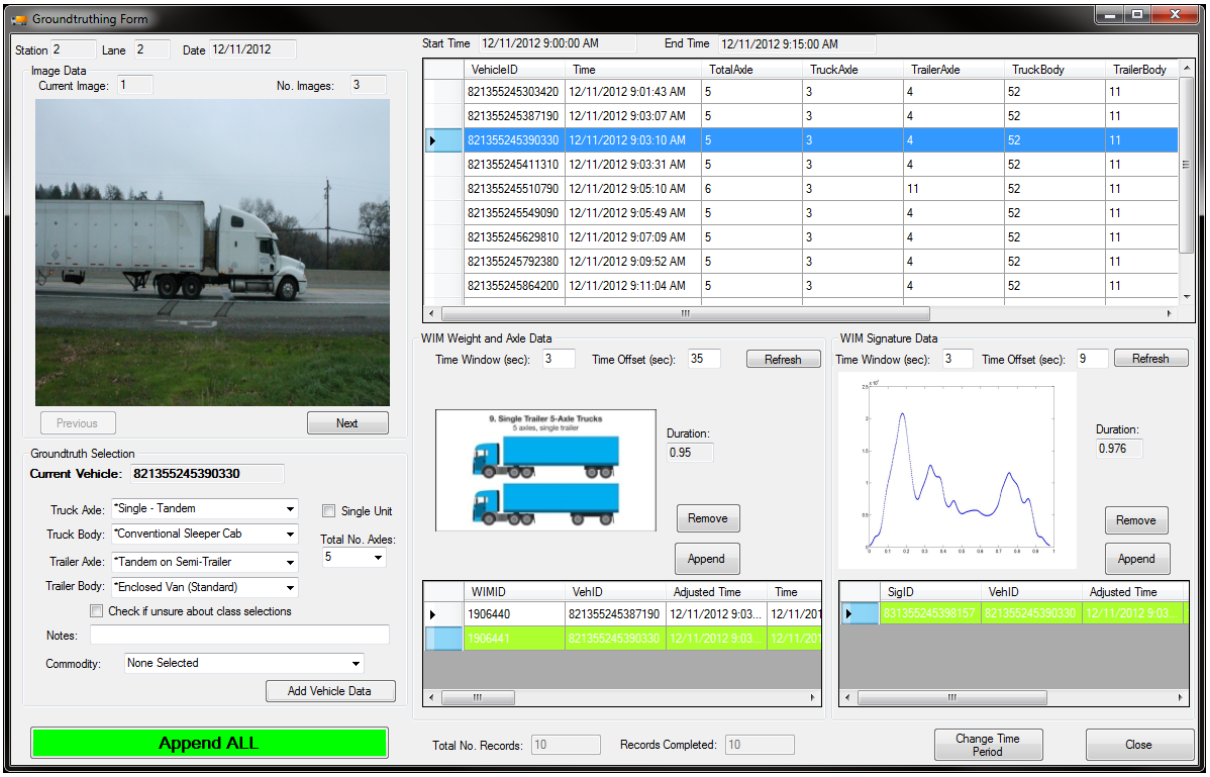
The main data used for body class model development was collected at Irvine, Fresno, Willows, and Redding WIM sites over several two to three day periods during the fall, winter, and spring seasons between 2012 and 2013 as summarized in Table 3-2. The data collection was conducted over 13 weekdays spanning a total of 97.25 hours. The majority of the data were collected during uncongested conditions. The median speeds are between 50 to 55 miles per hour across all sites. The Irvine site experienced a brief period of minor congestion on October 3rd with speeds below 50 mph. The Fresno site also experienced a brief period of congestion during the November 8th time period.

Table 3-2 Summary of Data Collection Time Periods

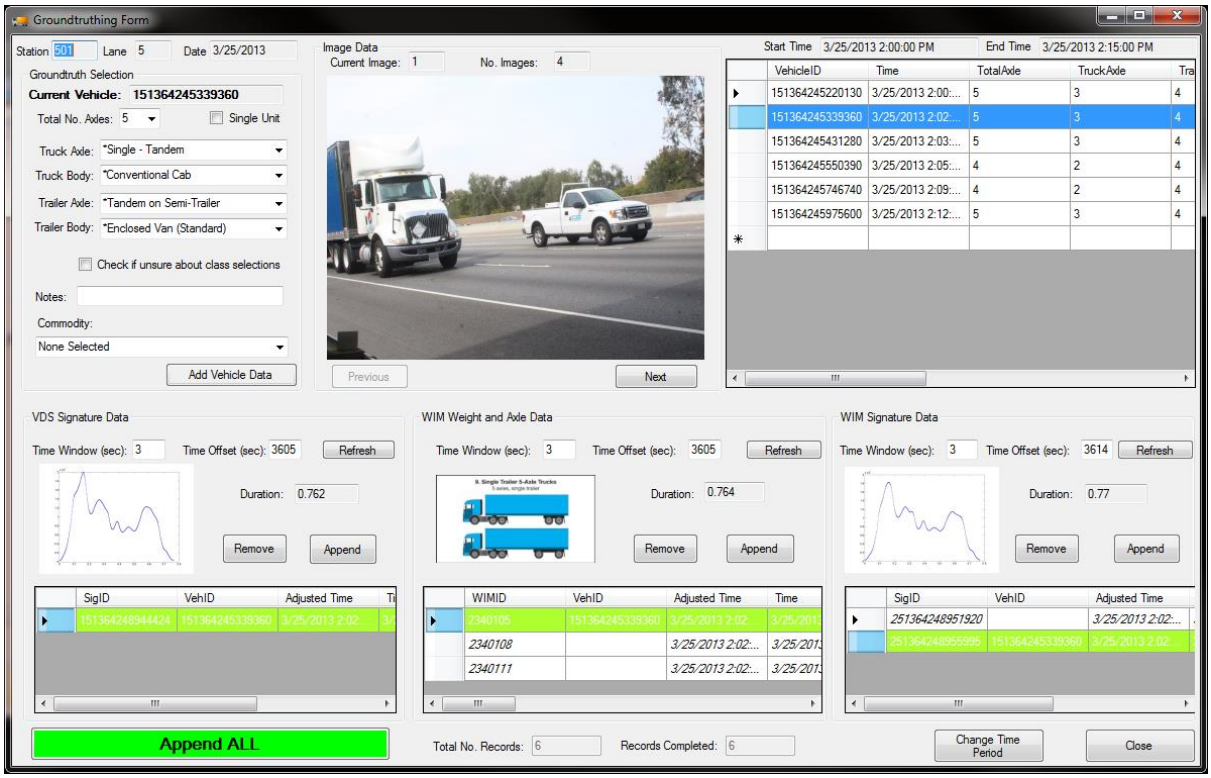
Site	Date	Day of Week	Season	Time Period	Total Hours	Average Speed (mph)
Irvine	Sept. 21, 2012	F	Fall	10:45AM – 6:00PM	7.25	56.4
	Oct. 2nd, 2012	T	Fall	1:00PM – 6:45PM	5.75	57.7
	Oct. 3rd, 2012	W		6:30AM – 9:15AM	2.75	48.2
	March 20th , 2013	W	Spring	6:30AM – 7:45PM	12.25	61.9
	March 25th, 2013	M		7:30AM – 4:15PM	8.75	59.1
Fresno	Nov. 7th, 2012	W	Fall	10:15AM -5:15PM	7.0	57.6
	Nov. 8th, 2012	T		6:15 AM – 4:45PM	10.5	56.8
Willows	Dec. 10th 2012	M		10:30AM – 4:45PM	6.25	62.1
	Dec. 11th,2012	T	Winter	7:15AM - 4:45PM	9.5	59.0
	Dec. 12th,2012	W		7:00 AM – 3:00 PM	8.0	61.9
Redding	Dec. 10th 2012	M		1:30 PM – 5:00 PM	3.5	57.8
	Dec. 11th,2012	T	Winter	7:00 AM – 4:45PM	9.75	57.7
	Dec. 12th,2012	W		7:00 AM – 1:00PM	6.0	58.8
Total	13 days				97.25	58.5

3.4 Data Processing

The data processing procedure, referred to as the groundtruth system, involved preprocessing the WIM, inductive signature, and still image data, linking the three data types, and identifying the vehicle configuration and body type from each photo record. A database was designed and developed in PostgreSQL to store and integrate inductive loop signature, WIM record, and still image data. A specially developed software user interface was developed in Visual Basic to efficiently process and classify collected vehicle samples. The user interface was designed to communicate with the database and allow the user to scroll through photos and select the vehicle class parameters while also linking inductive loop signature records and WIM data records to the appropriate vehicle record. Figure 3-5 shows two versions of the custom groundtruth system user interface, showing data collected from a WIM only in Redding and a paired ILD and WIM site in Irvine. Both interfaces display navigable still image, inductive signature and WIM data, with an additional data entry panel to enter detailed vehicle information. The ILD and WIM interface has an additional inductive signature display that corresponds to data collected from the adjacent ILD site. Five components of vehicle class information is recorded in this interface: truck axle configuration, trailer axle configuration, truck body configuration, and trailer body configuration, as well as total number of axles. A schematic of the database structure for the data groundtruth system in included in Appendix C.



(a) Customized User Interface for WIM Sites



(b) Customized User Interface for Combined WIM and ILD Sites

Figure 3-5 Customized User Interfaces for Data Groundtruth

3.5 Vehicle Classification Scheme

A key component in the development of the body classification models in this study was the creation of a comprehensive classification scheme that captures the diversity of truck bodies found in the data. An initial body class scheme was developed based on the VIUS defined body types (VIUS, 2002). VIUS was selected as it provided a high level of detail regarding body classification. VIUS separates trucks into three vehicle configuration groups: Passenger vehicles, single unit trucks, and semi-tractor trailer combination trucks. Body classes corresponding to trucks were further expanded based on observed field data.

3.5.1 Single Unit Trucks

Single unit trucks are defined by VIUS as any truck with or without a trailer that is not a truck or road tractor. Figure 3-6 illustrates examples of single unit trucks without trailers (a & b) and with small trailers (c). VIUS defines 22 single unit body classes.

Table 3-3 summarizes the five main body categories: vans, platforms, tanks, service, and specialty vehicles. The existing VIUS body classes were expanded into 28 body classes. The VIUS body class, 'Vans, insulated, non-refrigerated' are not visually distinguishable from basic enclosed vans, and were therefore grouped together, then separated across four body classes for modeling. The VIUS body class 'Concrete pumpers' was not found in the data, and was not included in model development. Body classes identified in the data but not defined in VIUS include pneumatic tanks, livestock trucks, and firetrucks. Additionally, the three body classes including two buses categories and recreational vehicles (RVs) are not included in the VIUS classification scheme, and were added to the model classification scheme. These modifications to the VIUS scheme resulted in a total of 31 body classes for single unit trucks used for model development.

(a) Enclosed van

(b) Trash, Garbage, or Recycling

(c) Enclosed Van with small trailer

Figure 3-6 Examples of Single Unit Trucks

Table 3-3 VIUS and Model Single Unit Trucks Body Classification

Body Category	VIUS Body Class	Model Body Class
Van	Van, basic enclosed (dry cargo) Van, insulated non-refrigerated	Conventional Enclosed Van
		Light Duty Enclosed Van
		Low Loading Enclosed Van
		Cab-Over Engine Enclosed Van
	Van insulated refrigerated	Conventional Reefer Enclosed Van
		Cab-Over Reefer Enclosed Van
	Van, open top (including low-side grain, fruit, potato bed, etc.)	Open Top Van
	Van, step, walk-in, or multistep (including hi-cube or cutaway)	Multi-Stop or Step Van
Tank	Tank, dry bulk	Tank
	Tank, liquids or gases	
	Vacuum	
	**	Pneumatic Tank
Platform	Flatbed (including any with added devices), stake, platform, etc.	Basic Platform
		Low Boy Platform
Service	Service, utility (telephone line, cable, pipeline, etc.)	Utility
	Service, other (mobile workshop, craftsman's vehicle", etc.)	
Specialty	Armored	Armored
	Beverage	Beverage
	Concrete mixer	Concrete Mixer
	Concrete pumper	*
	Crane	Winch or crane truck
	Curtainside	Curtainside Van
	Dump (including belly or bottom dump)	End Dump
		Bottom Dump
		Dumpster Transport
	Pole, logging, pulpwood, or pipe	Pole, logging, pulpwood, or pipe
	Street sweeper	Street sweeper
	Tow/Wrecker (including flatbed)	Platform for auto transport
		Wrecker
	Trash, garbage, or recycling	Garbage
	**	Livestock
	**	Firetruck
Other	Other	Other
Bus	**	Recreational Vehicle (RV)
	**	30ft Bus
	**	20ft Bus

* Not included in model body classification scheme
** Not included in VIUS

3.5.2 Combination Trucks

Combination trucks refer to multi-unit trucks that include single unit trucks pulling single or small trailers (Figure 3-7 a & b) as well as semi-tractor trucks pulling a semi-trailer or multiple trailers (Figure 3-7c & d). In the case of single unit trucks pulling a trailer, the drive unit is categorized by the drive unit body type classification as defined in the previous section, while the trailer body classification follows the scheme described in this section. Similarly, a semi-tractor combination truck pulling one or more trailers is defined in two components: the tractor unit and trailer unit(s). The classification schemes for both these components are presented in this section.

(a) Single unit truck with single trailer

(b) Single unit truck with small trailer (recreational vehicle)

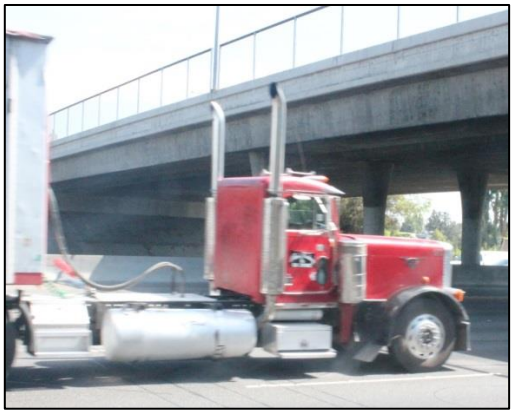
(c) Enclosed Van Semi Trailer

(d) Belly Dump Multi- Trailer

Figure 3-7 Examples of Combination Truck Trailer Body Classes

3.5.2.1 Tractor Units

The tractor unit refers to the drive unit of a semi- tractor trailer combination truck as shown in Figure 3-8 and is divided by VIUS into four standard engine configuration types described in Table 3-4. Conventional and cab-over engine configurations are further specified as sleeper (Figure 3-8 a and b) or non-sleeper (Figure 3-8 c and d). 'Cab forward engine' and 'cab beside engine' body classes defined in VIUS were not found in the data and are not included in the model scheme. Lastly, two specialty cab types were found in the data including cabs with attached cranes and cabs for auto transport (Figure 3-8 e) and included in the body classification scheme for model development. This yields a total of six tractor unit body classes included in the model classification scheme.



(a) Conventional cab without sleeper

(b) Cab over without sleeper

(c) Conventional cab with sleeper

(d) Cab over with sleeper

(e) Specialty cab with crane

Figure 3-8 Examples of Tractor Units

Table 3-4 VIUS and Model Drive Unit Body Classification

Body Category	VIUS Body Class	Model Body Class
Standard Tractor Units	Conventional cab with or without sleeper	Conventional cab
		Conventional with sleeper cab
	Cab over engine with or without sleeper	Cab over engine
		Cab over with sleeper cab
	Cab forward engine	*
	Cab beside engine	*
Specialty Tractor units	**	Cab with attached crane
	**	Cab for auto transport

* Not included in model body classification scheme
** Not included in VIUS

3.5.2.2 Trailer Units

As shown in Table 3-5, VIUS defines 17 trailer body types which can be grouped into four main categories: van, tank, platform, and specialty. The 17 body classes were expanded into 27 body types for the model classification scheme as shown in the rightmost column of Table 3-5. VIUS classes for basic enclosed vans and insulated non-refrigerated vans were combined because they could not be visually distinguished. This group was divided into vans with and without side skirts (e.g. 'skirted enclosed van) which are aerodynamic panels attached to the bottom sides of the trailer to decrease airflow through the trailer undercarriage. Further added categories include hoppers and agricultural vans.

Major additions to the VIUS body classification scheme include intermodal containers and smaller trailers. Intermodal container trailers are not included in VIUS since these belong to shippers rather than carriers or operators who participated in the VIUS survey. These trailers were further distinguished into the following categories: intermodal container chassis, 20ft, 40ft, and 53ft intermodal containers, and 40ft intermodal refrigerated containers. The distinction of these categories has significant implications on freight activity analysis. This is because 40ft containers are primarily used at sea port facilities, while 53ft containers are used only for domestic freight movement. In addition, refrigerated containers are typically used for hauling perishable goods. Hence, the movement of such containers may result in additional emissions and fuel usage, since refrigeration units on these containers need to be operated continuously. For the purpose of this research, small trailers may be pulled by a single unit truck or a tractor. The VIUS trailer body classes were expanded to include small trailer body types such as RV trailers, towed vehicles, and small dolly trailers.

Table 3-5 VIUS and Model Body Classification for Semi-Trailers for Existing VIUS Classes

Category	VIUS Body Class	Model Body Class
Van	Van, basic enclosed (dry cargo)	Enclosed van
	Van, insulated non-refrigerated	Skirted enclosed van
	Van, drop frame (excluding livestock)	Drop frame van
	Van, insulated refrigerated	Reefer enclosed van
Tank	Tank, dry bulk Tank, liquids or gases	Hot product tank
		Deep drop tank
		Food grade tank
		Petroleum tank
		Chemical tank
		Crude oil tank
		Air compression tank
		Propane tank
		Pneumatic Tank
Platform	Flatbed, platform, etc.	Basic platform
		Platform with devices
	Low boy (platform with depressed center)	Low boy platform
Specialty	Dump (including belly or bottom dump)	Bottom/Belly dump
		Bulk waste transport
		End dump
	Livestock (including livestock dropframe)	Livestock
	Curtainside	Curtainside van
	Mobile home toter	*
	Open tops (vans, low side grain, fruit, etc.)	Open top van
	Pole, logging, pulpwood, or pipe	Pole, logging, pulpwood, or pipe
	Automobile Carrier	Automobile transport
	Beverage	Beverage
	Trailer mounted equipment	*
	**	Hopper
	**	Agricultural van
Intermodal Containers	**	Container chassis
	**	40ft container
	**	40ft refrigerated container
	**	20ft container
	**	20ft container on 40ft chassis
	**	53ft container
Small Trailers	**	Recreational vehicle trailer
	**	Towed vehicle
	**	Small trailer/dolly

* Not included in model body classification scheme
** Not included in VIUS

3.6 Data Summary

FHWA class 5 single unit trucks and FHWA class 9 semi-trailer combination trucks were the most prevalent across all four data collection sites as summarized in Table 3-6. It should be noted that Table 3-6 reports lower than observed volumes of passenger vehicles. Passenger vehicles excluding larger

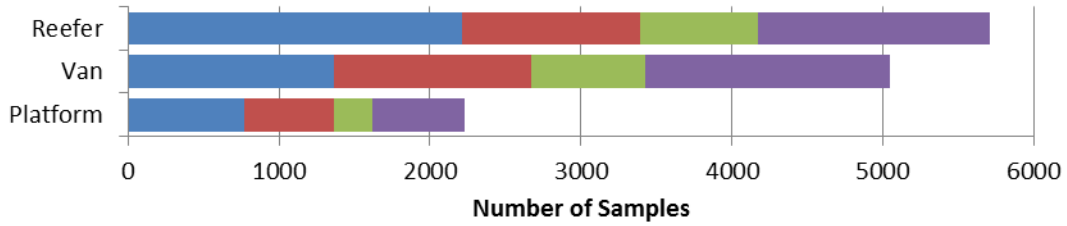
pick-up trucks or 12 passenger vans were not processed for the Fresno, Willows, or Redding data collection sites, or for the September and October data collection periods at Irvine to reduce the burden of processing the data groundtruth. All passenger vehicles, including sedans, SUVs, minivans, etc. are included in the March 20 and 25th data sets at the Irvine site.

Table 3-6 Volume by site and FHWA class

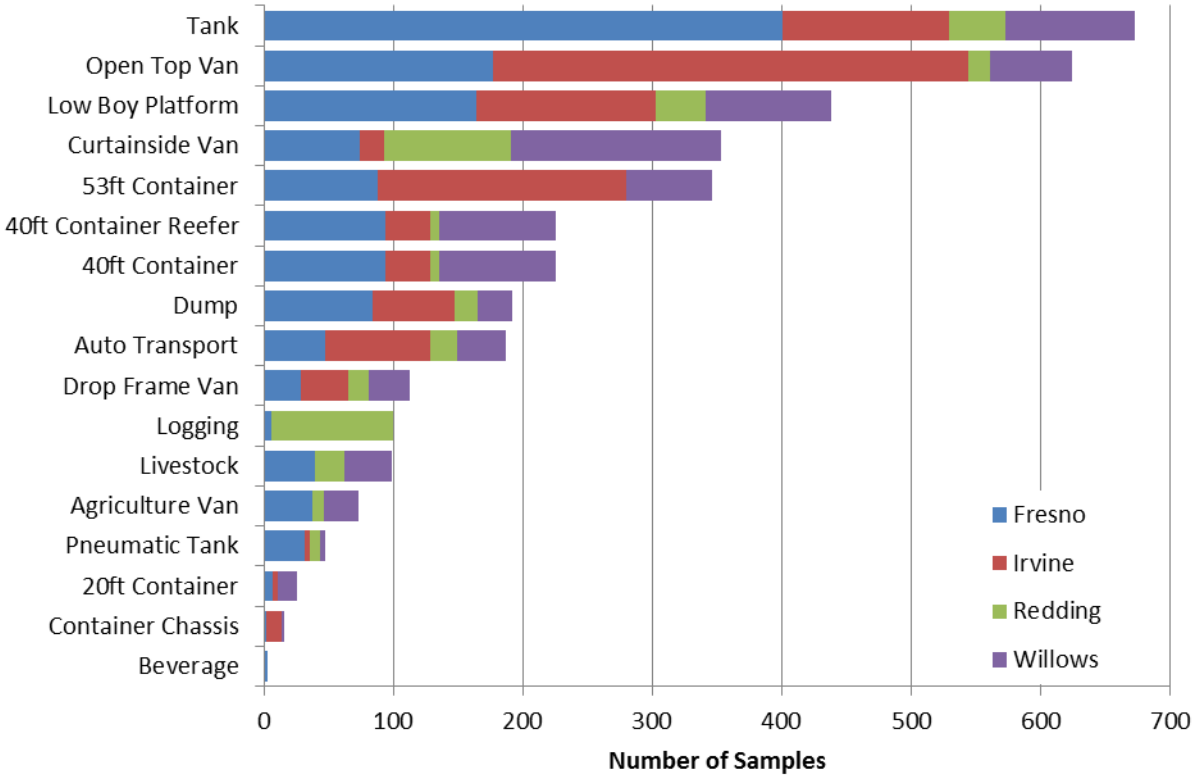
FHWA Class		Irvine	Fresno	Willows	Redding	Total
Passenger Vehicles	2	1,649	47	127	1	**1,824**
	3	4,345	745	1,187	22	**6,299**
Single Unit	4	176	58	11	8	**253**
	5	2,843	842	145	77	**3,907**
	6	515	238	64	64	**881**
	7	149	14	0	1	**164**
Single Trailer	8	455	261	82	112	**910**
	9	3,077	4,200	3,693	1,711	**12,681**
	10	11	20	7	7	**45**
	14	225	119	43	96	**483**
Multi Trailer	11	232	477	251	125	**1,085**
	12	12	63	65	44	**184**
	13	2	0	1	0	**3**
Other	15	148	422	88	88	**746**
Total		**13,839**	**7,506**	**5,764**	**2,356**	**29,465**

FHWA class 9 semi-trailer combination trucks exhibit one of the most diverse sets of body types. Therefore the summaries provided in this section focus on semi-trailer body types. There are a total of 12,681 processed vehicle records for semi-tractor trailer vehicle configurations in FHWA class 9. The most prevalent trailer body type is the enclosed van which comprises 65.1% of the data across all sites. Within the enclosed van body category, non-refrigerated vans represent 28.6%, refrigerated (e.g. reefer) vans represent 26.1%, and non-refrigerated skirted vans represent 10.4% of the total data. The second most populous semi-trailer body type is the basic platform which accounts for 9.5%, followed by tanks representing 4.8% of the total data. Figure 3-9 shows the number of samples for each semi-trailer body class across all four sites.

Enclosed vans, refrigerated enclosed vans, and platform semi-trailer body classes dominate the population across all four sites. Notable differences were observed in the number of tanks, open top vans, curtainside vans, 53ft containers, 40ft containers, and pole/logging/pipe trailers. These trailer body types are more industry specific and thus possess greater spatial variations due to diverse land uses across the data collection sites. For example, logging trucks were observed in higher proportion at the Redding site, which is located in a region with forestry-related industries. At the Irvine site, open top vans which transport garbage, refuse, and construction debris were observed in greater numbers as a result of urban land use.



(a) Majority Class Samples by Site



(b) Minority Class Samples by Site

Figure 3-9 Number of semi-trailer truck samples by body type across all sites

4 Model Development

4.1 Overview

This chapter presents all models that were developed as a part of this research study. These include the following:

i. Inductive signature-based truck body classification models,
ii. Truck volume estimation model,
iii. Optimal sensor location model, and
iv. Spatial propagation model

Two models were developed for truck body classification using inductive signatures. The first WIM-signature model is applicable at WIM sites retrofitted with inductive signature technology while the second inductive signature only model is applicable at ILD sites modified with inductive signature technology. The truck volume estimation model is used to estimate counts of trucks by body configuration at WIM sites not equipped with inductive signature technology as well as from archived WIM data to provide estimated counts of a reduced set of truck body categories compared with the signature-based truck body classification models. The optimal sensor location model was developed to identify the optimal locations for selecting sites for deploying the signature-based truck body systems. Lastly, the spatial propagation model allows truck weight data from existing WIM sites equipped with inductive signature technology to be extended to ILD sites with inductive signature technology.

4.2 Body Classification Models

Classification of trucks by body configuration is not a trivial task, given the wide diversity of body types within each vehicle configuration group. Furthermore, analysis of observed truck volumes by body type at the four data collection sites shows that each site contains a different distribution of truck body types, especially for minority body classes such as those that carry seasonal commodities such as agricultural vans, or vehicle classes with unique travel characteristics such as intermodal containers. In light of these observations, the body classification models developed in this project have the complex task of producing accurate classifications across a large array of body classes and generalizing across locations with varying distributions of those classes.

Previous vehicle classification models using inductive signatures show that different classifier models have varied competencies at predicting certain classes and no one model produced high CCR for all classes. Hence, the multiple classifier systems (MCS) method was adopted for this truck body classification modeling effort to increase the classification accuracy for minority body classes as well as to enhance the spatial and temporal transferability of the model. MCS which combine the predictions of a series of individual classifiers, leverage the talents of each classifier while minimizing the tradeoffs that would occur if only a single classifier was implemented, and is therefore well suited for the truck body classification problem. This is a significant contribution of this research, since MCSs have not been considered in previous work with inductive signatures nor have they been widely used in the review of vehicle classification literature.

Dietterich (2000) gives three general reasons why combining multiple models may improve performance. First, from a statistical perspective, when the amount of training data is too small, the classifier can produce many different predictions. But the combined prediction of multiple classifiers might be closer to the most accurate prediction. Second, from a computational perspective, classifiers which rely on hill climbing or random search algorithms are subject to local optima. Thus, an ensemble approach where local search starts from varying points and performs an optimal search in different ways can have a better chance of approximating the global optimum. Third, from a representational perspective, it is possible that the true relationship between inputs and outputs cannot be represented by any of the classifiers. In this case, the true classifier lies outside the scope of the individual classifiers, so combining classifiers can help expand the space of representable functions. Although there is no guarantee that combining models will result in higher performance than the most accurate individual classifier, at least the statistical, representational, and computational 'risks' are reduced while the performance and generalization capability are increased (Dietterich, 2000).

4.2.1 Selecting Base Classifiers

The individual classifiers, or base classifiers, which comprise the MCS, can be modifications of the same classifier or a combination of different classifiers called hybrid ensembles (Brown, 2005). For example, a set of neural networks with different number of hidden layers constitutes an ensemble while a set of three different classifiers-neural network, decision tree, and support vector machine- constitutes a hybrid ensemble. An important consideration in selecting an optimal set of base classifiers is to ensure that each classifier exhibits different error characteristics. Brown *(2005)* suggests that hybrid ensembles are best suited to ensure error diversity since each classifier represents the problem and search space in "radically different" ways leading to different patterns of generalization and class specific prediction accuracy. Therefore, a hybrid ensemble representing both simple statistical classifiers and computational, non-parametric classifiers is adopted in this paper. The ensemble consists of a Naïve Bayes Classifier, Decision Tree (DT), Support Vector Machine (SVM), Multilayer Feed Forward Neural Network (MLFF), and a Probabilistic Neural Network (PNN). To ensure error diversity the selected classifiers vary in their learning function, definition and interpretation, and possess different assumptions about collinearity of input features. Each base classifier was reasonably optimized with regards to parameter settings through cross validation techniques. The Naïve Bayes and PNN classifiers do not require parameter tuning; for the MLP, appropriate network architecture, i.e. number of hidden layers and hidden nodes was selected; for the DT, the pruning parameter was optimized to avoid over fitting; for the SVM a linear kernel was chosen and the penalty term was selected to improve classification accuracy.

4.2.2 Combining Classifiers

Base classifiers can be combined through simple majority voting strategies or through more complex formulations that attribute weights to each base classifier when combining predictions. Based on experimental studies, there is no significant difference in prediction accuracy between the simplest and most complex combination methods (Kuncheva , 2004). However, the simplest strategies such as majority voting do not make full use of the known competencies of base classifiers- information which can help increase the overall prediction accuracy of the MCS. Thus, the Naïve Bayes Combination (NBC) technique, which incorporates the accuracies of each base classifier into the final prediction, was selected.

NBC combines base classifier predictions following from Bayes theory (Kuncheva , 2004). Instead of assigning a weight to each base classifier, a value of support is obtained for each class. The value of support for a particular class is derived from the joint probability distribution between that particular class and the class predictions provided by each the base classifiers. NBC assumes that classifiers are mutually independent given a class label. Conditional independence is represented by the confusion matrix, or cross classification matrix, cm^i, resulting from application of each classifier on validation data where each entry, *(k, s),* represents the number of samples with true class *k* and assigned class *s*. The class with the highest value of support is assigned as the final prediction of the ensemble. The value of support, μ_k, is:

$$\mu_k \propto \left\{ \prod_{i=1}^{B} \frac{cm_{k,s}^i + 1/c}{N_k + 1} \right\}^{\beta}$$

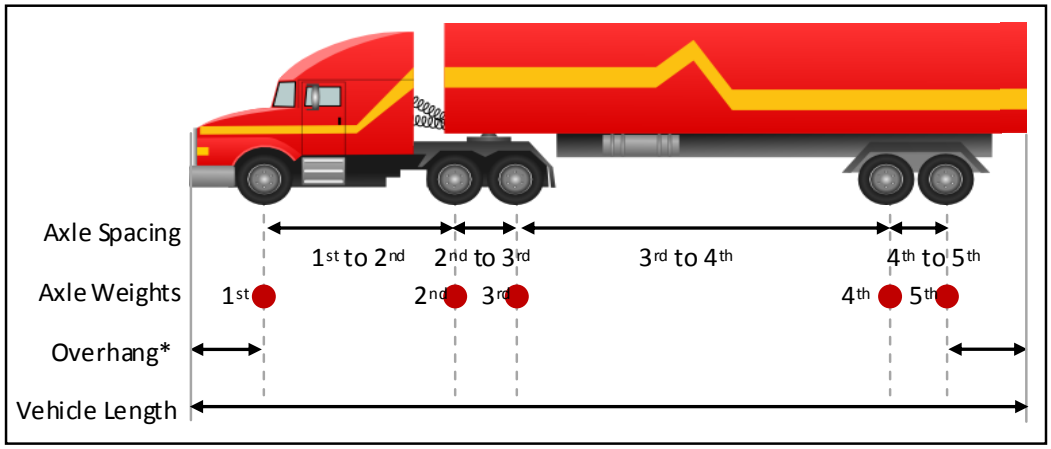
where

k = true class
s = assigned class
$cm_{k,s}^i$ = confusion matrix for base classifier i= 1…..B
N_k = proportion of samples in class k
c = number of classes
β = tunable constant

The formulation is adapted to account for zero values of $cm_{k,s}^i$ which left untreated would cause the calculation to be zero regardless of other classifier estimates. Titterington et al. (1981) suggests values for β between 0.5 and 1.

4.2.3 Model Inputs

The input feature sets comprised a combination of WIM measurements and inductive signature features. For the inductive signature only models presented in Section 4.2.4, the model inputs consist of only features derived from the inductive signature. For the WIM based models presented in Section 4.2.5, in addition to the inductive signature features, WIM features provide supplemental input to each model. WIM features include axle spacing and weights, vehicle length, and several derived features such as overhang and ratios between axle weights as shown in Figure 4-1. Overhang represents the front and rear portions of the vehicle extending beyond the axles, and is obtained as the arithmetic difference between the overall length and the sum of all axle spacing measurements. Weight ratios are calculated as the ratios between the steering, drive, or trailer axles.

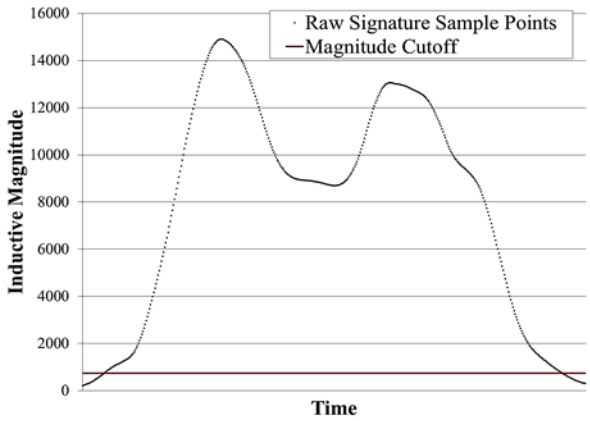


* Indicates derived measurement

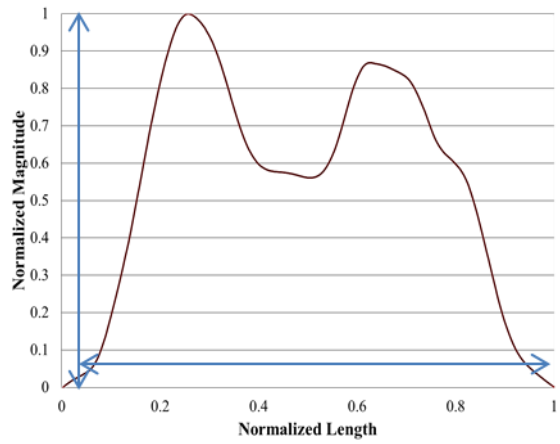
Figure 4-1 WIM system measurement and derived features for five axle semi tractor-trailer trucks

The signature feature set includes interpolated magnitudes and interpolated magnitude differences. Signatures are pre-processed as shown in Figure 4-2. First, the raw inductive vehicle signature is cleaned by applying a magnitude cutoff criterion to reduce measurement noise at the signature tails (Figure 4-2(a)). Second, the signature is normalized by its peak magnitude along the vertical axis and either the total duration of the inductive signature in the inductive signature only model or the overall vehicle length in the WIM-signature model along the horizontal axis (Figure 4-2(b)). Third, 20 to 30 equally spaced magnitudes are interpolated from the normalized signature using the cubic spline interpolation method (Figure 4-2(c)). A second feature set is computed as the difference between consecutive interpolated magnitude points (Figure 4-2(d)).

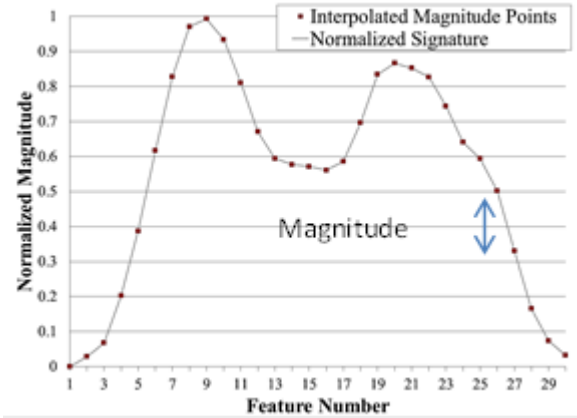
In the WIM-Signature model, for five axle semi-tractor trailer units, axle spacing measurements are used to parse inductive signatures into tractor and trailer portions, as shown in Figure 2-8. The parsed section of the signature representing the trailer is processed according to the same feature extraction procedure except the signature is normalized by the peak magnitude of only the trailer portion. This improves classification accuracy since the model will not be influenced by features from the tractor in predicting the body class of the trailer. The parsing procedure is only applied to FHWA class 9 five axle semi-tractor trailers.



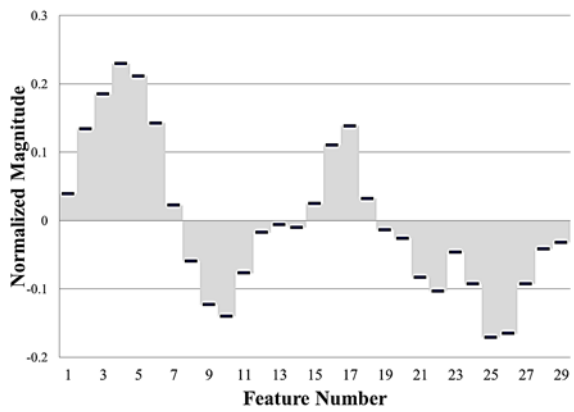
(a) Raw Signature



(b) Normalized Signature



(c) Interpolated Magnitude Features



(d) Magnitude Differences Features

Figure 4-2 Inductive signature feature extraction procedure

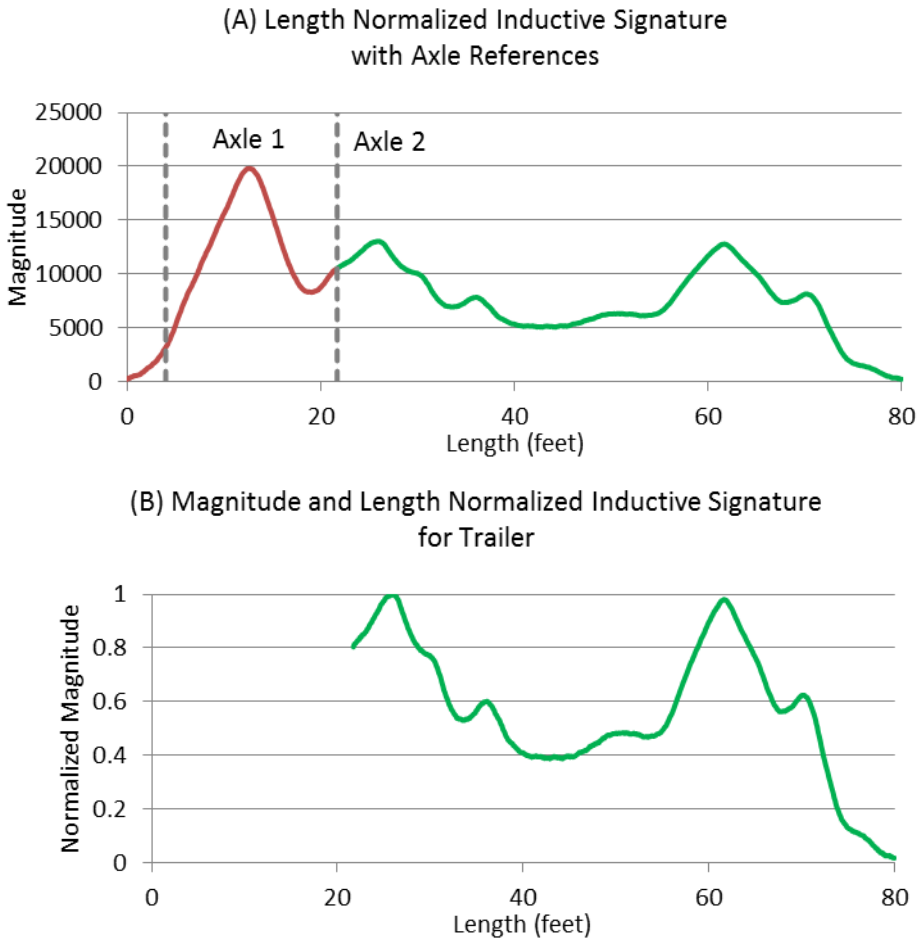


Figure 4-3 Inductive Signature Parsing using WIM axle spacing measurements

4.2.4 Inductive Signature Body Classification Model

The Inductive Signature Body Classification model requires only inductive signature data for model input and does not rely on WIM data. While the integrated WIM-Signature model is applicable at WIM sites equipped with inductive signature technology, the inductive signature only model in this section can be implemented at any ILD site equipped with 6-ft inductive loop sensors that has been upgraded with signature-capable ILD cards.

4.2.4.1 Model Structure

The inductive signature only model implements a three-tiered approach as shown in Figure 4-4. The first tier separates vehicles into two general vehicle configuration groups: single unit and multi-units. The second tier further divides the two groups into more specific body configuration groups. For single units, vehicles are classified as either passenger vehicles or single unit trucks. For multi-units, vehicles are classified as single units with trailers, semi-tractor single semi-trailer configurations, or semi-tractor multiple trailer configurations. The third tier consists of the body classification models for each body configuration group. All three tiers of the model use the same feature set comprised of interpolated magnitudes and magnitude differences derived from the normalized inductive signature. The first tier

was implemented as a decision tree. The second tier was implemented as a feed forward neural network with two hidden layers of 15 neurons. The third tier was implemented as a MCS.

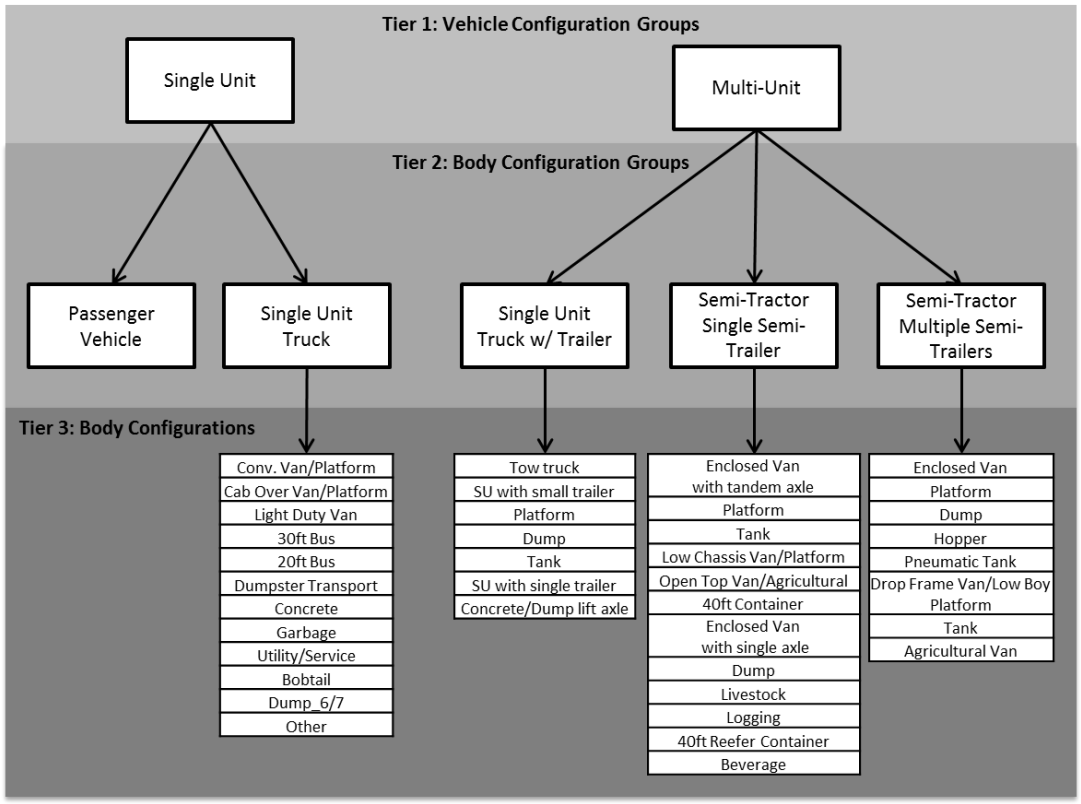


Figure 4-4 Model Structure for Inductive Signature Model

4.2.4.2 Results

The results are presented for the first and second tiers of the model followed by the results of the third tier body classification models for each of the body configuration categories. Results are presented in terms of the Correct Classification Rate (CCR), Absolute Percent Error (APE), and Mean Absolute Percent Error (MAPE). These are defined as follows:

$$CCR_b = \frac{N_b^{Correct}}{N_b^{Observed}} \times 100\ (\%)$$

$$APE_b = \frac{|N_b^{Observed} - N_b^{Predicted}|}{N_b^{Observed}} \times 100\ (\%)$$

$$\text{MAPE} = \frac{\sum_b (APE_b \times N_b^{Observed})}{N} \times 100\ (\%)$$

where

b = body class

$N_b{}^{Correct}$ = number of vehicles correctly classified in body class b

$N_b^{Observed}$ = number of vehicle observed in body class b (e.g. the true vehicle count)

$N_b^{Predicted}$ = number of vehicles predicted in body class b

N = total number of vehicles in test dataset

Tier 1 and 2 Results

The first tier of the model which separated single units from multi-unit vehicles had an overall CCR of 98.4%. The CCR for single units was 98.3% and 98.4% for multi-unit configured vehicles. Vehicles classified as single unit vehicles by the first tier were subsequently classified as either passenger cars (PC) with CCR of 95.3% or single unit trucks without trailers (SU) with CCR of 89.6%, for an overall CCR of 93.4%.

Vehicle classified as multi-unit vehicles by the first tier were subsequently classified into single units with trailers, single semi-trailers, or multiple semi-trailers, with CCR of 82.7%, 98.6%, and 96.1% respectively. The overall CCR for multi-unit vehicles was 96.3%.

The combined result of the first and second tiers is presented in the confusion matrix in Table 4-1. Common misclassifications occur for single units with trailers. A possible reason for the low performance of the single unit with trailers class is that signatures for these vehicles are similar to those of smaller three or four axle semi-tractor trailers. Secondly, smaller trailers towed by single unit trucks do not result in signatures that are distinctly different from those of larger single unit trucks without trailers.

Overall, the first and second tiers of the model produce reasonably accurate classifications. MAPE in volume of the combined first and second tiers is 1.3%. Volume APE ranges between 0.5% and 7.6%. Single units with trailers have the highest APE, but overall the model produces accurate volume estimates of each of the five vehicle-body configuration groups.

Table 4-1 Inductive Signature Only Model Tier 1 and 2 Cross Classification Table and Volume Accuracy

Vehicle-Body Configuration Groups	Single Unit		Multi-Unit			Total	CCR (%)
	PC	SU	SU w/ Trailer	Single Semi	Multi Semi		
Passenger Car	5,389	263	3			5,655	95.3
Single Unit (SU)	298	2,417	121	20		2,856	84.6
SU w/Trailer		104	925	186	23	1,238	74.7
Single Semi		47	89	7,415	9	7,560	98.1
Multi Semi			6	16	548	570	96.1
Total	**5,687**	**2,831**	**1144**	**7,637**	**580**	**17,879**	**93.4**
Volume APE (%)	**0.5**	**0.9**	**7.6**	**1.0**	**1.8**		

Tier 3 Results

Single Unit Truck without Trailers Classification

This model consists of 13 body classes. Table 4-2 summarizes the CCR results for each of the base classifier models and the NBC. NBC achieved superior performance over the best performing base classifier with CCR of 72.4%. For all but four body classes, the NBC approach performed better than the best base classifier. Nine of the 13 body classes have CCR above 70%, of which five achieve CCR of at least 80%, and finally, two models have CCR above 90.0%. Still several body classes have low classification accuracy. Single units without trailers have a wide variety of body types and due to their shorter length, have less distinguishing inductive signature features. This leads to low classification performance in this class. The most common errors resulted from vehicles being misclassified into the majority classes of conventional or cab-over van/platforms as well as utility/service trucks. This is most likely due to the variability in the signature features of these classes.

The MAPE in volume is 15.4% for the MCS modeling approach. Class specific volume APE ranges from 0.0 to 287.0% for the MCS approach. The largest discrepancies in volume arise from several low volume classes (< 30 samples) such as concrete mixers, dump trucks with triple tandem axles, and street sweepers. In these cases, the APE measurement somewhat exaggerated the small absolute differences in volume.

Table 4-2 Single Unit Truck without Trailer MCS Summary

Body Class	*Vol.*	Base Classifier Models (CCR %)					MCS (CCR %)
		MLFF	SVM	CPNN	DT	NB	NBC
Conv. Van/Platform	*333*	72.4	69.7	56.5	67.3	67.6	**74.5**
Utility/Service	*312*	74.4	68.3	59.3	64.4	58.7	**68.9**
Cab Over Van/Platform	*209*	30.6	56.0	23.4	35.9	45.5	**68.4**
30ft Bus	*114*	86.8	89.5	89.5	76.3	85.1	**89.5**
Bobtail	*107*	88.8	86.9	74.8	69.2	82.2	**89.7**
Garbage	*93*	77.4	91.4	90.3	74.2	66.7	**88.2**
Multi Stop Van/RV	*77*	39.0	35.1	29.9	44.2	15.6	**51.9**
20ft Bus	*74*	44.6	70.3	63.5	67.6	17.6	**78.4**
Dump/Tank	*66*	27.3	39.4	28.8	39.4	36.4	**36.4**
Dumpster Transport	*59*	49.2	61.0	50.8	35.6	52.5	**54.2**
Concrete	*21*	81.0	100.0	100.0	85.7	61.9	**95.2**
Dump w/ Triple Rear	*8*	62.5	75.0	62.5	12.5	62.5	**75.0**
Street Sweeper	*3*	66.7	100.0	100.0	33.3	0.0	**66.7**
OVERALL	***1,476***	***63.5***	***68.6***	***56.6***	***59.7***	***57.5***	***72.4***
Minority Classes	***622***	***64.3***	***72.5***	***66.6***	***61.3***	***55.5***	***74.3***

Single Unit Truck with Trailers Classification

The body classification model for Single Unit Trucks with Trailers covers nine truck-trailer combinations. A tractable set of unique truck-trailer combinations were observed in the data so separate models to predict trucks and trailers was not necessary. The body classes are presented as truck –trailer combinations, e.g. Dump- Dump is a dump truck pulling a dump trailer. **Table** 4-3 summarizes the CCR for each of the base classifiers and the two model combining strategies. Using the NBC method, the overall CCR is 94.2%. The best base classifier varied for each body class. For example, NB was best for single units with small trailers (SU small trailer) while SVM was best for RV w/ towed vehicles.

Single units with small trailers and RVs with towed vehicles are commonly cross classified due to the similarities in the signature shapes of these two classes. Concrete trucks with lift axles extended achieve superior performance in terms of CCR and precision. The MAPE in volume was 8.2% across all classes with four classes achieving APE below 10%. Tow trucks, tanks with tank trailers (Tank-Tank), and dump trucks with lift axles extended which had poor classification accuracy, likewise have low volume accuracy.

Table 4-3 Inductive Signature Only Model Tier 3 Single Unit Truck with Trailer MCS Results

Body Class	Vol.	Base Classifier Models (CCR %)					MCS (CCR %)
		MLFF	SVM	CPNN	DT	NB	NBC
SU small trailer	515	95.0	93.0	85.6	94.8	97.9	**96.3**
Dump-Dump	87	92.0	100.0	100.0	93.1	100.0	**100.0**
RV w/ Towed Vehicle	49	67.3	93.9	81.6	59.2	91.8	**85.7**
Concrete w/Lift Axle	34	100.0	100.0	88.2	91.2	79.4	**100.0**
Tank-Tank	30	83.3	100.0	93.3	66.7	70.0	**76.7**
Platform-Platform	20	35.0	25.0	50.0	30.0	70.0	**65.0**
Tow Truck w/ vehicle	8	87.5	50.0	75.0	87.5	62.5	**75.0**
Dump w/ Lift Axle	3	100.0	66.7	66.7	100.0	66.7	**66.7**
OVERALL	***746***	***90.9***	***92.1***	***86.3***	***89.1***	***94.5***	***94.2***

Tractor Classification for Combination Trucks

The tractor body classification model is applied to vehicles that have been classified as semi-tractor trailer combination trucks by the second tier of the model. There are three tractor body classes included in the model: sleeper cabs, cabs without sleepers, and other. The model uses a subset of the inductive signature features that pertain to the first part of the signature representing the tractor. The model is implemented as a multi-layer feed forward neural network as described in the previous chapter. The testing data included 4,100 sleeper cabs, 1,519 cabs without sleepers, and 26 other. The other category includes unconventional tractor types pulling semi-trailers such as wreckers, tanks, platforms, etc. Results are shown in Table 4-4 by CCR for each class. Overall the model achieves a CCR of 88.4%. The lowest performing group, 'Other' cabs, is typically misclassified as 'cabs without sleepers'.

Table 4-4 Inductive Signature Only Model Tractor Results

Body Class	Vol.	MLFF (CCR %)
Cabs without Sleepers	1,774	**78.1**
Sleeper Cabs	4,044	**93.2**
Other	20	**40.0**
OVERALL	***5,838***	***88.4***

Single Semi Trailers Classification for Combination Trucks

There are 19 trailer body classes included in the model for single semi-trailers. This category consists of semi-trailers with several different axle configurations including those with single axle trailers resembling FHWA class 8, those with tandem axle trailers resembling FHWA class 9, as well as triple or more axle trailers. Enclosed vans were separated into two groups which approximate FHWA class 9 (five axle) and 8 (three or four axle) semi-trailers. Each of the remaining 17 body classes contains both single and tandem rear axle configurations within its class. The classification performance of the MCS method is summarized in Table 4-5 along with the CCRs for each of the base classifiers.

The overall CCR of the MCS is 74.3% and 73.8% when considering only minority classes. In this model, minority classes refer to all body classes except enclosed vans and enclosed van reefers in FHWA 8 and 9. Seven of the 19 body classes have CCR between 70% and 80%, five have CCR between 80% and 90%, and two have CCR above 90%. Low performance in terms of CCR is observed for enclosed van reefers in

FHWA class 8, 53ft containers, and agricultural vans. Several unique body classes including logging, livestock, and beverage trailers have CCR above 80% and precision above 75%.

Enclosed Van body types including FHWA class 8 and 9 reefer and non-reefer configurations and 53ft containers were commonly cross classified. Essentially, the subtle difference in the signatures caused by the configuration of the chassis for these set of body types is not able to be picked up by the models. Also, vehicles are commonly misclassified as platform trailers. This is likely due to the diverse set of body types and axle configurations within the platform body class. For example, platform trailers can have split tandem axle configurations causing differing inductive signature shapes.

MAPE in volume was 11.3% for all classes and 17.5% for minority classes. Class specific APE ranged from 3.7% (Enclosed vans Reefer FHWA 9) to 149% (Drop Frame Vans). APE errors above 20% were observed for enclosed reefer vans in FHWA class 8, 40ft reefer container 20ft containers, low chassis trailers including low boy platform and drop frame vans, dump trucks, agricultural vans, and beverage trailers. The low CCR and APE in volume can be improved by further collapsing commonly cross classified classes. For example, if collapsed to nine classes, the total CCR rises to 90.5% and MAPE in volume becomes 9.6% with class specific CCR between 59 and 95% and APE between 7 and 100%.

Table 4-5 Semi Tractor Trailers MCS Summary

Body Class	Vol.	Base Classifier Models (CCR %)					MCS (CCR %)
		MLFF	SVM	CPNN	DT	NB	NBC
Enclosed Van (FHWA 9)	*2343*	31.2	33.1	87.5	63.8	46.4	**74.6**
Enc. Van Reefer (FHWA 9)	*1624*	54.4	71.9	22.9	71.2	44.3	**74.3**
Enclosed Van (FHWA8)	*89*	66.3	69.7	86.5	77.5	71.9	**83.1**
Enc. Van Reefer (FHWA 8)	*13*	76.9	46.2	0.0	30.8	38.5	**46.2**
53ft Container	*124*	71.0	74.2	21.8	64.5	37.9	**57.3**
40ft Container	*136*	68.4	51.5	19.1	81.6	19.9	**75.0**
40ft Container Reefer	*17*	88.2	88.2	82.4	70.6	82.4	**94.1**
20ft Container	*14*	71.4	85.7	0.0	57.1	71.4	**85.7**
Platform	*796*	57.4	60.2	70.4	64.2	83.8	**77.5**
Tank	*283*	58.3	65.0	78.1	71.7	64.3	**70.7**
Open Top Van	*185*	42.2	54.1	51.4	67.6	28.1	**60.5**
Auto	*86*	61.6	61.6	66.3	52.3	88.4	**77.9**
Low Boy Platform	*184*	58.7	75.5	97.3	63.6	71.2	**82.1**
Drop Frame Van	*51*	54.9	49.0	3.9	62.7	70.6	**60.8**
Dump	*54*	59.3	66.7	44.4	48.1	59.3	**70.4**
Logging	*15*	86.7	80.0	80.0	73.3	73.3	**80.0**
Livestock	*56*	78.6	83.9	67.9	78.6	28.6	**80.4**
Agriculture	*29*	44.8	79.3	13.8	51.7	13.8	**58.6**
Beverage	*14*	92.9	92.9	0.0	50.0	7.1	**92.9**
Overall	***6113***	**47.4**	**54.2**	**61.5**	**66.6**	**52.0**	**74.3**
Minority Classes	***2044***	**59.6**	**63.7**	**62.3**	**66.2**	**64.1**	**73.8**

Multiple Trailers Classification for Combination Trucks

There are seven trailer body classes included in the model semi-tractor trailers combination trucks with multiple trailers. Each trailer in the multi trailer configuration was of the same body class for all of the observed data. For example, two enclosed van trailers, or two tank trailers, but not one enclosed van and one tank trailer. The body classes listed in Table 4-6 represents the body class of both trailers in the multi-trailer configuration. The class labeled 'platforms/tanks' contains tank-tank and platform-platform trailers, not to be confused with tank-platform trailers. The model was trained with separate distinctions for platforms and tanks, but later the classes were merged due to the inability of the models to effectively distinguish between these two classes. Likewise, the low chassis van/platform class represents a merged class, rather than a mixed trailer configuration.

Table 4-6 summarizes CCR for each of the base classifiers and the two model combining strategies. As with the previous models, the MCS approach achieves overall higher CCR compared to the best performing base classifier and is more consistent across body classes than any of the of the base classifiers. Using MCS, all seven body classes achieve above 70% CCR with five of these having CCR above 90%. In examining the cross classification matrix for the MCS approach it can be seen that common misclassifications occur between enclosed vans and platforms/tanks. The MAPE for the multi-trailer model was 7.0%. Enclosed vans, platform/tanks, and dump multi-trailers all have APEs lower than 7%. The APEs of the remaining classes ranges from 22 to 50%.

Table 4-6 Multiple Semi Tractor Trailer Combination Trucks MCS Summary

Body Class	Vol.	Base Classifier Models (CCR %)					MCS (CCR %)
		MLFF	SVM	CPNN	DT	NB	NBC
Enclosed Van	*253*	92.1	93.3	77.6	87.4	77.6	**92.9**
Platform/Tank	*121*	91.7	86.8	25.6	71.1	95.0	**90.1**
Dump	*126*	84.1	88.9	90.5	84.9	80.2	**90.5**
Pneumatic Tank	*36*	72.2	75.0	80.6	72.2	66.7	**75.0**
Hopper	*46*	84.8	37.0	54.3	95.7	93.5	**91.3**
Agricultural Van	*2*	100.0	0.0	100.0	50.0	0.0	**100.0**
Low Chassis Van/Pltfr.	*20*	95.0	80.0	100.0	75.0	100.0	**85.0**
OVERALL	***604***	***88.8***	***84.9***	***69.1***	***82.8***	***82.7***	***90.4***
Minority Classes	***351***	***86.3***	***78.9***	***63.0***	***79.5***	***86.3***	***88.6***

4.2.4.3 Discussion

The inductive signature only model used inductive signature data as the sole input. The model was structured into three tiers. The first tier distinguished between single units and multi units. The second tier separated vehicles in each of the first tier bins into axle configuration classes. For single units, these are passenger cars and single unit trucks without trailers. For multi units, these are single units with trailers, semi-tractor trailer combination trucks with single semi-trailers, and semi-tractor trailer combination trucks with multiple semi-trailers. Combined together, Tiers 1 and 2 achieved CCR of 93.4% and volume error of only 1.3%. The four body classification models on Tier 3 are summarized by their training and testing dataset sample sizes, number of body classes, CCR, and APE in Table 4-7. The overall classification vehicle scheme of the inductive signature only model is presented in Appendix F.

The major contribution of the Inductive Signature only model is that it greatly increases the level of information available from an inductive loop sensor. Current methods found in the vehicle classification

literature can at best distinguish vehicles into a few length based bins using conventional inductive loop detectors, and methods using inductive signatures are capable of classifying vehicles into a handful of body groups. The method presented in this section distinguished amount 47 body classes representing a vast improvement over existing methods.

Notable body class distinctions for single unit trucks include 30 and 20ft buses, concrete mixers, garbage trucks, bobtails, and street sweepers. Knowledge of these body classes can assist in separating freight from service trucks for freight modelling as well as for emissions estimation. Other notable body class distinctions for single semi unit trucks include distinction between five and three or four axle semi-trucks, container transport trailers including 40ft, 40ft reefer, and 20ft containers, and specialty trailers like dump, beverage, livestock, and logging trailers. Commodity specific body types can be especially useful for commodity based freight modeling.

Of the 47 body classes, 34 have CCR above 70% and 19 have APE in volume lower than 10%. Single unit and single semi-trailer models have the largest variety of body types and therefore also possess higher volume error and lower classification accuracy. Low performance in these classes is due to the varied axle configurations in the class. For example, single units can be two to four axle trucks and semi-trailers can be three to five axle trucks. Inductive signatures are not apt at distinguishing axle configuration thus this diversity maybe partly to blame for lower performance in these classes.

Table 4-7 Inductive Signature Only Model Tier 3 Results Summary

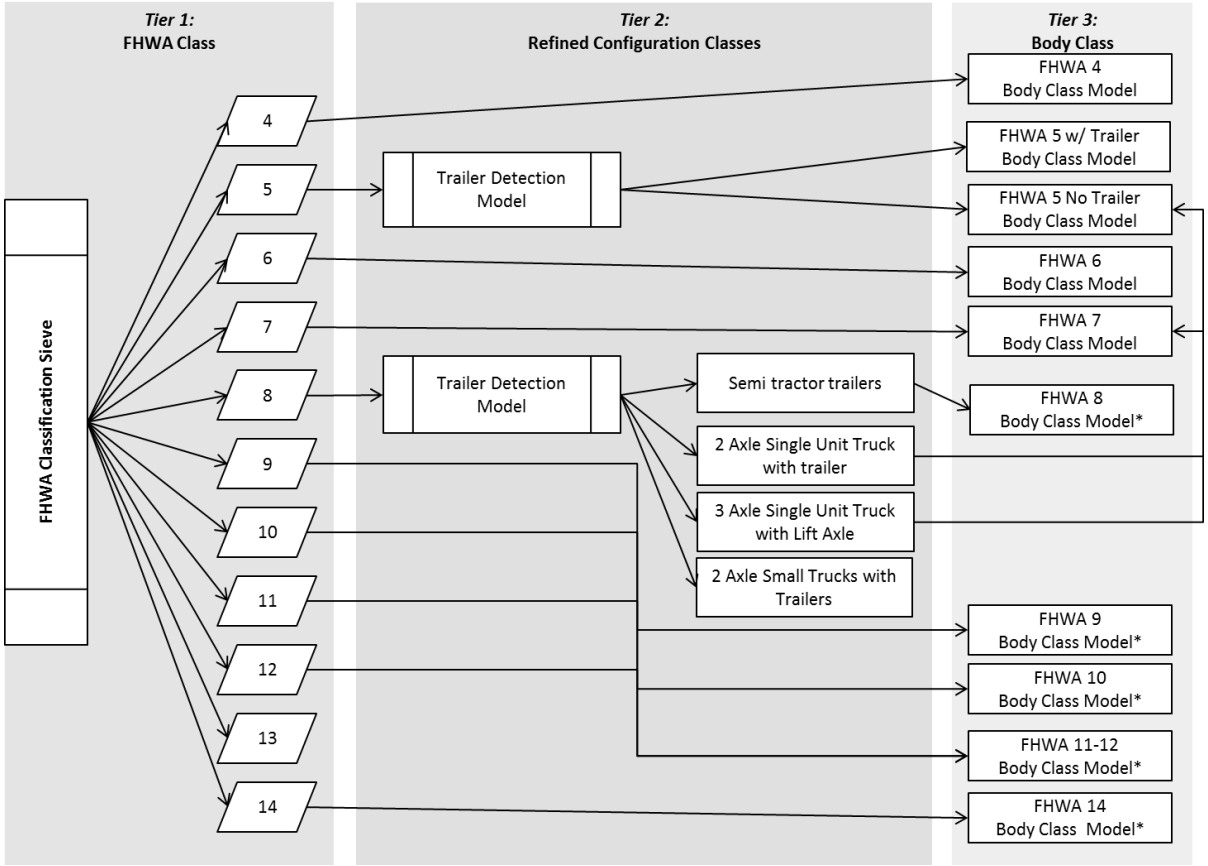
Model	Training Samples	Testing Samples	Body Classes	CCR (%)	Volume MAPE (%)
Single Units	1,553	1,476	13	**72.3**	**15.4**
Single Units with Trailers	714	746	8	**94.2**	**8.2**
Single Semi Trailers	3,720	6,113	19	**74.2**	**11.3**
Multiple Semi Trailers	375	605	7	**90.4**	**7.0**
Overall	**6,362**	**8,940**	**47**	**34 models CCR > 70%**	**27 models APE < 10%**

4.2.5 Inductive Signature and WIM Body Classification Model

For each traversing vehicle, an inductive signature was collected concurrently with conventional WIM measurements. Vehicle axle configuration is first determined from axle spacing and total vehicle weight measurements according to the FHWA axle configuration classification sieve (Quinley , 2010). Body class is subsequently determined using inductive signature data.

4.2.5.1 Model Structure

The overall model structure is shown in Figure 4-5. Following the data paring and parsing procedure, the former of which will be described in Chapter 5, the model is divided into three tiers. A separate body classification model was estimated for each axle configuration group corresponding to the FHWA scheme, yielding a total of eight truck body classification models. Each model uses a slightly varied set of inputs and several models incorporate a multi-tiered approach to separate unique axle configurations before predicting body class. For multi-unit trucks with tractors and trailers, separate models were developed for classification of the tractor and trailer units.

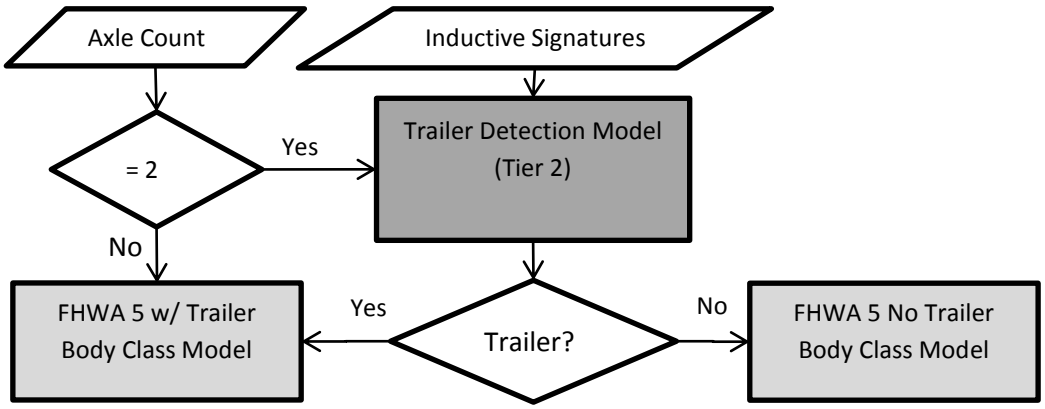


* Indicates separate body class model for tractor and trailer units

Figure 4-5 Model Structure for WIM and Inductive Signature Model

With the exception of the modifications described below, each of the eight body classification models was implemented as a hybrid MCS combined via NBC. Cross validation was used to train each base classifier in the ensemble. For NBC, trained models were first applied to a reference data randomly sampled from the test data to produce a confusion matrix for each base classifier. The five class predictions were subsequently combined according to the NBC technique.

For single unit trucks (FHWA 5), a two tier model was implemented to initially separate trucks with trailers from those without. This tiered approach was necessary due to occasional missed trailer axle detections observed in the WIM data. The first tier, implemented as a MLFF, was based on inductive signature features alone. The second tier subsequently classified non-trailer trucks by body type following the MCS method. The trailer detection model used for two axle trucks is shown in Figure 4-6.



*Due to low sample size, no model was estimated for this class

Figure 4-6 FHWA Class 5 Model Framework

Similarly, vehicles classified into FHWA class 8 by the axle-based classification sieve were first separated into four further refined axle-based categories: three or four axle semi-tractor trailers, two axle single unit trucks with trailers, two axle small trucks with trailers, and three axle trucks with lift axles. Approximately 71% of the trucks assigned to FHWA class 8 were three or four axle semi-tractor trailers and body class was subsequently determined for trucks in this refined axle group. Examples of the four types of axle configurations contained in FHWA class 8 are shown in Figure 4-7. Therefore, vehicles classified into FHWA class 8 by the axle-based classification sieve were first separated into four further refined axle-based categories: three or four axle semi-tractor trailers (Figure 4-7a), two axle single unit trucks with trailers (Figure 4-7b), two axle small trucks with trailers (Figure 4-7c), and three axle trucks with lift axles (Figure 4-7d). A MLFF neural network was used to classify vehicles designated as FHWA class 8 into the four refined axle categories. Only those predicted to be three or four axle semi-tractor trailers, approximately 71% of the FHWA class 8 records, were subsequently classified by body type. Vehicles predicted to be two axle single unit trucks with trailers were assigned to the 'FHWA class 5 with trailer' model. Vehicles predicted to be three axle trucks with lift axles were assigned to the 'FHWA class 7' model. Lastly, vehicles predicted to be two axle small trucks (FHWA class 3) with trailers were not further classified as these vehicles types are outside the scope of the body classification models.

(a)Three or Four Axle Semi-Tractor Trailers

(b) Two Axle Single Unit Truck with Trailers

(c)Two Axle Small Trucks with Trailers

(d) Three Axle Trucks with Lift Axles

Figure 4-7 Examples of Axle Groups with FHWA Class 8

Five axle tractor trailers (FHWA 9) were first separated into two axle configuration groups: three axle semi-tractors pulling a two axle semi-trailer (FWHA 9a) and three axle straight trucks pulling a two axle full trailer (FHWA 9b, also referred to in the California scheme as Class 14). Axle spacing and weight thresholds from the California adaptation of the FHWA classification scheme (Lu et al, 2002) were used to divide trucks as FHWA 9a or 9b.

Multi-trailer units (FHWA 11 and 12) differ in the axle count and configuration but analysis of the groundtruth data showed that they share the same set of trailer body types, so the two classes were merged into one model. Multi-trailer configurations consisted of two trailers of the same body type, so the body class model outputs a single prediction representing all trailers pulled by a truck. Seven or more axle multi-unit trucks (FHWA 13) tend to be specialized equipment movers or other unique body types, so no model was developed for this class.

4.2.5.2 Results

In this section, results are presented in terms of the correct classification rate (CCR), volume APE, volume MAPE, and summary of MCS base classifier performance. In the summary table of the MCS base classifier performance, models with less than ideal performance, deemed to be less than 60% CCR, have been highlighted in red to emphasize the cases where an individual base classifier would not perform well.

The terms used to assess the models include class specific CCR, Absolute Percent Error (APE), and Mean Absolute Percent Error (MAPE) as defined previously.

Single Unit Truck Classification

Single Unit Trucks include FHWA classes 4, 5, 6, and 7. Results for these classes are summarized in Table 4-8 through Table 4-11.

FHWA Class 4

Vehicles classified into FHWA class 4 are meant to be of two or three axle buses, the classification sieve captured several non-bus body classes including vans and platforms, as well as recreational vehicles (RVs). The model was trained with 63 samples and tested on 62 samples. A summary of the MCS base classifier performance is provided in Table 4-8. The overall CCR for the MCS approach with NBC voting

was 95.2%. None of the five base classifier models exceeded the CCR of the NBC approach across all body classes. CCR was near or above 90% across all four body classes. The model underestimates the volume of 30ft buses with single rear axles by 12% but the overall MAPE stands at 9.7%.

Table 4-8 FHWA 4 Body Classification Model Results

Body Class	*Vol.*	Base Classifier Models (CCR %)					MCS (CCR %)
		MLFF	SVM	CPNN	DT	NB	NBC
Van or Platform	*26*	92.3	92.3	76.9	88.5	96.2	**92.3**
30ft Bus Tandem	*11*	100.0	100.0	90.9	81.8	45.5	**90.9**
30ft Bus Single	*25*	96.0	0.0	84.0	100.0	100.0	**100.0**
RV	*0*	-	-	-	-	-	**-**
Overall CCR (%)	**62**	**95.2**	**56.5**	**82.3**	**91.9**	**88.7**	**95.2**

FHWA Class 5

Vehicles detected with only two axles by the WIM system or those identified by the axle detection model shown above as having no trailer are included in the FHWA class 5 without trailer body classification model. A total of 10 groups were created from the 21 body classes observed in the data. Groups were formed based on shared body characteristics, general usage characteristics, and the models ability to distinguish between particular classes. Of note are platforms and vans which could not be distinguished effectively by the models and were therefore lumped into a body class group. Also, platform trucks were separated into four types that represent body class characteristics matching other vehicles in the dataset. For example, 'cab over platforms' were grouped with 'cab over vans' since both have similar body characteristics and could not be distinguished by the models. Vehicles grouped into the 'other' category were not able to be distinguished as any of the 18 listed body classes.

The CCR results of the MCS base classifiers and combination strategies are summarized in Table 4-9. The overall CCR of the NBC approach was 75.3%. Of all the base classifier models, the SVM classifier performs best overall with 71.3% CCR, however has CCR below 60% for several classes while the NBC approach achieves CCR above 60% for all but one class ('other' trucks). The majority of body classes have CCR above 70%. Low performing classes include light vans/RVs, 12 passenger vans, and 'other' trucks. Misclassifications tend to occur by assigning vehicles into the van/platform and utility/platform/pickup categories. The MCS with NBC approach yields a MAPE in volume of 6.8% with volume APE for each class between 0.0% and 25.0%.

Table 4-9 FHWA Class 5 without Trailer MCS Summary

Body Class	*Vol.*	Base Classifier Models (CCR %)					MCS (CCR %)
		MLFF	SVM	CPNN	DT	NB	
Cab Over Van/Platform	*215*	41.4	73.5	47.9	54.0	63.3	71.6
Conv. Van/Platform	*180*	86.1	89.4	58.9	80.0	86.7	88.3
Utility/Platform/Pickup	*174*	66.1	62.1	49.4	62.6	64.4	73.6
Light Van/RV	*104*	67.3	51.9	47.1	61.5	54.8	67.3
20ft Bus	*71*	83.1	87.3	69.0	71.8	18.3	83.1
Tow Truck/Platform	*61*	67.2	47.5	26.2	49.2	60.7	70.5
12 Pass Van	*41*	73.2	78.0	58.5	63.4	19.5	65.9
30ft Bus	*32*	87.5	96.9	93.8	81.3	84.4	96.9
Other	*22*	63.6	18.2	59.1	45.5	22.7	22.7
Bobtail	*12*	91.7	91.7	100.0	91.7	83.3	91.7
Overall CCR (%)	***912***	**67.1**	**71.3**	**53.5**	**64.4**	**61.5**	**75.3**

FHWA Class 6

There were 215 samples in the training dataset representing 15 distinct body class which were collapsed into the eight defined in the model. Dump, dumpster transport, and garbage trucks are the most prevalent in this class. The CCR of each base classifier by body class and the MCS combining strategies are shown in Table 4-10. The MCS with NBC combination has an overall CCR of 80.5% with class CCR all above 60%. The model performs very well in identifying bobtail tractors with 91.5% CCR and 95.6% precision, and trucks with trailer assigned to FHWA class 6 with 92.9% CCR and 100.0% precision. Common cross classification occurred into the 'platform/van/tank/other' class. However this is to be expected given the wide diversity of body types and feature distributions within this class. The MCS approach has a MAPE in volume of 9.4% with class specific APE ranging from 0.0% to 27.6%. Garbage trucks and concrete mixers possess the largest APEs in volume, while bobtails, buses, and dumpster transport truck have APE in volume below 10%.

Table 4-10 FHWA 6 MCS Summary

Body Class	*Vol.*	Base Classifier Models (CCR %)					MCS (CCR %)
		MLFF	SVM	CPNN	DT	NB	NBC
Dumpster	95	72.6	74.7	42.1	76.8	67.4	**78.9**
Bobtail	94	92.6	97.9	85.1	89.4	90.4	**91.5**
Platform/Van/Tank/Other	89	59.6	39.3	53.9	65.2	92.1	**73.0**
Dump	88	75.0	76.1	40.9	47.7	60.2	**78.4**
Garbage	29	72.4	89.7	86.2	51.7	37.9	**79.3**
Concrete	16	81.3	93.8	93.8	68.8	18.8	**68.8**
FHWA 6 w/ trailer	14	57.1	92.9	85.7	85.7	35.7	**92.9**
Bus	1	0.0	100.0	100.0	100.0	0.0	**100.0**
Overall CCR (%)	***426***	***74.4***	***75.1***	***60.3***	***69.5***	***71.1***	***80.5***

FHWA Class 7

FHWA Class 7 three axle single unit trucks are divided into four body classes. From the observed data it was found that two axle single unit trucks with extended lift axles were often categorized as three axle trucks by the FHWA classification sieve. Therefore the body classes include two axle dump trucks and concrete mixers with lift axles in addition to three axle dump trucks ('dump triple') and garbage trucks ('garbage triple'). The 72 samples in the training data come from the Irvine and Fresno sites as none

were observed at the Willows or Redding sites. This was to be expected since three axle single unit trucks tend to be urban service trucks, e.g. garbage or dump trucks, and would therefore not be found in more rural areas like Redding and Willows.

The test dataset was comprised of a limited set of 19 samples. Table 4-11 summarizes the performance of the base classifiers and MCS combination methods on the test data. The MCS with NBC performs with 100.0% CCR across all body classes, significantly improving upon all of the individual base classifiers and the MV combination method. Since the MCS with NBC had 100.0% CCR and there were no cross classifications. The MAPE in volume was 0.0%.

Table 4-11 FHWA 7 MCS Summary

Body Class	*Vol.*	Base Classifier Models (CCR %)					MCS (CCR %)
		MLFF	SVM	CPNN	DT	NB	NBC
Garbage Triple Tandem Axle	9	77.8	0.0	100.0	100.0	66.7	100.0
Concrete Tandem w/ Lift	7	100.0	100.0	42.9	100.0	42.9	100.0
Dump Triple Tandem Axle	1	100.0	0.0	100.0	100.0	100.0	100.0
Dump Tandem w/ Lift	2	100.0	100.0	50.0	100.0	100.0	100.0
Overall CCR (%)	19	89.5	47.4	73.7	100.0	63.2	100.0

Multi-Unit Truck Tractor Classification

The current model for multi-unit truck tractor body classification covers only FHWA class 9, however it can be expanded later to include FHWA classes 8 and 10 through 13. The tractor body classification scheme is a two class scheme: (1) cabs with sleepers and (2) cabs without sleepers. The tractor body class was modeled using a Naïve Bayes model.

Figure 4-8 shows the distribution of the axle spacing data for the two tractor body classes: cabs without sleepers and sleeper cabs. For the tractor model, the spacing between the 1st and 2nd axles, vehicle length, and an interaction term constructed by multiplying the vehicle length and spacing between the 1st and 2nd axles were used as inputs to the classification model. The spacing between the 2nd and 3rd axle did not vary significantly by tractor body type and was not included in the model. Cabs without sleepers have shorter spacing between the 1st and 2nd axles than cabs with sleepers. The median distances between the 1st and 2nd axles were 13.3ft and 17.3ft for cabs without sleepers and cabs with sleepers, respectively. The median vehicle length for cabs without sleepers was 66.4ft and 74.0ft for sleeper cabs.

The interaction term was included to further delineate between the two tractor body types across sites. It was observed that the tractor axle spacing (e.g. spacing between axles 1 and 2) for cabs without sleepers at Redding was significantly larger than that at the other three sites. However, the overall vehicle lengths of cabs without sleepers observed at the Redding site were shorter than at the other three sites. Interacting the terms allowed for better delineation between sleeper and non-sleeper cabs that was consistent across all observed sites.

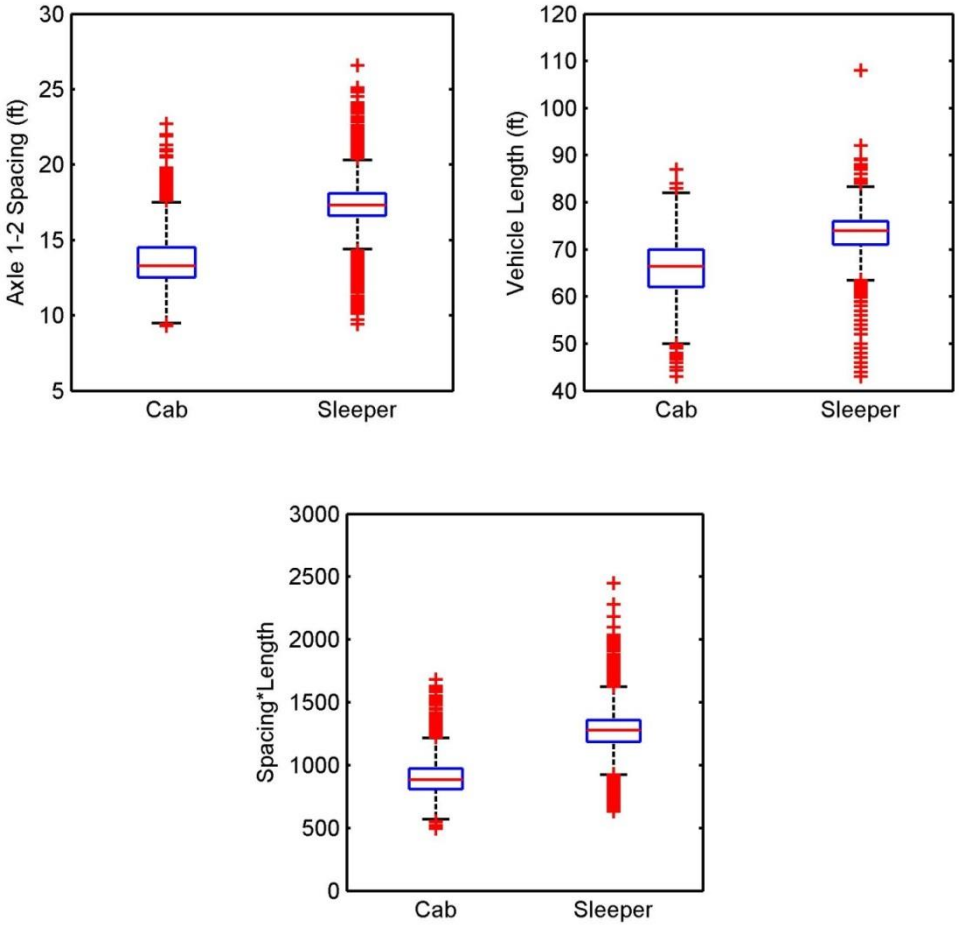


Figure 4-8 Box plots of spacing between the 1st and 2nd axles, vehicle length, and vehicle length interacted with axle spacing by tractor body class

As summarized in Table 4-12, the total CCR was 91.9% for the 4,999 test samples. Cabs without sleepers had a CCR of 83.4% while sleeper cabs had a CCR of 94.6%. The overall CCR were 90.1%, 96.2%, and 94.0% for Fresno, Willows, and Redding, respectively.

Both the training and test data from Redding had lower than average accuracy for cabs without sleepers. As previously mentioned, the spacing of the 1st and 2nd axles of cabs without sleepers observed at Redding was larger than the average across the other four sites. This configuration was found at each of the four sites but dominated the Redding site data. Tractors without sleeper cabs that have longer tractor chassis might be considered as an additional unique body type, however, there is no clear visual features by which the body could be identified to provide groundtruth model data.

Table 4-12 Results of the WIM-Only Tractor Classification Model

	Cabs without sleepers	Sleeper Cabs	Count	CCR
Fresno	86.4%	91.8%	1,692	**90.1%**
Willows	79.8%	97.4%	1,474	**96.2%**
Redding	60.3%	97.2%	670	**94.0%**
Overall	**83.4%**	**94.6%**	**4,999**	**91.9%**

The Irvine data was held out during model training so that it could be used independently for spatial transferability testing. Table 4-13 summarizes the spatial transferability results. The overall CCR was 88.0%. The CCR for cabs without sleeper was 83.6% and for sleeper cabs the CCR was 91.2%. The results support the conclusion that the model is reasonably transferable to locations that were not included in model training.

Table 4-13 Spatial Transferability Results of the WIM-Only Tractor Classification Model

	Cabs without sleepers	Sleeper Cabs	Count	CCR
Irvine	83.6%	91.2%	1,163	**88.0%**

Multi-Unit Truck Trailer Classification

Multi-Unit Trucks include FHWA classes 8, 9, 10, 11, 12, and 13. Results for these classes are summarized in Table 4-14 through Table 4-17.

FHWA Class 8

The FHWA Class 8 Trailer Axle Detector Model distinguishes four refined axle configurations found within the FHWA class 8 data using an MLFF neural network. The overall CCR is 93.6% with class specific CCR between 78.7 and 100.0%. The MAPE in volume is 8.9%. Those vehicles identified as two axle single unit trucks with lift axles and two axle single unit trucks with trailers, are then assigned to the body class models for FHWA class 7 and 5 with trailers, respectively. Vehicles identified as two axle small trucks with trailers are terminally identified. Vehicle predicted to be three or four axle semi-tractor trailers are fed into the FHWA class 8 body class model.

There were seven unique trailer body classes observed for FHWA class 8 three or four axle semi-tractor trailer combination trucks. Of the seven trailer body classes, five trailer body class groups were formed. Enclosed van trailer are the dominant class accounting for 76% of the training samples and 80% of the test data. Unique minority trailer body classes include beverage, livestock, and agricultural trailers.

The results of the base classifiers and MCS combining methods are shown in Table 4-14. The MCS with NBC achieves 90.9% and 71.1% CCR for the overall and minority classes, respectively. The majority of cross classification occur between platform and van trailers. If van and platform trailers were to be combined into a single class, the CCR of the combined class would be 97.6%. Low chassis van/platform and beverage trailers had CCR of 77.8 and 100.0%, respectively. Unfortunately agricultural van trailers were not found in the test data set. However, the validation dataset performance shows that the base classifier models were able to predict agricultural van trailers with CCR of 50%. The MCS based model

generally produces accurate volume estimates with MAPE of 4.3%. Platforms are the largest source of volume error with APE of 20.0%.

Table 4-14 FHWA Class 8 MCS Summary

Trailer Body Class	*Vol.*	Base Classifier Models (CCR %)					MCS (CCR %)
		MLFF	SVM	CPNN	DT	NB	NBC
Van	*149*	90.6	99.3	94.0	83.9	63.1	**96.0**
Platform	*20*	45.0	15.0	40.0	60.0	90.0	**55.0**
Low Chassis Van/Platform	*9*	88.9	55.6	77.8	77.8	33.3	**77.8**
Beverage	*9*	100.0	100.0	100.0	33.3	88.9	**100.0**
Agricultural Van	*0*	-	-	-	-	-	**-**
Overall CCR (%)	***187***	**86.1**	**88.2**	**87.7**	**78.6**	**65.8**	**90.9**
Minority Class CCR (%)	***38***	**68.4**	**44.7**	**63.2**	**57.9**	**76.3**	**71.1**

FHWA Class 9a

FHWA class 9 semi-tractor trailer combination trucks have the widest diversity of body types compared with all other FHWA classes. The training data consisted of 3,205 samples from the Irvine, Willows, Redding, and Fresno data collection sites. The 20 unique body classes were collapsed into 16 body class groups, two of which contain multiple body types: (1) the platform group contains basic platforms, bulk waste transport, and 20ft containers on platforms; (2)the tank group contains liquid, dry bulk, and pneumatic tanks. Grouping of the trailer body types was based on physical and use characteristics as well as cross classifications evident in the training results. Unique body classes include 20ft, 40ft, and 53ft intermodal containers as well as refrigerated vans and containers ('reefer'). Commodity specific classes include tank, automotive transport, livestock, agricultural, and logging trailers.

Table 4-15 summarizes the base classifiers and MCS model combining strategies for the FHWA class 9 trailer body class models. The MCS with NBC achieves 75.5% CCR overall and 77.7% for minority classes. The SVM base classifier has the best performance for minority classes (79.2%) but sacrifices classification accuracy for the majority class as a result. Four of the 16 trailer body class groups have CCR above 90%, five between 80 and 90%, and four between 70 and 80% using the MCS with NBC approach. Lower than acceptable classification performance resulted for 53ft containers. Cross classifications are common between enclosed vans, reefer vans, and 53ft containers due to their similar length, overhang, and chassis characteristics. Classification accuracy of enclosed vans could be increased to 92.1% CCR by merging these three classes. The model performs exceptionally well for many of the unique minority classes. For example, 40ft containers, 20ft containers, automobile transport, and livestock trailers have CCR above 90% and precision above 70%. The MCS with NBC applied to the test data across all sites results in MAPE of 12.0% and 11.1% for all classes and minority classes, respectively. High APE was observed for 40ft reefer containers and drop frame vans, both of which were overestimated. Commodity specific body classes such as auto transport, logging, livestock, and agricultural van trailers had APE in volume less than 20%.

Table 4-15 FHWA Class 9 MCS Results Summary

Trailer Body Class	*Vol.*	Base Classifier Models (CCR %)					MCS (CCR %)
		MLFF	SVM	CPNN	DT	NB	NBC
Enclosed Van	2,210	35.8	53.8	74.8	69.3	61.2	**71.9**
Reefer Vans	1,564	64.5	82.4	24.3	72.4	61.0	**74.9**
53ft Box Container	116	69.0	71.6	7.8	59.5	13.8	**55.2**
40ft Container	131	74.0	77.1	66.4	58.8	16.8	**81.7**
40ft Reefer Container	16	87.5	93.8	18.8	81.3	87.5	**93.8**
20ft Container	13	92.3	84.6	100.0	92.3	100.0	**100.0**
Platforms	734	51.0	72.5	89.5	77.4	88.6	**81.3**
Tank	259	65.3	74.1	81.5	76.4	71.8	**78.4**
Open Top Van	152	61.2	79.6	16.4	67.8	33.6	**80.3**
Auto	67	89.6	92.5	94.0	74.6	82.1	**94.0**
Low Boy Platform	166	80.1	81.9	90.4	81.9	88.6	**88.6**
Dump	51	62.7	80.4	41.2	45.1	68.6	**66.7**
Drop Frame Van	43	65.1	76.7	32.6	58.1	58.1	**74.4**
Logging	14	78.6	78.6	85.7	85.7	85.7	**85.7**
Livestock	45	95.6	100.0	64.4	84.4	62.2	**95.6**
Agricultural Van	22	90.9	72.7	4.5	77.3	40.9	**63.6**
Overall CCR (%)	**5,603**	**52.9**	**69.2**	**59.4**	**71.5**	**63.7**	**75.5**
Minority Class CCR (%)	**3,393**	**64.1**	**79.2**	**49.4**	**72.9**	**65.3**	**77.7**

FHWA Class 9b (California Class 14)

The body class model for FHWA class 9b (California Class 14) single unit trucks with single trailers consists of five body classes. The body classes represent the truck and trailer body of the vehicle. For example, 'Dump-Dump' refers to a single unit dump truck pulling a single dump trailer. The training set consisted of 145 samples, dominated by 'dump-dump' trucks. Overall the CCR for the test data of 244 samples was 96.7% for the MCS with NBC method. Dump-Dump, Tank-Tank, and RV-Small trailer classes have CCR above 90%. Platform-Platform trucks are the only underperforming class due to cross classifications as tank-tank and dump-dump trucks. MAPE in volume of the MCS with NBC was 1.7%. All body classes had APE in volume of less than 2%. Only the Platform-Platform class exceeded 10% APE.

Table 4-16 FHWA 14 MCS Summary

Trailer Body Class	*Vol.*	Base Classifier Models (CCR %)					MCS (CCR %)
		MLFF	SVM	CPNN	DT	NB	NBC
Dump-Dump	171	94.7	98.2	83.6	84.2	98.2	**99.4**
Tank-Tank	31	87.1	90.3	71.0	90.3	90.3	**90.3**
RV-Small Trailer	25	92.0	100.0	100.0	88.0	100.0	**100.0**
Platform-Platform	17	64.7	52.9	11.8	52.9	52.9	**76.5**
Livestock-Livestock	0	-	-	-	-	-	**-**
Overall CCR (%)	**244**	**91.4**	**94.3**	**78.7**	**83.2**	**94.3**	**96.7**

FHWA Class 10

The FHWA Class 10 model was limited by the number of samples available for model training and testing. Due to low sample size, the MCS approach could not be applied and a single SVM model was used for prediction instead. The six unique trailer body classes were collapsed into four groups: 20ft intermodal containers, platforms, low chassis vans and platforms, and enclosed vans. The overall CCR

was 100.0% for the test data consisting of a limited set of 10 samples. The model preforms well given the limited training data.

FHWA Class 11 and 12

From the nine unique trailer body classes representing dual semi-trailers, seven groups were created. The majority of samples are enclosed vans, platform, and bottom dump trailers. Minority body types included pneumatic tanks, hoppers, and agricultural vans. The MCS results are shown in Table 4-17 for the test set of 302 samples. The overall CCR for the MCS with NBC was 92.6%. The MLFF neural network base classifier has an overall higher CCR than the NBC method, however, it achieve slightly lower accuracy for bottom dump and tank trailers. The MAPE in volume of the MCS with NBC was 8.0%. The largest APE in volume results from platforms being misclassified as tanks, all other class specific APE in volume were less than 15%.

Table 4-17 FHWA 11 and 12 MCS Summary

Trailer Body Class	*Vol.*	Base Classifier Models (CCR %)					MCS (CCR %)
		MLFF	SVM	CPNN	DT	NB	NBC
Platform	108	92.6	88.9	42.6	77.8	87.0	**91.7**
Van	88	95.5	96.6	87.5	88.6	90.9	**95.5**
Bottom Dump	67	98.5	100.0	98.5	97.0	100.0	**100.0**
Hopper	33	81.8	63.6	45.5	75.8	81.8	**75.8**
Pneumatic Tank	25	88.0	80.0	64.0	72.0	68.0	**88.0**
Tank	3	66.7	66.7	66.7	66.7	33.3	**100.0**
Agricultural Van	2	100.0	0.0	100.0	50.0	0.0	**100.0**
Overall CCR (%)	***326***	***92.9***	***89.3***	***68.7***	***83.7***	***87.7***	***92.6***

4.2.5.3 Discussion

Both the results and depth of detailed body classes depicted in these models go well above the capabilities of previous classification models which were limited to at most five body classes using current in-pavement technology (Liu et al., 2011). A major advantage of integrating the WIM site with inductive signature capabilities is that the body classification model is able to predict among body types found within an axle configuration class whereas previous inductive signature based models had the added challenge of predicting among all vehicle body and axle configurations. In all, eight separate body classifications models were developed from an extensive data set of 18,967 truck records distinguishing an unprecedented total of 26 single unit truck, 25 semi-trailer body configurations, and 12 multi-unit trucks as shown in the summary in Table 4-18. Of the 63 body classes, 51 had CCR greater than 70% and 33 had APE in volume less than 10%. The overall classification vehicle scheme of the combined WIM and inductive signature model is presented in Appendix G.

Table 4-18 WIM-Signature Model Summary

Model	No. Training Samples	No. Testing Samples	No. Body Classes	CCR (%)	Volume APE (%)
FHWA 4	63	62	4	95.2	9.7
FHWA 5 without Trailers	1,172	912	10	75.3	6.8
FHWA 6	215	342	8	80.5	9.2
FHWA 7	72	19	4	100.0	0.0
FHWA 8 Semi Trailers	224	187	5	90.9	4.2
FHWA 9	3,198	5,603	16	75.4	12.2
FHWA 10	17	13	4	92.3	7.7
FHWA 11 and 12	508	326	7	92.7	8.0
FHWA 14	145	244	5	96.7	1.7
Overall	**5,614**	**7,708**	**63**	**52 with CCR > 70%**	**37 with APE < 10%**

The model for FHWA class 4 effectively distinguishes buses from vans and platforms which have been misclassified by WIM controllers into this class. Hence, this model actually improves classification accuracy according to the FHWA scheme since FHWA class 4 is meant to only contain buses. For two axle single unit trucks (FHWA 5), body configurations are extremely heterogeneous, and axle measurements and weights overlap significantly, however the model still effectively separated buses, passenger vans, utility and pickup trucks, bobtails (tractor drive units), and enclosed vans. Models for FHWA classes 6 and 7 differentiated several industry specific categories including concrete, dump, garbage, dumpster transport, and vans. These are important distinctions if one desires to compare possible freight and non-freight related vehicle volumes, since passenger related vehicles like buses or service oriented trucks such as pickups are not freight carriers.

Five axle combination trucks (FHWA 9) possessed the widest diversity of trailer body types. Enclosed vans formed the majority, while several unique and industry specific classes like logging and intermodal containers comprised the minority. Remarkably, the model can even distinguish refrigerated from non-refrigerated vans as well as refrigerated from non-refrigerated 40ft intermodal containers. This level of distinction paves new possibilities for advanced tracking of perishable commodities, especially those that are port-related. Also, because length is included in the model, 20ft intermodal containers can be classified with high accuracy. Other commodity specific body types such as automobile transport, logging, and livestock trailers had CCR above 85%. Unfortunately, the model is not able to accurately distinguish 53ft intermodal containers from enclosed vans due to overlap in signature and axle configuration features thus it is necessary to collapse these classes into a single group for this study.

Across all axle configuration groups, the MCS approach which combined five base classifiers – MLP, SVM, CPNN, DT, and Naïve Bayes – using NBC had higher overall CCR than any single base classifier model. While several base classifiers undoubtedly demonstrated adequate performance at predicting certain classes, there was always a tradeoff with low performance of the same classifier on other classes. To illustrate this important facet of MCS, the class specific CCRs less than 60% were highlighted in red in each of the MCS summary tables. Each classifier has a set of classes where it performs quite poorly, hence the need for the MCS method which consequently produced fewer class specific CCRs under 60%.

In the FHWA class 9 model, for example, the SVM base classifier produced more accurate predictions than any base classifier for livestock trailers (100.0%) but had the lowest performance among all base classifier for enclosed van trailers (53.8%). If the SVM model had been applied alone, the accuracy in predicting enclosed vans would have been sacrificed in favor of livestock trailers. But by combining all models, the CCR for enclosed vans was elevated to 71.9%, well above the ability of the SVM model, while still maintaining the high accuracy for livestock trailers (95.6%). Accordingly, for all eight models (FWHA class 4 through 11/12) the overall CCRs were above 75% using the MCS approach. Alternatively, had only the MLFF approach been implemented, for example, only six of the eight models would achieve CCRs above 75%.

Since NBC calculates an array of probabilities for the set of possible body classes based on the joint distribution arising from each base classifier's cross classification matrix, the strengths of each base classifier are captured and weaknesses are controlled. For example, an unknown vehicle could be classified as a reefer by the CPNN model, which had a CCR of only 24% for this class because most reefers were misclassified as enclosed vans. In NBC, the CPNN model will contribute little 'evidence' toward the estimated probability of the vehicle being a reefer (i.e. the number of records that were truly reefers and predicted as such is low) but contribute significant 'evidence' toward the estimated probability of the vehicle being an enclosed van (i.e. the number of records that were truly reefers but classified as enclosed vans is high). In the simplest possible interpretation, this approach allows the best model to be used for each class. Without this approach, there is no way to predetermine which base classifier should be applied to an unknown vehicle.

4.2.6 Spatial Transferability Analysis

Spatial transferability analysis was performed for the WIM-Signature FHWA class 9 semi-tractor trailer body class model since this model possess the widest array of body classes and spatial differences in body class distributions were observed to be particularly significant for this class of trucks. To assess the spatial transferability of the model, the data from the Irvine data collection site was held out from model training, and then the trained MCS with NBC method was applied to the data collected from the Irvine site.

Table 4-19 presents the CCR for the Irvine data. The CCR of the Irvine data consisting of 589 samples was 64.3% overall and 55.6% for the minority classes. The MAPE in volume was 30.4% and 41.3% for the overall and minority classes, respectively. The highest error arises from 53ft containers. In all, lower performance resulted from leaving the Irvine data out from the model training, however, this was to be expected because each of the four sites possessed widely different body class distributions. Thus, leaving any site out of the training would mean excluding unique samples from the data. For the most accurate model, data from the Irvine site and potentially other sites around the state should be included to fully encompass the diversity of vehicle types.

Table 4-19 Spatial Transferability Analysis Cross Classification Table for Irvine Data

Trailer Body Class	Total Number of Test Samples	CCR (%)
Enclosed Van	231	77.9
Reefer	104	51.0
53ft Container	44	25.0
40ft Container	10	40.0
40ft Reefer	0	-
20ft Container	0	-
Platform	62	72.6
Tank	19	42.1
Open Top Van	56	44.6
Auto Transport	11	90.9
Low Boy Platform	27	88.9
Drop Frame Van	10	90.0
Dump	15	66.7
Logging	0	-
Livestock	0	-
Agricultural Van	0	-
Total	**589**	**64.3**

4.3 Methods for Backcasting Body Classification Data

The motivation for this method was to find a way to get more truck body class information out of historical WIM data for the purposes of validating the California Freight Forecasting Model (CSFFM) to prior validation years. The objective of this model was to estimate site and time specific truck body configuration volumes using existing WIM site data and to tie together body configuration volume and GVW distribution to provide the much needed relationship between GVW distribution and body class.

Rather than predicting individual vehicle classifications as the models in Section 4.2, the model presented in this section produces aggregate trailer body configuration volumes and GVW distribution estimates at the hourly, weekly, or seasonal levels for a specific WIM location. This method allows more information to be extracted from axle-based measurement data without requiring modifications to existing infrastructure such as installing inductive signature capable detector cards, thus better leveraging already heavy investments in WIM systems. This level of aggregation can better meet the needs of the various transportation agencies tasked with freight planning, air quality monitoring, and general truck data collection.

The model was developed for five axle semi-tractor trailers that have been classified according to their axle spacing and weight measurements to be FHWA class 9 trucks. Five axle semi-tractor trailer combination trucks have distinct axle configuration characteristics, prevalence in the traffic stream relative to other truck types, and unique implications for freight and emissions analysis. Thus,

identifying the body classes of five axle semi-tractor trailer combination trucks is of interest. Two models were developed: (1) a tractor body class model and (2) a trailer body class model.

4.3.1 Modeling Framework

Classification algorithms which rely on WIM data alone have been shown to be incapable of making body configuration predictions (FHWA, 1999). Hence, this model estimates overall body configuration volume and GVW distribution using a modified decision tree (MDT) algorithm. Whereas a standard decision tree (DT) assigns singular predictions to each sample collected at a terminal node of the tree, our MDT approach instead applies predetermined probabilities of each body configuration to the collection of samples that accumulate at each terminal node. Then, body configuration volume can be estimated by summing across all terminal nodes.

4.3.1.1 Modified Decision Tree (MDT) Model

A DT is a non-parametric, supervised classification method. DTs are viable for large data sets, intuitively represented, and due to their non-parametric implementation, do not require calibration. This technique has been used in a variety of fields for these reasons (Chang and Wang, 2006). In the transportation domain, DTs have been used for safety and accident analysis (Chang and Wang, 2006; Yan and Radwan, 2006) and for axle-based classification (FHWA, 1999).

A binary Classification and Regression Tree (CART) (Breiman, 1984) was selected as the MDT implemented in this paper. Used in their standard form, a DT declares one target class at each terminal node. However, since we aim to produce volume estimates at an aggregate level rather than individual vehicle classifications, the probabilities of body configurations estimated at each terminal node were used to produce body configuration volume rather than assign a single prediction to all collected samples. Predetermined probabilities are applied to distribute the samples into body configurations at each terminal node. The predetermined probabilities result from training the tree on observed field data. The input variables of the MDT represent each body configuration's physical attributes which were statistically shown to be invariant by site or time, thus the terminal node's body configuration proportions from the observed field data (i.e. training data) can be used to predict the body configuration volumes when the same input variables are used to bin new, unseen data via the MDT. The estimated volumes of each body configuration from each terminal node are then aggregated across all terminal nodes to produce the body configuration volume estimates:

$$V_{body\ configuration\ i} = \sum_n V_i^n = \sum_n Pr_i^n \times v^n$$

where V_i = total volume of body configuration i , V_i^n= volume of body configuration i at terminal node n, Pr_i^n = probability of body configuration i at a terminal node n, v^n= total volume collected at a terminal node n

To evaluate model performance, the absolute percentage error (APE) was used. APE measures the deviation of estimated from actual truck volumes by body configuration. The overall APE representing all sites by body configuration is the volume weighted average of each body configuration's APE at each site:

$$APE_{body\ configuration\ i}^{site\ t} = \frac{|Actual_i^t - Estimated_i^t|}{Actual_i^t} \times 100\ (\%)$$

$$\text{APE}_{\substack{all\ sites \\ body\ configuration\ i}} = \frac{\sum_t (APE_i^t \times Actual_i^t)}{\sum_t Actual_i^t} \times 100(\%)$$

4.3.1.2 Gaussian Mixture Model (GMM) for estimating GVW distribution

A GMM was applied to estimate GVW distributions by body configuration using the body configuration volume estimates obtained from the MDT model. A GMM is a linear composition of Gaussian distributions, $\mathcal{N}(\mu_m, \Sigma_m)$ combined via a mixing proportion parameter, p_m (Hastie, 2009):

$$f(x) = \sum_{m=1}^{M} p_m \cdot \mathcal{N}(x; \mu_m, \Sigma_m)$$

where m = number of mixture components, $\mathcal{N}(\mu_m, \Sigma_m)$ = Gaussian distribution with mean μ and covariance matrix Σ, and $\mathrm{p_m}$ is the mixing proportion

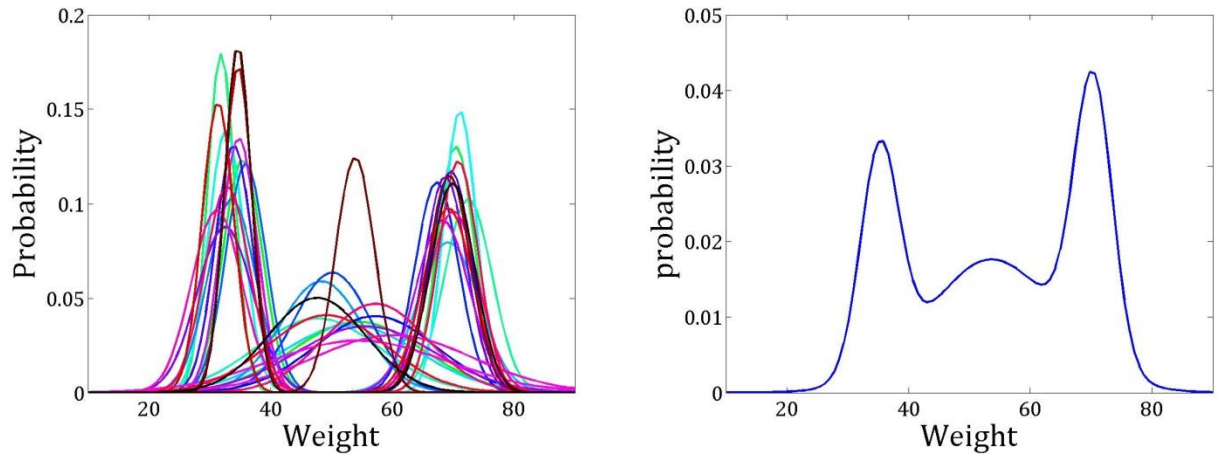
To estimate GVW by body configuration, a GMM consisting of three mixture components was estimated from the GVW data collected at each terminal node, *n*, of the MDT. The three mixture components approximate the unloaded, partially loaded, and fully loaded GVW distributions. Next, the GMMs at all terminal nodes were aggregated to form a body configuration-specific GMM consisting of three independent Gaussian distributions from each of *n* terminal nodes, e.g. *n* by three Gaussian mixture components. The mixture proportions, $p'^{\,n}_{\,m}$, in the body configuration specific GMM are derived from the volume of each body configuration, *i*, at each node, *n*, such that $p'^{\,n}_{\,m}(i) = v^n_m(i) \cdot p^n_m(i)$ where $p^n_m(i)$ is the mth mixing proportion of the nth terminal node for body class *i*. The final body configuration specific GVW distribution is as follows:

$$f(i) = \sum_{n=1}^{N} \sum_{m=1}^{M} p'^{\,n}_{\,m}(i) \cdot \mathcal{N}(z;\ \mu^n_m, \Sigma^n_m) = \sum_{n=1}^{N} \sum_{m=1}^{M} v^n_m(i) \cdot p^n_m(i) \cdot \mathcal{N}(z;\ \mu^n_m, \Sigma^n_m)$$

where i = body configuration, n = terminal node, m = number of mixture components

It should be noted that only nodes which have sufficient volumes to construct the Gaussian distributions, deemed to be greater than 30 trucks, were used to build the mixture distribution. The model is sensitive to the time aggregation interval in this regard since volumes at each terminal node depend on the body configuration distribution and total volume at a site.

In this study, the trained MDT model had 13 terminal nodes each resulting in a three mixture GMM of GVW. Therefore, the Gaussian mixture distribution for each body configuration consists of up to 39 individual Gaussian distribution components. The total number of Gaussian models is variable since nodes with less than 30 records are disregarded. Figure 4-9(a) shows the 39 individual Gaussian distributions estimated at each node for vans. Figure 4-9(b) represents the final G mixture distribution of GVW for vans from the 39 individual distributions combined via the mixing proportions derived from body configuration volumes at the nodes.



(a) 39 individual gaussian distributions for vans (b) Final GVW distribution for vans

Figure 4-9 Deriving GVW distribution for vans

4.3.2 Data

Three WIM sites described in Chapter 3 (i.e. Fresno, Redding, and Willows) were selected for model development with a fourth site (Irvine) selected for spatial and temporal transferability testing. In total, 10,904 five axle semi-trailer truck records were collected across multiple days and manually identified by body configuration. Data were split into training and testing datasets by day with approximately 70 percent of the data chosen for training and the remaining selected for testing as shown in Table 4-20.

Table 4-20 Summary of Data for Backcasting Model Development

Site Location	Site Description		Collection dates	Number of Truck Samples Collected	Number of Class 9 Truck Samples Collected (%)
SB SR-99 at Fresno	Agricultural, Semi-Urban	*Train*	Nov. 08, 2012	6,715	4,101 (61%)
		Test	Nov. 07, 2012		
SB I-5 at Redding	Rural	*Train*	Dec. 11, 2012	2,326	1,694 (73%)
		Test	Dec. 10, 2012		
NB I-5 at Willows	Rural	*Train*	Dec. 10 - 11, 2012	4,493	3,507 (78%)
		Test	Dec. 12, 2012		
SB I-5 at Irvine	Urban	*Test*	March 20th and 25th, 2013	4,677	1,632 (35%)

Vehicle characteristics

Five axle semi-trailer trucks (also known as 3S2 trucks) contain the largest variety of truck body configurations. Five trailer body groups were formed to capture this diversity: vans, platforms, tanks, 40ft box containers, and 'other'. Each group shares similar physical, industry and/or commodity attributes. 40ft box containers were considered as a separate category in this study since estimating volumes of 40ft box containers would be valuable for capturing intermodal transportation movements.

An exploratory analysis was performed to identify the optimal set of axle configuration variables that are required to determine body configuration. In addition to the spacing between the 3rd and 4th axle and overall vehicle length, a derived measure called overhang was used. Spacing between the 3rd and 4th axles is the measured length in feet between the last tractor axle and the first trailer axle. Length

refers to the distance in feet from the nose of the tractor to the tail of the trailer. Overhang represents the front and rear portions of the vehicle outside the axles and is obtained as the arithmetic difference between the overall length and the sum of all axle spacing measurements. Figure 4-10 shows the boxplot depicting the descriptive statistics of these variables by body configuration, illustrating the potential of each variable to distinguish the body types defined in this modeling effort. Van trailers have the longest length and overhang, and are distinctive from other body configurations. Generally, platforms have the second longest length. Tank type trailers have distinctive, shorter overhang. 40ft box containers have relatively longer overhang compared to their short length. A Kolmogorov-Smirnov hypothesis test confirmed that the five body groups are indeed differentiable by length, axle spacing, and overhang. It should be noted that axle loads and GVW were not used due to spatial dissimilarities in weight. Further, to alleviate measurement inconsistencies due to sensor calibration, a normalization technique was applied. The spacing between the 2nd and 3rd axles of the tractor unit, which is fixed on tractors and does not vary geographically or by body configuration, was used to normalize each WIM measurement (e.g. length, spacing, and overhang).

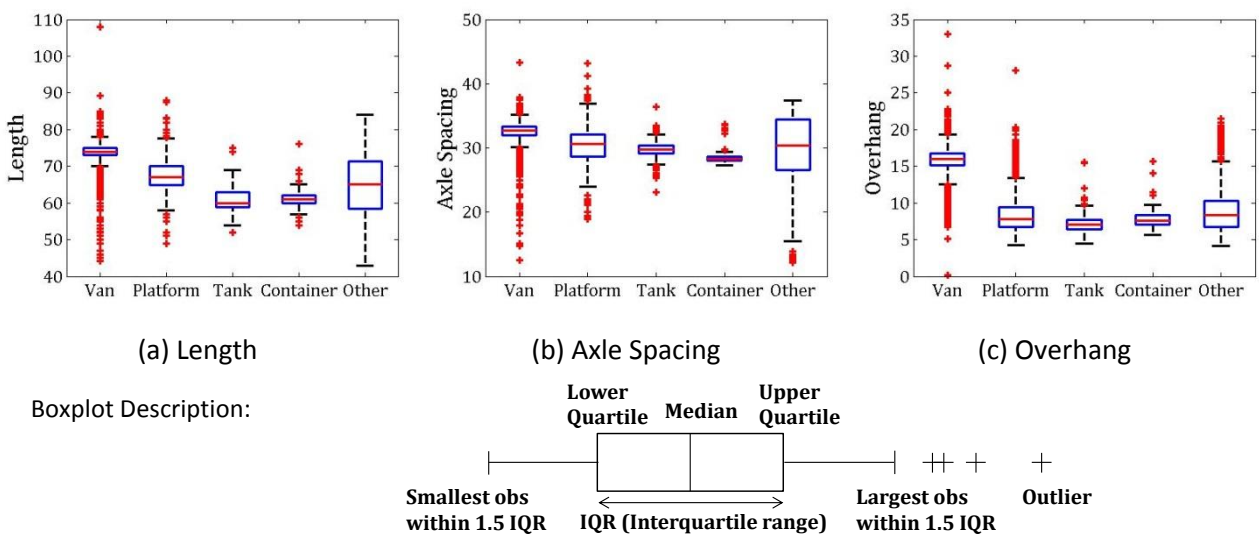


Figure 4-10 Descriptive statistics of WIM variables; (a) Length (b) Axle spacing between 3rd and 4th axle (c) Overhang.

4.3.3 Results

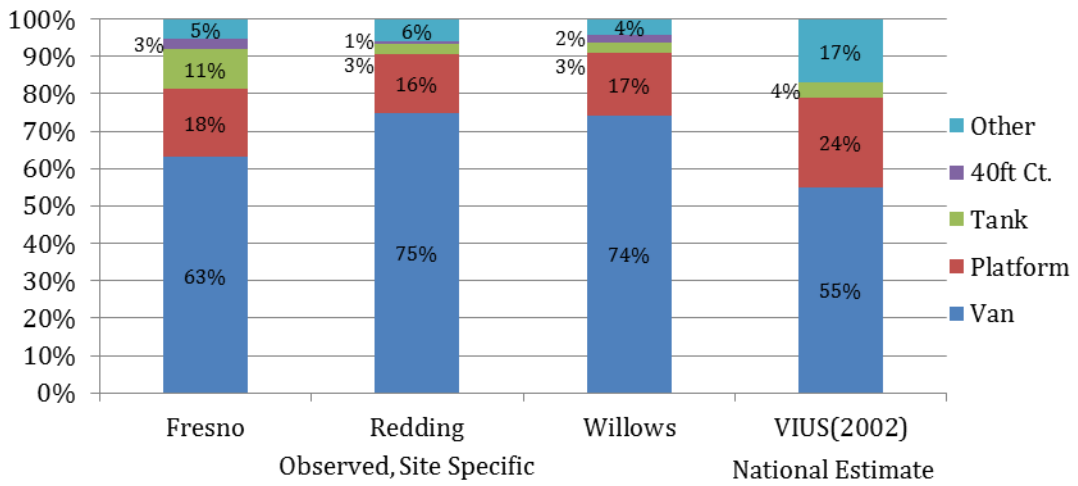
The trained MDT is presented in Appendix D with probabilities of body configurations displayed at each terminal node along with the branching criteria. Table 4-21 summarizes the results on the training and test datasets. APE across all sites and body classes for the training and test data sites were 3.9 percent and 5.3 percent, respectively. For the test data, the APE in volume across all sites for vans and platforms were 1.7 percent and 5.4 percent, respectively. The APE in volume by site ranged between 0.4 percent to 1.9 percent for vans, and 2.8 percent to 13.2 percent for platforms. However, the volumes of tanks, 40ft box containers, and other trailer types had relatively higher APEs, ranging from 18.1 percent to 44.2 percent. These higher APEs were due in part to the heterogeneity in physical characteristics of trucks within those body configuration groups.

Table 4-21. Results of Decision Tree Model

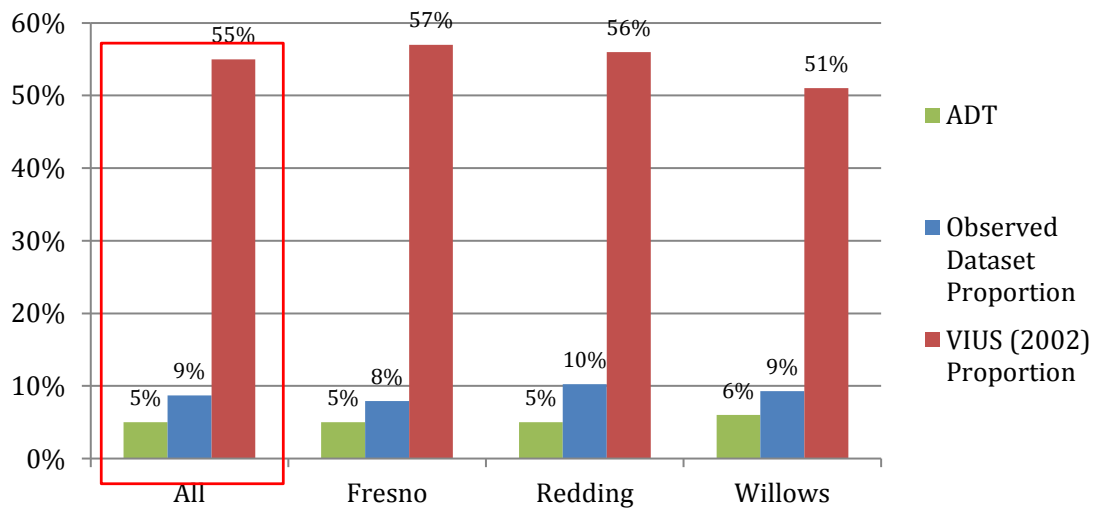
APE(%) by Sites		Test Data Set						Train Data Set					
		Body types					Overall	Body types					Overall
		Van	Pt	Tank	40ft Ct	Other	APE %	Van	Pt	Tank	40ft Ct	Other	APE %
All Sites	Actual Volume	2353	515	208	108	139	**3323**	4180	1207	349	116	307	**5979**
	APE %	1.7%	5.4%	21.2%	31.5%	22.3%	**5.3%**	2.2%	3.7%	20.6%	13.8%	4.9%	**3.9%**
Fresno	Actual Volume	1131	286	160	61	72	**1710**	1509	437	253	63	129	**2391**
	Estimated Volume	1151	294	131	47	87		1540	418	227	70	136	
	APE %	1.8%	2.8%	18.1%	23.0%	20.8%	**5.0%**	2.1%	4.3%	10.3%	11.1%	5.4%	**3.8%**
Redding	Actual Volume	223	53	11	4	29	**320**	1028	217	39	8	82	**1374**
	Estimated Volume	224	60	9	3	25		1041	220	29	9	75	
	APE %	0.4%	13.2%	18.2%	25.0%	13.8%	**4.7%**	1.3%	1.4%	25.6%	12.5%	8.5%	**2.5%**
Willows	Actual Volume	999	176	37	43	38	**1293**	1643	373	57	45	96	**2214**
	Estimated Volume	980	189	50	24	50		1598	389	93	37	97	
	APE %	1.9%	7.4%	35.1%	44.2%	31.6%	**5.9%**	2.7%	4.3%	63.2%	17.8%	1.0%	**4.8%**

For comparison purposes, two baseline approaches were developed. The first is based on body configuration proportions calculated from observed, site specific data. This method applies a static proportion of vehicle types that does not vary by time of day, but only by location. The second is based on national level body configuration proportions collected in VIUS which have been adjusted by Average Annual Truck Miles by body configuration. VIUS does not include intermodal container trucks so this body configuration cannot be estimated under this baseline approach. Figure 4-11(a) shows the body configuration proportions for each site and for VIUS.

To apply the baseline approaches to the test data, the static proportions were used to stratify the test data by body configuration. The same training and testing datasets used for the MDT approach are used so the results are direct comparisons between the MDT and two baseline approaches. Prediction accuracies of the baseline approaches are shown in Figure 4-11(b). The overall APEs are 9 percent and 55 percent for the observed and VIUS methods, respectively. Even though the static proportion derived from the observed data estimate truck body configuration volumes with error ranging from 8 percent to 10 percent, this method is extremely labor intensive since manual data collection at targeted WIM sites is required to obtain the static proportions. In addition, compared to the MDT approach which had an overall error of only 5 percent, VIUS proportion approaches are not viable alternatives.



(a) Proportion of Body Configuration from Observed Data and VIUS (2002)



(b) APE from Baseline Approach

Figure 4-11 Baseline approaches and APE; (a) Proportion of body configuration from observed and VIUS (2002) data (b) APE from baseline proportion approach.

4.3.4 Sensitivity Analysis

Two sources of sensor measurement error may arise in length and weight measurements: systematic errors due to calibration and random errors due to vehicle dynamics over the sensors (Nichols and Bullock, 2004*)*. A sensitivity analysis was performed to ensure that the MDT model is robust to both error sources. To evaluate the sensitivity to systematic errors, a constant 10% increase was applied to each measure (length, spacing, and overhang) in the test dataset. For random errors, random disturbances were applied to each measure through repeated random sampling from a normal distribution, $\mathcal{N}(0, 0.05)$. Two different MDT models, one trained on the non-normalized data (i.e. non-normalized tree, NNT) and a second trained on a dataset normalized by the tractor axle spacing (i.e.

normalized tree, NT) were evaluated under two error scenarios: (1) systematic error only, and (2) systematic and random error.

With normalization, errors greatly decrease for all sites as summarized in Table 4-22. The second column of Table 4-22 shows the error that results when no errors are applied. The normalization approach reduced overall APE from 15 percent to 5 percent under scenario (1) and from 13 percent to 6 percent under scenario (2). Fresno performed the worst with 20 percent APE under scenario (2) and 18 percent APE scenario (2) when no normalization was applied. Redding and Willows had more stable performance with APE below 10 percent for scenarios (1) and (2). In all, the normalization approach effectively corrected for both systematic and random errors.

Table 4-22 Results of Sensitivity Analysis on Backcasting Model

Error Scenario	No Applied Error	Scenario 1: Systematic Error		Scenario 2: Systematic and Random Error	
APE (%)	NT**	NNT*	NT**	NNT*	NT**
All Sites	5%	15%	5%	13%	6%
Fresno	5%	20%	5%	18%	7%
Redding	5%	9%	5%	8%	5%
Willows	6%	9%	6%	8%	6%

*NNT : Non-Normalized Tree, ** NT : Normalized Tree

Spatial and temporal transferability was analyzed to ensure the MDT approach could generalize to sites and time periods not included in model development. The MDT model was applied to the Irvine site, which was not used in model development. Unlike the three sites used to develop the model, the Irvine site is an urban location in close proximity to two major sea ports. General findings conclude that the model is both spatially and temporally transferable. While sites used in model development had 5 percent overall APE, the Irvine site had an overall APE of 8 percent (Figure 4-12). From these findings, we conclude that the model can be readily applied to other WIM sites and has a potential to be used for historical analysis.

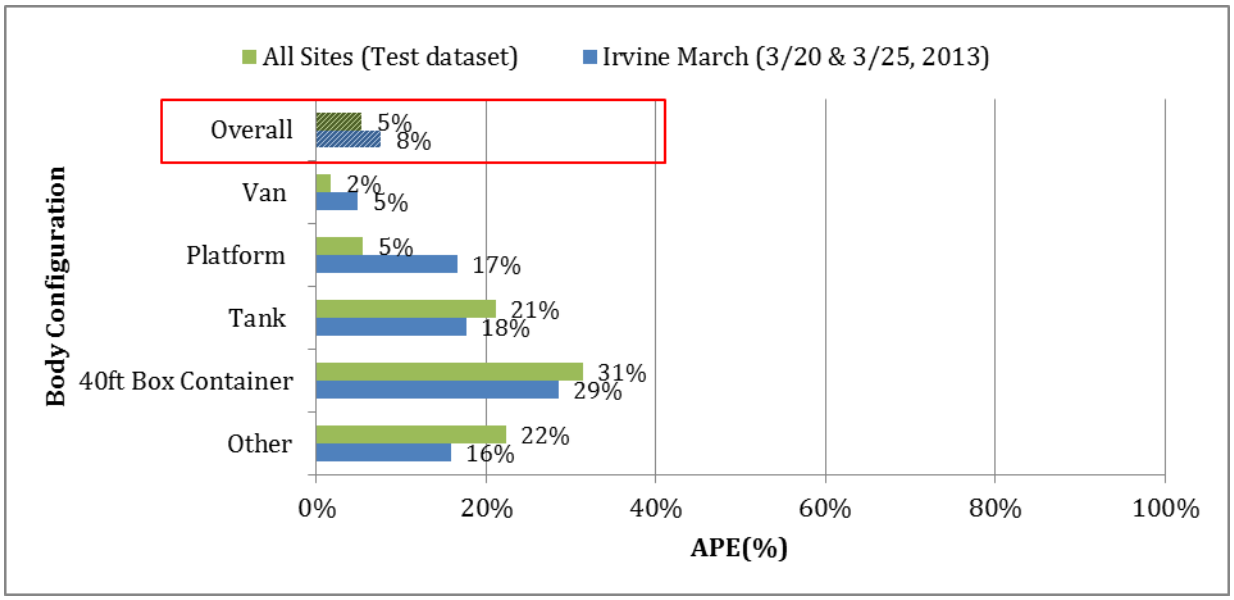


Figure 4-12 Results of transferability analysis

Time of day (TOD) patterns distinguished by body configuration can capture the daily variations of commercial vehicle travel activity. Figure 4-13 highlights the model's ability to accurately track TOD patterns by body configuration at the Irvine site. It is clear that vans and platforms follow different TOD trends, which would not be captured using static proportion methods.

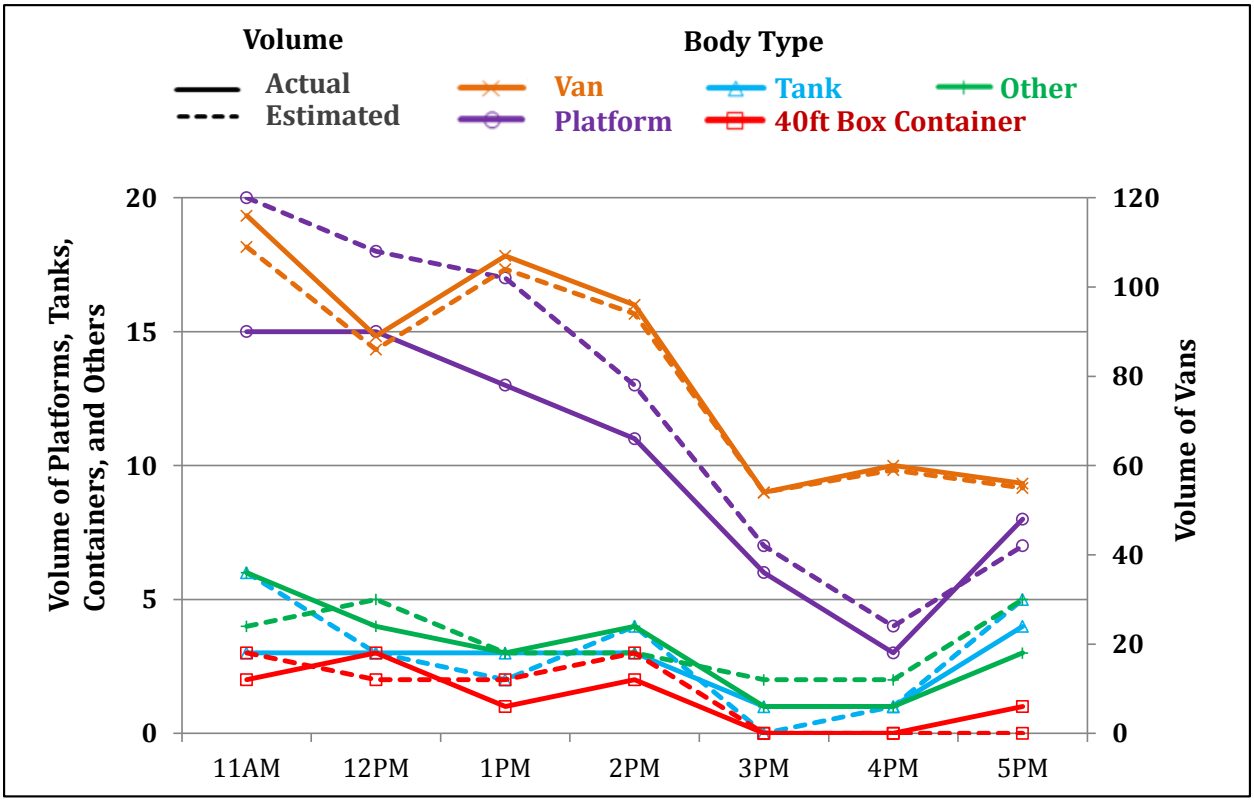


Figure 4-13 Hourly volume estimation using MDT model at Irvine (March 20, 2013)

4.3.5 GVW Distribution Estimation

Figure 4-14 shows the resulting GMMs for (a) vans, (b) platforms, (c) tanks, and (d) 40ft intermodal containers along with a normalized histogram of observed GVWs for each corresponding body configuration. To obtain sufficient volumes needed to estimate each Gaussian mixture distribution, data from all four sites were aggregated. In practice, larger datasets representing weekly or monthly aggregation levels would be more than sufficient. There are 30 bins in the normalized histogram each with a width of 3 kips. The resulting GMMs for vans and platforms depict tri-modal GVW distributions, while tanks follow a bimodal distribution. This reflects the loading characteristics of tanks which due to drive instability arising from partially loaded liquids tend to travel either fully loaded or completely empty whereas vans and platforms travel fully, partially, or un- loaded. In addition, despite the small volume observed for 40ft intermodal containers, the GMM was capable of replicating the high proportion of lighter or empty loads observed. This could indicate return trips or transports of less dense commodities.

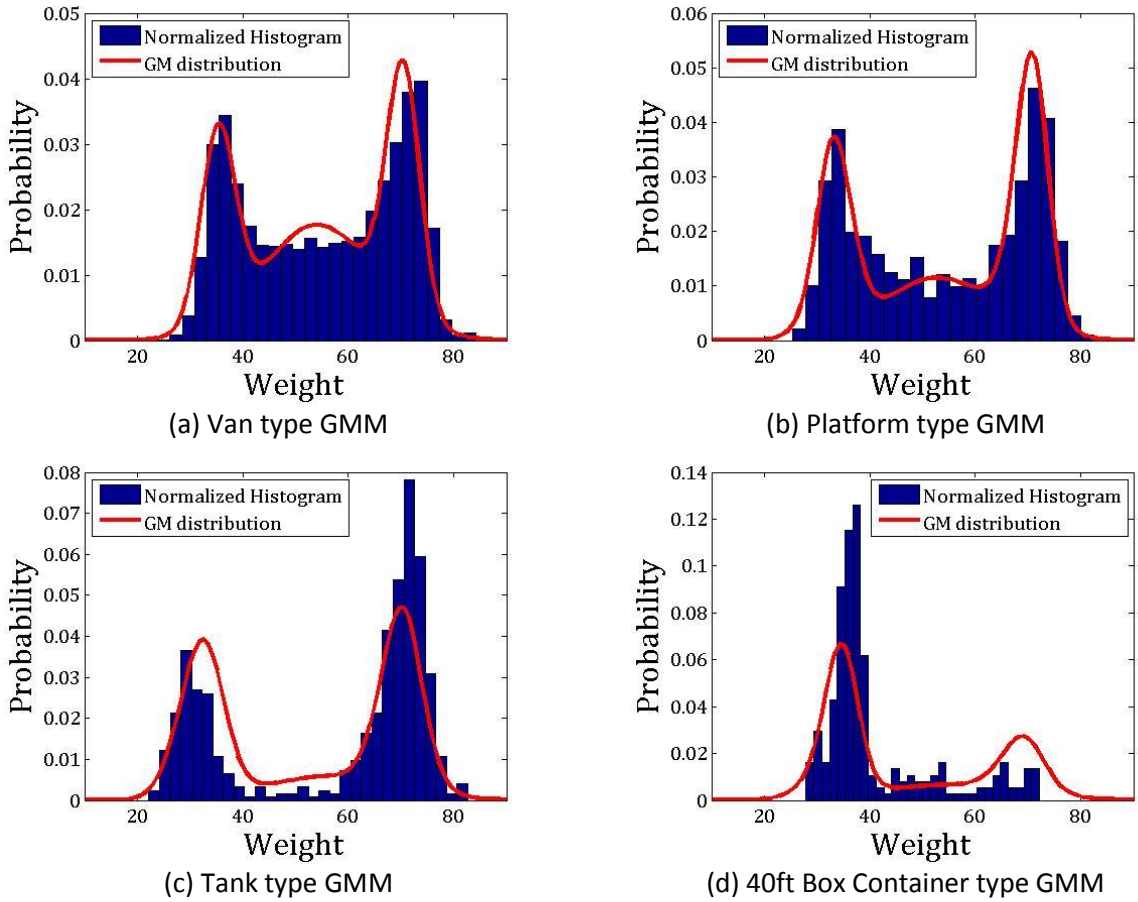


Figure 4-14 Weight histogram and GMM; (a) van (b) platform (c) tank (d) 40ft box container.

A Kolmogorov-Smirnov hypothesis test was performed for all body configurations to confirm the statistical fit of the models. The test concluded that the null hypothesis, i.e. the estimated mixture model and the normalized histogram are from the same population, cannot be rejected at a significance level of 1% ($\alpha = 0.01$) for all body configurations aggregated across all stations. Therefore, the estimated mixture models visually and statistically capture the peaks and overall shape of the observed data.

4.3.6 Conclusion

In this section, a MDT method to estimate five-axle semi-tractor trailers (3S2) volume and GVW by body configuration based on existing WIM site data was shown to have acceptable performance. Across all sites and body configurations, the MDT model predicted body configuration volumes with only 5 percent error. The model employs a simple normalization procedure to ensure robustness to systematic errors due to sensor calibration and random errors due to vehicle dynamics over the sensors. Across all sites, the normalization procedure reduced volume estimation errors caused by systematic and random measurement errors by approximately 7 percent. Furthermore, the model demonstrated spatial and temporal transferability when applied to a test site with a body configuration distribution that diverged from that of the sites used for model development. Compared against a pair of baseline approaches that estimated body configuration volumes using national survey and observed, site specific data the MDT method was shown to produce more accurate volume predictions than either baseline approach.

In addition to body configuration volume estimation, a method to estimate GVW distribution by body configuration was developed. Statistical results show that the GVW estimation model was able to replicate the observed GVW distribution for each of the five body configuration classes.

There are several immediate applications of body class volume and GVW distribution estimates which will be investigated in future studies. First, detailed information on truck volume by body configuration tied to GVW distribution can be used for payload factor estimation used in freight modeling. Second, GVW distributions by body class can be used to assess the volume of empty trips at the link or route level providing more advanced methods for validating empty movements in freight modeling. Lastly, by applying the MDT model to historical WIM data, body class volume changes over time can be assessed at the site level which can be tied to historical accident data to identify accident hot spots stratified by body class and perhaps link body class with safety impacts.

4.4 Optimal Site Selection for System Deployment

4.4.1 Background

The problem of selecting the optimal deployment sites was initially investigated by considering candidate WIM sites across the State of California, and designing optimization functions coupled with sets of constraints to recommend the ideal locations for deployment.

This model was not used in the final site selection effort of this study due to the limited spatial scope of the final deployment strategy (within the California San Joaquin Valley) as well as operational constraints involved with final selection of deployment sites. The final deployment strategy was to perform deployments within the San Joaquin Air Basin, spanning San Joaquin, Stanislaus, Merced, Madera, Fresno, Kings, Tulare, and approximate half of Kern County. However, only thirteen existing data WIM sites are found within this region as shown in Figure 4-15.

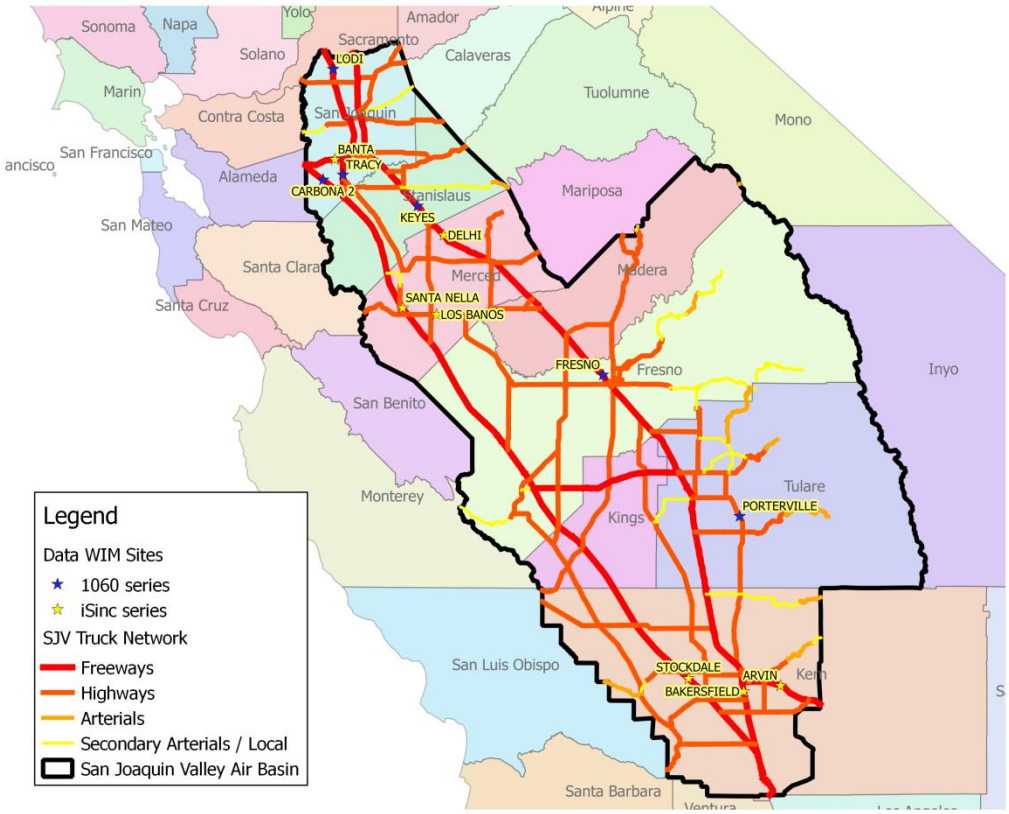


Figure 4-15 Location of Data WIM sites in the San Joaquin Air Basin

Although most of these sites were equipped with 1060 series WIM controllers at the start of this study, many had been upgraded or were being planned for upgrade to the newer iSinc series controllers during the deployment phase of this study. During the deployment phase, it was determined that six WIM sites (Banta, Santa Nella, Los Banos, Delhi, Stockdale, Bakersfield and Arvin) were equipped with the iSinc controller, one (Porterville) was planned for iSinc upgrade. This limited the set of feasible candidate deployment sites, as the iSinc controller is beyond the scope of this project for implementation of the integrated WIM and inductive signature system. In addition, the Fresno WIM site was undergoing construction, and the Lodi WIM site was found to be equipped with 6'x12' inductive loop sensors, which are not compatible with the model developed in this study, which is based on 6'x6' loops. Hence, only three feasible data WIM sites were left as candidates for deployment of the truck body classification system.

Notwithstanding, the models developed for the optimal location of sites for WIM and inductive signature deployment can still be used for future statewide expansion efforts of the system developed in this study. The models can also be easily modified to include existing candidate ILD sites in addition to WIM sites to expand the scope of inductive signature deployment.

Two alternative methods to choose optimal deployment locations based on the analysis of sampled truck GPS data are presented in this section. The first method selects locations based on maximizing the total number of origin-destination flows captured while considering the distribution of truck body

classes by OD pair. The second method attempts to locate the best possible set of stations needed to perform vehicle re-identification along the highway network to provide further insights into truck trip behavior. Both methods would provide guidelines for optimal deployment.

4.4.2 Sensor Location Data

The heavy truck GPS data used in this study was obtained from the American Transportation Research Institute (ATRI). In this dataset, anonymized randomly generated identification numbers (IDs) were used to maintain the confidentiality of truckers and trucking companies. Position and timestamp information was obtained from trucks equipped with automatic vehicle location equipment at predetermined intervals. However, these update intervals varied by trucks. This dataset comprises four weeklong subsets in the middle of each quarter in 2010 to provide seasonal representation of truck travel activities within California, yielding a total of over 8 million truck positions. It was assumed that the truck trajectories derived from this dataset provide a reasonable representation of heavy duty truck activity in California. The truck trajectories were subsequently matched with WIM site locations to determine the number of trips captured by each WIM station, as well as the ordered sequence of WIM sites traversed by each truck trip. Figure 4-16 shows the truck trajectories and locations of existing WIM sites in California.

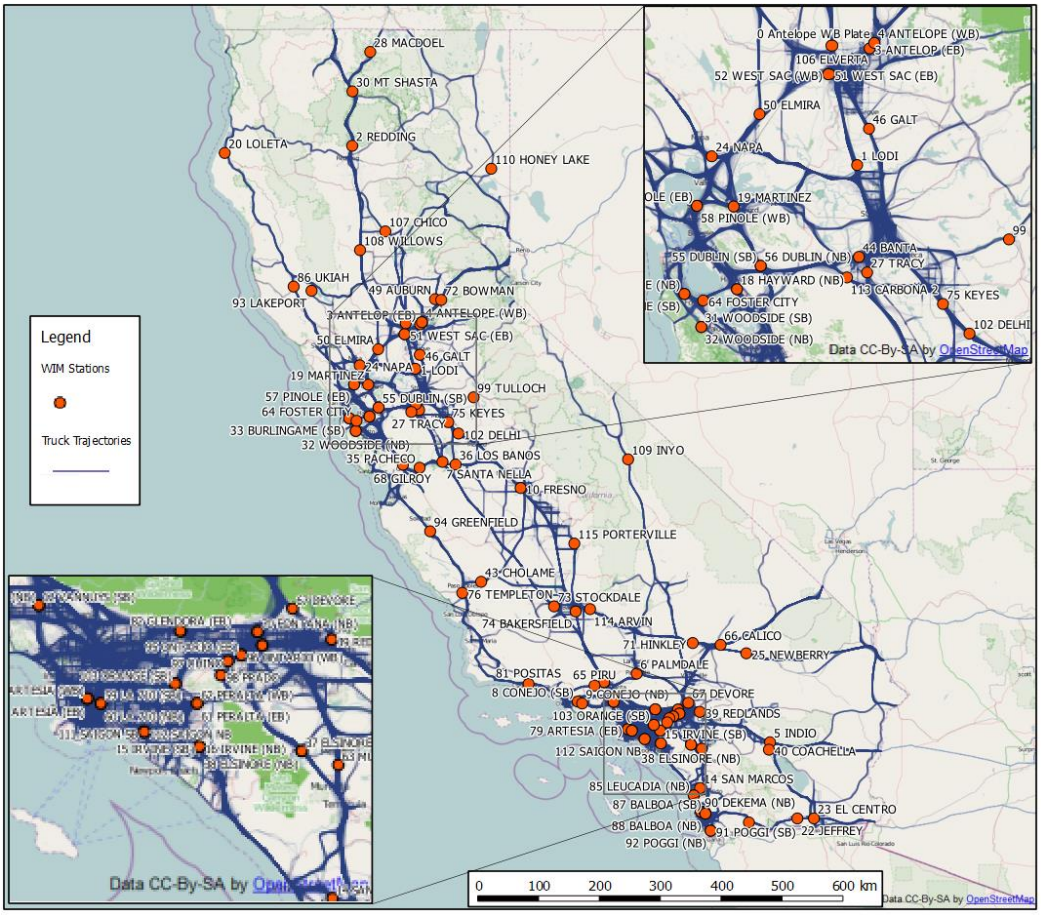


Figure 4-16. Heavy Truck GPS Trajectory Coverage and WIM Station Locations in California

4.4.3 Sensor Location Problem

The location allocation of roadway sensors is a critical issue when limited budget restricts the number sensors that can be deployed. Sensor placement problems have been extensively studied in transportation research in recent years for applications such as origin-destination (OD) flow estimation (Lam and Lo, 1990; Yang and Zhou, 1998), travel time estimation (Li and Ouyang, 2011), freeway bottleneck identification (Liu and Danczyk, 2009), network level flow estimation (Gentili and Mirchandani, 2005; Gentili and Mirchandani, 2012). These location problems and their associated models provided an ample platform from which a sensor location problem for the inductive signature technology deployment was developed.

Two methods are presented in this Section. In the first method, called the flow-interception approach, the objective is to deploy a selected number of locations that will maximize the total number of trip samples captured. Such an implementation would be desirable for an agency that would like to survey the highest possible representation of trucks traveling on the network. The second method—the re-identification approach—attempts to locate the detector stations such that deployed locations have the potential to capture not only the maximum number of trip segments, but also have the captured segments reflect the largest possible portion of each trip as well, constrained to the locations available for deployment. Vehicle re-identification aims to match vehicles crossing two different locations based on vehicle attribute data. Vehicle re-identification data can be used to estimation trip OD matrices and vehicle path flows. Inductive signature based vehicle re-identification methods have been shown to be feasible for short distances. There has been interest in developing an inductive loop signature-WIM based heavy truck monitoring system, in which data from inductive loop detectors and WIM stations are integrated[3]. Such an integrated technology can offer a great potential to identify heavy truck movements on the freeway. Cetin et al. studied the re-identification of trucks over long distance can be performed. They developed algorithms to match commercial vehicles that cross two WIM stations with a higher accuracy in Oregon by using vehicle length and axle information.

4.4.3.1 Flow-interception Sensor Placement (FSP)

The Flow-interception Sensor Placement (FSP) method is designed to maximize the total amount of net OD trips captured on a network. Genetili and Mirchandani reviewed a series of sensor location problems and categorized those problems into two main types; (1) Full Flow-Observability problem and (2) Partial Flow-Observability problem, where the former minimizes the number of sensors to cover the entire network flow level and the latter tries to maximize intercepting flow volumes given the number of sensors. In their study, two Partial Flow-Observability problems (M3 and M6) are considered for FSP application. However, M6 is found to be a greedy strategy for selecting locations because it does not consider link flow dependence, resulting in double-counting trips that traverse more than one sensor. So, M3 was used as the base model for the flow-interception sensor placement (FSP) where sensors are located on links such that the max amount of route flows are measured together with a positive fraction of trips between any observed OD pairs. The FSP formulation is as follows:

Maximize $\sum_{i \in I} \mu_i y_i$ (1)

s.t. $\sum_{j \in J_i} x_j \geq y_i \quad \forall i \in I$ (2-a)

[3] http://www.volpe.dot.gov/sbir/sol12_2/topics.html#FH4

$$\sum_{j \in J} x_j = P \tag{2-b}$$

$$\sum_{j \in J} \delta_i^j x_j \geq 1 \quad \forall i \in I_r \tag{2-c}$$

$$y_i \in \{0,1\} \qquad \forall i \in I \tag{2-d}$$

$$x_j \in \{0,1\} \qquad \forall j \in J \tag{2-e}$$

$$\delta_j^i \in \{0,1\} \qquad \forall i \in I, \forall j \in J \tag{2-f}$$

Where μ_i: flow-based weight factor for different vehicle types on path i
I and J: set of all travel paths with non-zero flows and the set of candidate sensor locations on the network, respectively.
f_i: traffic volume along path $i \in I$
$\mu_i = f_i \sum_{b \in B} w_b \rho_b^i \qquad \forall i \in I$
B: set of vehicle types that can be measured
w_b: weight factor for body type $b \in B$
ρ_b^i: proportion of vehicle type b contributing path i, $\sum_{b \in B} \rho_b^i = 1.0 \qquad \forall i \in I$
P: maximum number of sensors that can be installed on freeways

$$\delta_j^i = \begin{cases} 1, & \text{location } j \text{ intercepts flow } f_i \\ 0, & \text{otherwise} \end{cases}$$

The objective function in Equation (1) maximizes the total intercepted net path flow for all given OD pairs. Two binary decision variables, x_j and y_i are included, which determine whether the candidate location j is selected and whether path i is covered by sensor placement, respectively. Constraint (2-a) ensures that the decision variable y_i to be equal to zero if there is no sensor placed on path i. Constraint (2-b) locates exactly P sensors on links while constraint (2-c) forces a particular set of paths I_r to be covered if needed. The incident parameter, δ_j^i, indicates one if the candidate location j is on path i, and zero otherwise. It is noted that constraint (2-c) can conflict with (2-b) if P is not large enough to cover I_r ($P < |I_r|$).

A set of vehicle types, B denotes possible types of vehicles passing over the candidate sensor locations. A proportion of flow for vehicle type $b \in B$ on path i can be expressed as ρ_b^i. The flow-based weight factor, μ_i can be defined as a flow multiplied by weight factors for each proportion of vehicle types. If the weight factor, w_b is the same for all vehicle types, then the model becomes an ordinary flow-interception model considering all vehicle types equally. It is noted that without constraint (2-c), the problem represents the maximum covering location problem, which is known as NP-hard.

Heavy Truck Body Type Distribution

One of the main requirements of FSP is the truck type distribution associated with each vehicle path. However, the GPS truck trajectories do not provide information on vehicle body types and associated commodity types. Hence, the body type information was obtained from an ongoing effort to develop the California Statewide Freight Forecasting Model[4] (CSFFM). CSFFM is a commodity based model to

[4] http://www.dot.ca.gov/hq/tsip/otfa/csffm/index.html

forecast and analyze freight movement in and out of California. The model outputs include location-based commodity productions, consumptions, commodity flows, and vehicle flows by different transportation mode (truck, rail, and air). To identify trucks' OD and paths, the truck position records in truck GPS trajectories were then associated with the CSFFM's freight analysis zones (97 zones in California) based on their GPS points. In addition, each truck trajectory was projected on CSFFM's statewide transportation network to identify an ordered set of WIM stations associated with each truck route.

CSFFM payload factors and body types were employed to determine the fraction of body types associated with each OD pair. The commodity-based payload factors were derived from the California portion of the 2002 Vehicle Inventory Use Survey (VIUS), which provide conversion factors from the total tonnage value associated with each of fifteen commodity groups defined in CSFFM to the corresponding number of vehicles for each OD pair. Table 4-23 shows the eight different body types defined in CSFFM. The proportions of body types based on the California portion of VIUS are subsequently applied to CSFFM commodity groups in Table 4-24.

Table 4-23. VIUS Body Types in CSFFM

Body Types		CA portion of VIUS	National VIUS
1	Concrete mixer and concrete pumper	4.49%	3.62%
2	Dump	12.56%	21.27%
3	Flatbed, stake, platform, and low boy	27.30%	19.49%
4	Service	6.34%	4.52%
5	Tank	23.73%	8.28%
6	Trash, garbage, recycling, and vacuum	5.53%	4.21%
7	Van	17.05%	30.00%
8	Others	3.00%	8.61%
Total		100.00 %	100.00 %

Table 4-24. Body Type Distribution by Commodity Group (FHWA Class 9)

Commodity Group	Body Type	%	Commodity Group	Body Type	%
1 Agriculture products	2	0.0185	8 Manufactured Products	2	0.0787
	3	0.2407		3	0.3371
	5	0.0926		7	0.4494
	7	0.5741		8	0.1348
	8	0.0741	9 Chemical, Pharmaceutical products	3	0.0345
2 Wood, Printed products	2	0.7111		5	0.2759
	3	0.1333		7	0.6897
	7	0.0444	10 Non-metallic mineral	7	1.0000
	8	0.1111	11 Metal manufactured products	2	0.0519
3 Crude petroleum	5	1.0000		3	0.4545
4 Fuel, Oil products	2	0.2857		7	0.4545
	3	0.1786		8	0.0390
	5	0.0357	12 Waste materials	2	0.5517
	7	0.5000		3	0.1724
5 Gravel, Non- Metallic minerals	3	0.0556		7	0.2759
	7	0.8889	13 Electronics	3	0.0652
	8	0.0556		5	0.1087
6 Coal, Metallic minerals	2	0.7222		7	0.8261
	3	0.1111	14 Transportation equipment	2	0.2500
	7	0.1111		3	0.2500
	8	0.0556		7	0.5000
7 Food, Beverage, Tobacco products	3	0.2391	15 Logs	3	0.8571
	5	0.2174		8	0.1429
	7	0.5435			

The following method is applied for the calculation of ρ_b^i stated in FSP:

$$T_b^k = \frac{m^{k1}\varphi_{1b}}{p_{1b}} + \frac{m^{k2}\varphi_{2b}}{p_{2b}} + \frac{m^{k3}\varphi_{3b}}{p_{3b}} + \cdots + \frac{m^{k15}\varphi_{15b}}{p_{15b}} = \sum_{s\in S}\frac{m^{ks}\varphi_{sb}}{p_{sb}} \quad \text{(3-a)}$$

$$\rho_b^k = T_b^k / T^k \quad \text{(3-b)}$$

Where
k: k-th OD pair, $k \in K$
s: commodity group, $s \in S$
T_b^k: number of trucks in vehicle type b for k-th OD pair
m^{ks}: tonnage of commodity group s for k-th OD pair, $k \in K$, $s \in S$
φ_{sb}: faction of commodity group s with vehicle type b, $b \in B$
p_{sb}: payload of truck type b transporting commodity group s

ρ_b^k: estimated proportion of body type b for OD pair k
T^k: total number of vehicles assumed for k
T_b^k : the number of trucks in truck type b for the given OD pair k, $k \in K$

4.4.3.2 Re-identification Sensor Placement (RSP)

The objective function and constraints employed in RSP are based on the study by Sherali et al (2006) and Liu and Danczyk (2009), where it is formulated as a non-linear optimization problem. The objective function maximizes the benefits that can be represented as the utility of the placement of two sensors. In Sherali's study, benefit was represented as travel time variability while Liu and Danczyk represented benefits as non-negative speed gradients. The objective of the RSP problem for vehicle re-identification is to select optimal pairwise locations such that the benefit is derived in terms of the number of truck routes captured by the sensor pair. Unique to re-identification problems are the constraints which affect the accuracy of re-identification, such as distance and traffic flow between sensors.

The formulation for the RSP problem is as follows. Each path $i \in I$ is specified by its traffic volume f_i which is assumed to be known via GPS trajectory sampling. If f_i passes at least two sensors, the route portion between two sensors can be identified as a vehicle re-identification route. Each path i passes a set of candidate sensors on highway network, J_i.

RSP-1) Maximize $\sum_{i \in I} \sum_{h \in J_i} \sum_{e \in J_i} b_{he} x_h x_e$ (4)

s.t. $\sum_{j \in J} x_j \leq P$ (5-a)

$\sum_{j \in J} c_j x_j \leq U$ (5-b)

$x_j \in \{0,1\} \qquad \forall j \in J$ (5-c)

I: set of OD paths on the network
J: set of all candidate locations for re-identification sensors
f_i: traffic volume on path i
b_{he}: benefit factor between head (h) and rear (e) sensors
P: maximum number of sensors that can be installed on freeways
c_j: unit cost of installing sensors at any site j
U: maximum installing budget

We assign the upstream and downstream sensor traversed by a truck trip as the head ($h \in J_i$) and rear ($e \in J_i$) sensors, respectively. If there exist at least two sensors at x_h and x_e along path i, its utility can be recorded as b_{he}. The objective function (4) seeks to maximize the total benefit based on sensor placements among candidate sites, J. Constraint (5-a) asserts that the total number of sensors should not exceed the available maximum number, R, while constraint (5-b) limits the total cost of installing sensors at each site j with a unit cost c_j. The decision variable x_j is defined as binary, where 1 represents the placement of a sensor at location j and 0 otherwise in (5-c).

The benefit factor (b_{he}) constituting the objective function parameters here reflects the ability to capture highway paths regarding the variability of freight truck movement. Since the number of sensors on path i is a key factor in estimating the true OD paths, the locations of upstream and downstream sensors are critical. A path coverage concept is introduced as a benefit factor of RSP. Consider that the

length of an individual vehicle path i connecting its true origin and destination is denoted by L_i. Then, let R_i denote a set of candidate sensors on path i. These sensors play a critical role in reconstructing the individual path. If there are two sensors placed on path i, denoting $R_i = (r_h, r_e)$ for upstream (head) and downstream (rear), the identified distance by two sensors, $d_i(r_h, r_e)$ along path i can be calculated by actual vehicles trajectories along truck routes. If $d_i(r_h, r_e)$ represents a significant proportion of L_i, as described in (Li and Ouyang, 2011). On the other hand, if $d_i(r_h, r_e)$ is very short compared to L_i, it can be inferred that the two sensors do not provide much information about path i.

$$b_{he}^{d} = \begin{cases} 0, & r_e < r_h \\ \frac{d_i(r_h, r_e)}{L_i}, & \text{otherwise} \end{cases} \tag{6}$$

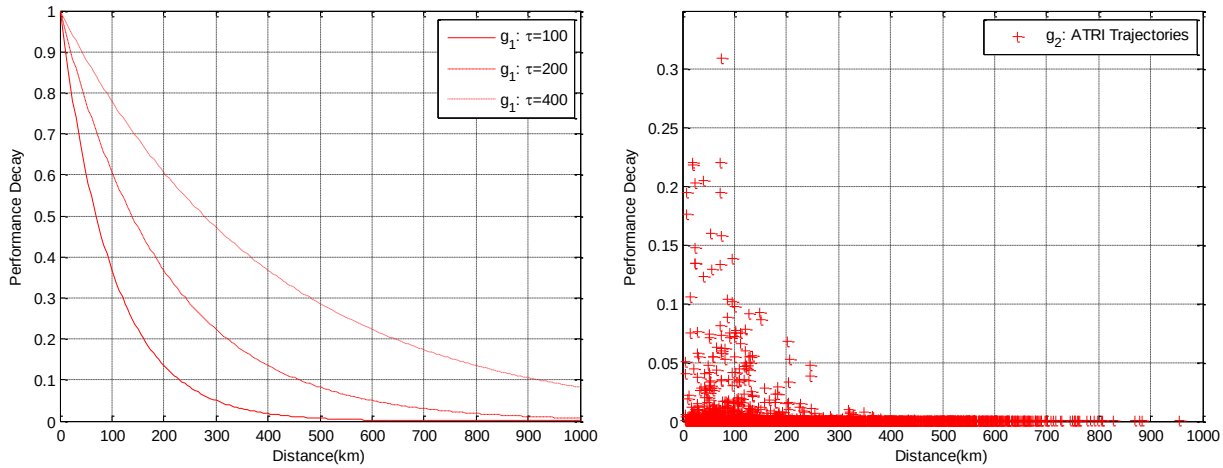
There are currently no comprehensive empirical data for re-identification performance of the integrated WIM and inductive signature technology system. However, a recent study by Cetin (2011) evaluated factors which affect the matching accuracy between WIM station pairs, such as distance, travel time variability, truck volumes, and sensor accuracy and consistency. They demonstrated that re-identification with WIM can be performed over large distances, and WIM sensor accuracy and traffic volume were identified as the major factors determining the search space at the upstream location.

In this study, the variability of travel time between upstream and downstream is considered to potentially affect the accuracy of re-identification because a search space for a downstream vehicle is determined based on the travel of detections at head and rear stations. The width of time windows usually increases as the distance between upstream and downstream detectors increases, generally yielding a lower probability of vehicle matching. Therefore, one assumption is that the error rate of re-identification increases as the distance between head and rear sensors increase in (7).

$$\text{Re-identification Error Rates} \propto d_i(r_h, r_e) \tag{7}$$

$$g_1(r_h, r_e) = e^{-s/\tau} \tag{8}$$

We assume a distance-based exponential decay function shown in Equation (8), where τ is a decay rate with a positive constant. Let s denote an input variable s for the function, g. In this study, we assume that s can be replaced with $d_i(r_h, r_e)$. Figure 4-17(a) shows three performance decay curves assumed by different parameter sets with short (τ=100), medium (τ=200), and long distance performance (τ=400).



(a) Distance-based Decay: g_1 (b) Location-specific Decay: g_2

Figure 4-17. Re-id Performance Decay Assumptions: (a) Distance-based, (b) Location-specific.

A second approach assumes that the decay factor, g_2 can be based on location-specific traffic flow characteristics between the upstream to downstream sensors. Suppose that two detectors, $[r_h, r_e]$, are installed on a freeway. Two data types can be used to estimate the likelihood of finding the correct match—traffic flow from upstream (h) to downstream (e), f_{he} and all traffic counts at both upstream and downstream sites, $f_{h\in J}$ and $f_{e\in J}$in Equation (9). The basic idea of this approach is simply assuming a search space given the number of vehicles coming from the upstream. For example, if two locations are close and the most vehicles passing the upstream location are going to the downstream location, there would be higher probability that the downstream vehicles can be found at the upstream candidate set, which means that both f_{he} and $f_{h\in J}$ are higher. On the other hand, if the majority of volumes at the upstream site take other paths instead of going to the downstream site, there would be fewer downstream vehicles coming from the upstream sites, resulting in a higher $f_{h\in J}$, but lower f_{he}. The same concept can be extended to traffic counts ($f_{e\in J}$) at the downstream site. Note that g_2 is a location-specific factor estimated from the ground truth samples. The observed traffic counts and the flow between upstream and downstream can be abstracted from truck trajectories.

$$g_2(r_h, r_e) = \frac{f_{he}}{f_{h\in J}} \cdot \frac{f_{he}}{f_{e\in J}} \tag{9}$$

Segment combinations of possible upstream and downstream WIM sites were extracted from all truck trajectories. A total of 151,760 segments were identified with 2,460 segments of unique upstream and downstream WIM site pairs. Figure 4-17(b) shows the assumed performance decay (g_2) for each combination of upstream and downstream detectors by distance. As expected, higher rates are generally found within the shorter distance range. The highest value is 0.309 corresponding to truck movement between site 30 (Mt. Shasta on I-5 South) and site 2 (Redding on I-5 South). Almost 60 percent of vehicles that passed the upstream site also traversed the downstream site. The performance decay associated with WIM site pairs is independent of the presence of other upstream or downstream sites.

Since multiple sensors can be utilized for one path, i.e. $|R_i|$ > 2, a subset of sensor pairs was assumed such that there are $|R_i|$-1 levels of pairs from which to derive benefit factors. For the example of $R_i = \{r_1, r_2, r_3, \dots, r_s\}$, there are s-1 benefit factors contributed sequentially by each sensor-based segment $\{r_1, r_2\}$, $\{r_2, r_3\}$,..., $\{r_{s-1}, r_s\}$. The total benefit for multiple sensors on a path can be summed up as described in Equation (10).

$$b_{he} = \sum_{h,e \in R_i} b_{he}^d \, g(h, e) \tag{10}$$

4.4.4 Solutions

FSP Solution

The proposed FSP problem takes a linear programming form. This is useful for finding an exact solution using conventional linear programming solvers if the problem is not extremely large in scale. A preliminary computational experience with C++ and CPLEX 12.4 Concert Library showed reasonable computational times around 5 minutes on Intel Core 2 Duo 2.6 GHz CPU and Windows XP. The origins and destinations of 131,201 truck trajectories are grouped by 97 CSFFM Freight Analysis Zones (FAZs) in California. A total of 13,009 unique OD paths were identified with 83 distinct WIM locations. Hence, each OD pair could be associated with multiple distinct paths. Truck trajectories with one end of their trip outside of California had the external end matched to the zones closest to the state boundary. The number of WIM sites traversed by paths ranged between one and eleven sites.

FSP Results for Heavy Truck Body Types

The aforementioned body type distribution in Figure 4-24 was used to estimate the distribution of body types for each commodity group. Truck body types G1, G4, and G6 were not considered as they are not associated with freight movement. The selection priorities were performed using a stepwise approach to increase the maximum number of locations selected. Table 4-25 (a) and (b) show the selected locations that result from applying six different truck body type criteria to the FSP optimization problem. As shown in the first row of the table, the 'All groups' criteria gave equal priority to all truck body types. The remaining criterions correspond to placing priority on only one of the body types (G2, G3, G5 G7 and G8) when determining the optimal locations.

Table 4-25. Comparison of FSP location selection results for different truck body type priorities

(a) 1 to 5 selected stations

Criteria		Selected WIM Sites				
		P = 1	*P* = 2	*P* = 3	*P* = 4	*P* = 5
All groups		95	95, 102	82, 95, 102	3, 82, 95, 102	3, 77, 82, 95, 102
G2	Dump	47	47, 102	47, 95, 102	44, 47, 95, 10	44, 47, 95, 102, 103
G3	Flatbed, stake, platform and lowboy	95	82, 95	82, 95, 102	44, 82, 95, 102	44, 77, 82, 95, 102
G5	Tank	95	82, 95	82, 95, 102	44, 82, 95, 102	44, 77, 82, 95, 102
G7	Van	95	3, 95	3, 77, 95	3, 77, 95, 102	3, 77, 82, 95, 102
G8	Others	82	82, 102	82, 95, 102	1, 82, 95, 102	3, 44, 82, 95, 102

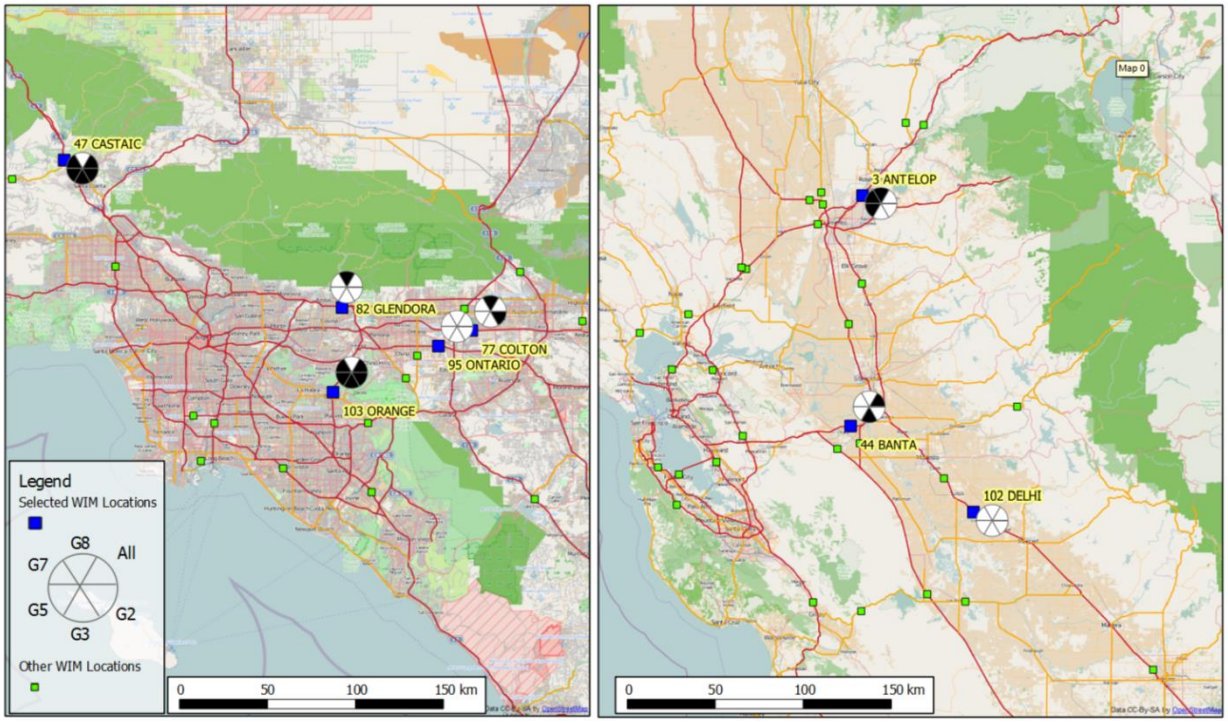
*[95]: Ontario on CA SR-60, [47]: Castaic on I-5, [82]: Glendora on I-210, [102]: Delhi on CA SR-99, [3]: Antelop on I-80

(b) 20 selected stations

Criteria		Selected WIM Sites
		P = 20
All groups		3, 5, 10, 15, 24, 29, 37, 44, 47, 57, 59, 63, 67, 69, 75, 77, 82, 95, 103, 114
G2	Dump	3, 7, 10, 15, 17, 29, 44, 47, 55, 57, 59, 61, 68, 75, 82, 87, 95, 103, 111, 114
G3	Flatbed, stake, platform and lowboy	3, 5, 7, 10, 15, 24, 29, 37, 41, 44, 47, 59, 61, 67, 69, 75, 77, 82, 95, 103
G5	Tank	1, 5, 7, 10, 15, 24, 35, 37, 41, 44, 46, 47, 59, 61, 67, 69, 77, 82, 95, 102
G7	Van	3, 5, 10, 15, 24, 29, 30, 37, 44, 47, 57, 59, 63, 67, 69, 75, 77, 82, 95, 114
G8	Others	2, 3, 5, 7, 10, 15, 29, 35, 41, 44, 47, 59, 61, 67, 69, 77, 82, 95, 102, 114

Figure 4-16 shows a comparison of site selection results using the six different selection criteria for selecting five optimal locations. Sites chosen by one or more of the selection criteria have their corresponding slices shaded white. Both sites 95 and 102 were chosen across all criteria, while site 47 (Castaic on I-5) and 103 (Orange on CA SR-57) were chosen only when optimizing the location selection for G8 (Other) trucks. The results show that site 95 (Ontario on CA SR-60) had the highest priority across when higher priority was given to groups G3, G5, and G7. This indicates that the model is sensitive to different body type priorities. The location was also reported as one of the top three ranked sites (95, 77, and 82) based on volumes of 5-axle tractors pulling semi-trailers (FHWA class 9), indicating an agreement between the truck GPS data with WIM data. Although site 95 was not initially selected under criterion G2 or G8, it was selected when the total number of sites (P) specified was three, indicating that it is generally an ideal site for flow interception. Also, site 102 (Delhi on CA SR-99) was selected across

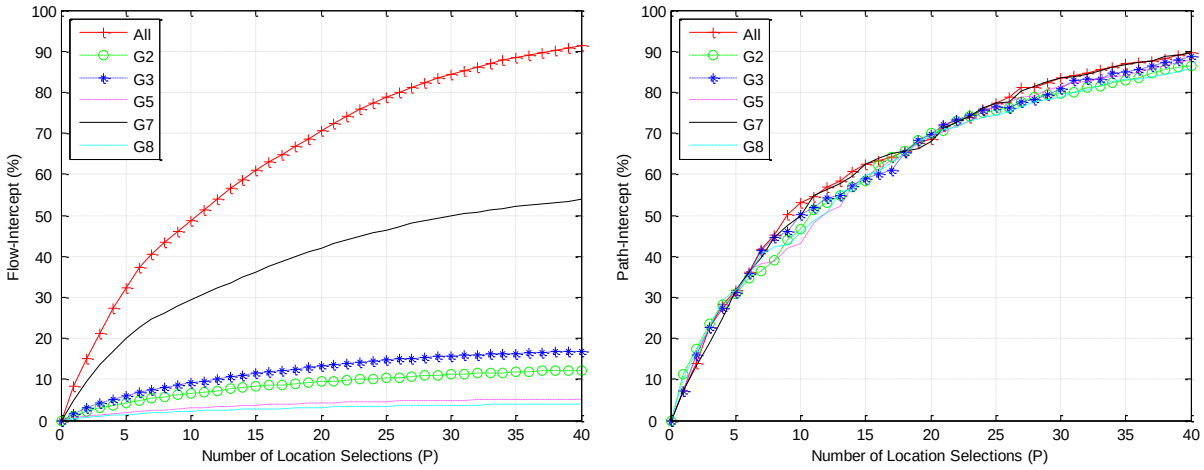
all criteria when the total number of sites was specified to be 5 sites (P = 5), but was not included when P=20 for all but two criteria: G5 and G8. This indicates that a significant proportion of truck trips traversing site 102 may be captured by one or more combinations of alternative sensors in close proximity. Indeed, site 75 (Keyes on CA SR-99), which is located 13 miles north of 102 on SR-99 in Central California, was selected for the other criteria in the absence of site 102, and both these sites were never chosen simultaneously for each criterion. This confirmed that the FSP model gives priority to selecting non-overlapped locations.



* Fully white pie chart slice denotes that the location is selected for all scenarios.

Figure 4-18. Comparison of FSP site selection criteria results for *P* = 5

A sensitivity analysis was performed across a range of sensor deployments, for the number of selected sites, P ranging from 1 to 40. Figure 4-19 shows the flow-interception and the path-interception using the six different criteria. When deploying the body type classification technology for up to 40 locations, the flow-interception can capture more than 90 percent of the given truck flows. G7 shows the dominant flows among the body types considered. When the G7 criterion is used and deployed at 40 locations, the G7 trucks captured by these selected locations represent 54 percent of total flows. There is no significant difference in path interception across all criteria.



(a) Flow Interception (b) Path Interception

Figure 4-19. FSP results with truck body types

RSP Solution

The RSP problem is a constrained mixed-integer quadratic programming problem. Since the benefit factors in RSP can vary depending on the combination of selected locations, benefit complexity makes it a challenge to solve with traditional nonlinear solver software packages. Thus, a Genetic Algorithm (GA) method was employed to solve the RSP problem. GA is an adaptive meta-heuristic algorithm used to solve combinatorial optimization problems. The GA method involves a chromosome structure that represents a solution of the problem and the solution evolves by copying chromosomes and swapping partial chromosomes over generations. Although GAs do not guarantee optimal solutions, they are efficient in seeking approximate solutions in combinatorial optimization problems. Each chromosome is made up of candidate locations representing binary decision variables. An initial population is given by randomly distributed servers over the candidate locations. For genetic operators, Elitism, Crossover and Mutation techniques are employed with the maximum number of generations (e.g., 500 generation span) for each scenario.

As shown in Table 4-26, three types of re-identification performance decay scenarios were considered, with g_0, g_1, and g_2 representing no decay, distance-based decay, trip-based location specific decay, respectively. Unlike FSP, only trips traversing two or more WIM stations on an observed path were considered. A total of 58,762 individual trips were identified to have two or more sensors among 131,201 truck paths. The average number of WIM stations on a path is 2.62 stations. 35,012 (59.6%), 15,252 (26.0%), and 5,898 (10.0%) trajectories contained two, three, and four WIM sites, respectively, while only about 2,600 (4.4%) trajectories contained five or more WIM sites.

Table 4-26. RSP location selection results

Number of Selections, P	Decay Assumption				
	g_0 (No decay)	g_1, τ =400	g_1, τ =200	g_1, τ =100	g_2 (Trip-based)
P = 2	41: Vacaville, I-80 57: Pinole, I-80	41: Vacaville, I-80 57: Pinole, I-80	41: Vacaville, I-80 57: Pinole, I-80	41: Vacaville, I-80 57: Pinole, I-80	41: Vacaville, I-80 57: Pinole, I-80
P = 3	3: Antelope, I-80 41: Vacaville, I-80 57: Pinole, I-80	3: Antelope, I-80 41: Vacaville, I-80 57: Pinole, I-80	3: Antelope, I-80 41: Vacaville, I-80 57: Pinole, I-80	3: Antelope, I-80 41: Vacaville, I-80 57: Pinole, I-80	3: Antelope, I-80 41: Vacaville, I-80 57: Pinole, I-80
P = 4	3: Antelope, I-80 41: Vacaville, I-80 57: Pinole, I-80 72: Bowman, I-80	3: Antelope, I-80 41: Vacaville, I-80 57: Pinole, I-80 72: Bowman, I-80	3: Antelope, I-80 41: Vacaville, I-80 57: Pinole, I-80 72: Bowman, I-80	3: Antelope, I-80 41: Vacaville, I-80 57: Pinole, I-80 72: Bowman, I-80	3: Antelope, I-80 41: Vacaville, I-80 57: Pinole, I-80 72: Bowman, I-80
P = 5	1: Lodi, I-5 3: Antelope, I-80 41: Vacaville, I-80 57: Pinole, I-80 72: Bowman, I-80	1: Lodi, I-5 3: Antelope, I-80 41: Vacaville, I-80 57: Pinole, I-80 72: Bowman, I-80	1: Lodi, I-5 3: Antelope, I-80 41: Vacaville, I-80 57: Pinole, I-80 72: Bowman, I-80	3: Antelope, I-80 41: Vacaville, I-80 51: West Sac, I-80 57: Pinole, I-80 72: Bowman, I-80	3: Antelope, I-80 41: Vacaville, I-80 46: Galt, SR-99 57: Pinole, I-80 72: Bowman, I-80
P = 6	1: Lodi, I-5 3: Antelope, I-80 29: Arco, I-5 41: Vacaville, I-80 57: Pinole, I-80 72: Bowman, I-80	1: Lodi, I-5 3: Antelope, I-80 29: Arco, I-5 41: Vacaville, I-80 57: Pinole, I-80 72: Bowman, I-80	1: Lodi, I-5 3: Antelope, I-80 29: Arco, I-5 41: Vacaville, I-80 57: Pinole, I-80 72: Bowman, I-80	3: Antelope, I-80 41: Vacaville, I-80 46: Galt, SR-99 51: West Sac, I-80 57: Pinole, I-80 72: Bowman, I-80	3: Antelope, I-80 10: Fresno, SR-99 41: Vacaville, I-80 ~~46: Galt, SR-99~~ 57: Pinole, I-80 ~~72: Bowman, I-80~~ 75: Keyes, SR-99 102: Delhi, SR-99
P = 7	~~1: Lodi, I-5~~ 3: Antelope, I-80 10: Fresno, SR-99 ~~29: Arco, I-5~~ 41: Vacaville, I-80 57: Pinole, I-80 72: Bowman, I-80 75: Keyes, SR-99 102: Delhi, SR-99	1: Lodi on I-5 3: Antelope on I-80 29: Arco on I-5 41: Vacaville on I-80 57: Pinole on I-80 72: Bowman on I-80 108: Willows on I-5	1: Lodi on I-5 3: Antelope on I-80 29: Arco on I-5 41: Vacaville on I-80 51: West Sac on I-80 57: Pinole on I-80 72: Bowman on I-80	1: Lodi on I-5 3: Antelope on I-80 29: Arco on I-5 41: Vacaville on I-80 ~~46: Galt, SR-99~~ 51: West Sac on I-80 57: Pinole on I-80 72: Bowman on I-80	3: Antelope on I-80 10: Fresno on SR-99 41: Vacaville on I-80 57: Pinole on I-80 72: Bowman on I-80 75: Keyes on SR-99 102: Delhi on SR-99
P = 8	3: Antelope, I-80 10: Fresno, SR-99 41: Vacaville, I-80 57: Pinole, I-80 72: Bowman, I-80 74: Bakersfield, SR-99[1] 75: Keyes, SR-99 102: Delhi, SR-99	1: Lodi, I-5 3: Antelope, I-80 29: Arco, I-5 41: Vacaville, I-80 51: West Sac, I-80 57: Pinole, I-80 72: Bowman, I-80 108: Willows, I-5	1: Lodi, I-5 3: Antelope, I-80 29: Arco, I-5 41: Vacaville, I-80 46: Galt, SR-99 51: West Sac, I-80 57: Pinole, I-80 72: Bowman, I-80	1: Lodi, I-5 3: Antelope, I-80 29: Arco, I-5 41: Vacaville, I-80 44: Banta, I-205 51: West Sac, I-80 57: Pinole, I-80 72: Bowman, I-80	3: Antelope, I-80 10: Fresno, SR-99 41: Vacaville, I-80 57: Pinole, I-80 72: Bowman, I-80 74: Bakersfield, SR-99[1] 75: Keyes, SR-99 102: Delhi, SR-99
P = 9	3: Antelope, I-80 10: Fresno, SR-99 41: Vacaville, I-80 47:, Castaic, I-5[1] 57: Pinole, I-80 ~~72: Bowman, I-80~~ 74: Bakersfield, SR-99[1] 75: Keyes, SR-99 82: Glendora, I-210[1] 102: Delhi, SR-99	1: Lodi, I-5 3: Antelope, I-80 29: Arco, I-5 41: Vacaville, I-80 46: Galt, SR-99 51: West Sac, I-80 57: Pinole, I-80 72: Bowman, I-80 108: Willows, I-5	1: Lodi, I-5 3: Antelope, I-80 29: Arco, I-5 41: Vacaville, I-80 ~~46: Galt, SR-99~~ 51: West Sac, I-80 57: Pinole, I-80 72: Bowman, I-80 75: Keyes, SR-99 102: Delhi, SR-99	1: Lodi, I-5 3: Antelope, I-80 29: Arco, I-5 41: Vacaville, I-80 44: Banta, I-205 46: Galt, SR-99 51: West Sac, I-80 57: Pinole, I-80 72: Bowman, I-80	3: Antelope, I-80 10: Fresno, SR-99 41: Vacaville, I-80 47: Castaic I-5[1] 57: Pinole, I-80 72: Bowman, I-80 74: Bakersfield, SR-99[1] 75: Keyes, SR-99 102: Delhi, SR-99
P = 10	1: Lodi, I-5 3: Antelope, I-80 10: Fresno, SR-99 41: Vacaville, I-80 47:, Castaic, I-5[1] 57: Pinole, I-80 74: Bakersfield, SR-99[1] 75: Keyes, SR-99 82: Glendora, I-210[1] 102: Delhi, SR-99	1: Lodi, I-5 3: Antelope, I-80 29: Arco, I-5 41: Vacaville, I-80 44: Banta, I-205 46: Galt, SR-99 ~~51: West Sac, I-80~~ 57: Pinole, I-80 72: Bowman, I-80 102: Delhi, SR-99 108: Willows, I-5	1: Lodi, I-5 3: Antelope, I-80 10: Fresno, SR-99 ~~29: Arco, I-5~~ 41: Vacaville, I-80 44: Banta, I-205 51: West Sac, I-80 57: Pinole, I-80 72: Bowman, I-80 75: Keyes, SR-99 102: Delhi, SR-99	1: Lodi, I-5 3: Antelope, I-80 29: Arco, I-5 41: Vacaville, I-80 44: Banta, I-205 ~~46: Galt, SR-99~~ 51: West Sac, I-80 57: Pinole, I-80 72: Bowman, I-80 75: Keyes, SR-99 102: Delhi, SR-99	3: Antelope, I-80 10: Fresno, SR-99 41: Vacaville, I-80 47: Castaic I-5[1] 57: Pinole, I-80 72: Bowman, I-80 73: Stockdale, I-5[1] 74: Bakersfield, SR-99[1] 75: Keyes, SR-99 102: Delhi, SR-99

[1]Southern California Locations

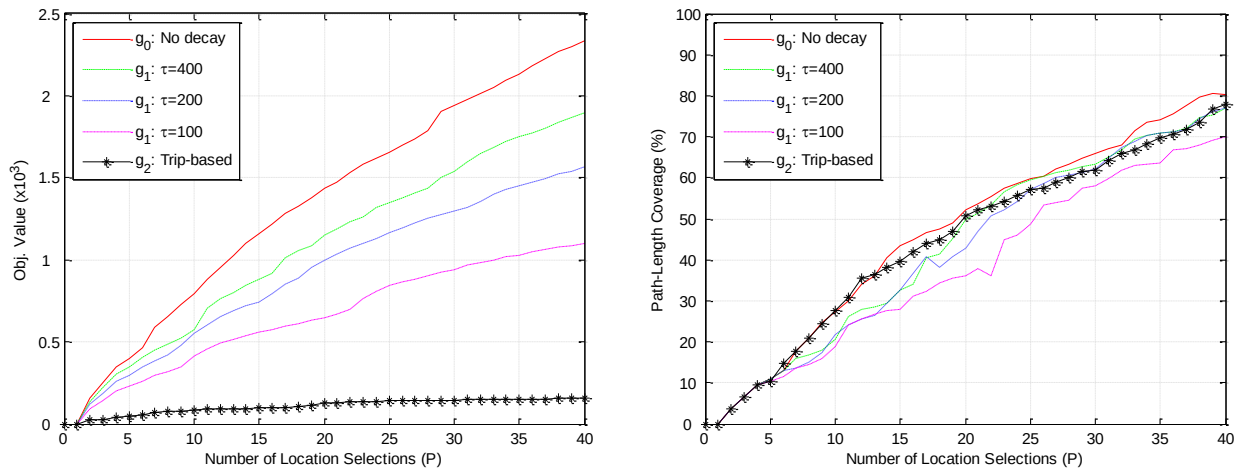
Table 4-26 reports the RSP location selection results showing between two and ten locations with different performance decay scenarios. New locations from the increase in selected locations are shown

in red. For all scenarios, sites 41 and 57 are the best candidates for the maximum re-identification benefit if only two locations are allowed within budget. These two detectors are located 58 km (36 miles) apart along I-80 in Northern California. No differences were found across scenarios in *P* = [2, 3, 4]. For *P* = 5, g_1 (τ=100) and g_2 show different results with sites 51 (g_1) and 46 (g_2), although all selected sites are in Northern California. The solution generally adds an additional site based on the previous solutions, due to the pairwise selection characteristics. There are exceptions, however. For example, g_2 drops sites 46 and 72 from *P* = 5 and adds sites 10, 75, and 102 for *P* = 6. This reveals that the location selection can significantly change depending on the benefit from different pairwise combinations. There is also a preference for Northern California locations across all scenarios, with only g_0 and g_2 choosing Southern California locations starting at *P* = 8.

Figure 4-20 shows a spatial comparison the RSP location selection results obtained with g_0, g_1 (τ=100), and g_2 for *P*=10. It shows that g_0 expands the re-identification network from San Francisco (site 57) to Los Angeles (site 82), whereas the result of g_1 remains in Northern California. For g_2, the location selection that started from sites 41 and 57 extends to site 47 in Northern Los Angeles when considering nine locations in Table 4-26 (b). The three locations selected in Northern Los Angeles, 47, 73, and 74 are critical points for identifying the truck route diversion on I-5 and CA-99, which is an intuitive result.

Figure 4-20. Comparison of RSP location selection results for g_0, g_1, τ =100 and g_2, with (*P* = 10)

Figure 4-21 shows sensitivity analysis with up to 40 sites selected. The objective values in Figure 4-21(a) are not as smooth as the curves in Figure 4-19(a). This is to be expected since the pairwise selection method is based on a meta-heuristic algorithm. However, the results are very consistent with the applied decay assumptions. Without the performance decay assumption (g_0), the result shows the highest objective values, and then τ = 400, 200, and 100 are ranked sequentially. As expected, g_2 shows the lowest objective values due to the trip-based performance decay assumption that considers a potential search space for candidate vehicles, but it is conceivable that g_2 is the most reliable because the assumption is based on the observed truck flows. Figure 4-21(b) shows path-length coverage (%) which can be defined as the distance of covered segments between upstream and downstream locations divided by the total segment length. The scenario g_0 shows covering more than 80 percent when 40 locations are installed. Not much difference in the path coverage can be seen because performance decay impacts are not taken into account in this figure, but the scenario with τ = 100 shows the lowest values.



(a) Objective Values (%) (b) Path-length Coverage (%)

Figure 4-21. RSP Results with Decay Performance Assumptions

4.4.5 Conclusion

In Section 4.4, two applications for optimally allocating the integrated WIM and inductive signature systems at existing WIM were developed: flow-interception sensor placement (FSP) and re-identification sensor placement (RSP). Truck GPS trajectories were utilized with 97 CSFFM FAZs to identify truck's OD and paths. A total of 83 existing WIM sites were considered as candidate sites. The FSP model is capable of selecting locations which emphasize different body types by employing flow-based weight factors. The assumed distribution of eight different truck body types was based on the 2002 VIUS data for trucks in California. The RSP model investigates the best locations for heavy truck re-identification by selecting pairwise locations, and is shown to be sensitive to the assumed re-identification decay factor. Based on RSP, sites 41 and 57 have a strong influence on the site selection including many candidate sites in Northern California. Candidate locations in Southern California were not selected due to the limited numbers of sensors in that region despite higher traffic counts at those sensors. This may be because few of the existing WIM sites in Southern California are located in such a way that most sensors pairs do not capture a significant number of shared flows despite high volumes at individual locations.

Although the re-identification decay factors assumed in the RSP may not be representative of true re-identification performance, they can be easily substituted when such information becomes available through future studies. The methods described in this Section were designed for the implementation of inductive signature technology at existing WIM sites, the concepts presented for FSP can be extended to the deployment of other truck surveillance infrastructure with the use of available truck GPS data. Similarly, the RSP can be applied determine the optimal locations for deploying other truck tracking technologies such as Bluetooth or RFID. In these cases, re-identification decay would not be a concern.

4.5 Propagation of Gross Vehicle Weight to Inductive Loop Detector Sites

Truck weight data is critical for pavement management, emissions estimation, and freight modeling but is not widely collected since it requires WIM or static scales. Inductive Loop Detector (ILD) sites, even those equipped to collect ILD signatures, do not measure weight nor do GPS based tracking methods. However, if weight measurements can be obtained at ILD sites then truck weight data will be available at more than just the sparse WIM site locations thus providing valuable data for the abovementioned applications. The objective of this section is to estimate gross vehicle weight (GVW) distributions at loop detector stations using a combination of body class volume and spatial relationships between ILD and WIM sites.

4.5.1 Methods

As shown in the models presented in Section 4.3 for backcasting, GVW distributions can be modeled as GMM. Each site is assumed to possess a GVW distribution following a GMM composed of three components represented by three means, three variances, and three missing components as follows:

$$f(x) = p_1 \cdot \mathcal{N}(x; \mu_1, \Sigma_1) + p_2 \cdot \mathcal{N}(x; \mu_2, \Sigma_2) + p_3 \cdot \mathcal{N}(x; \mu_3, \Sigma_3)$$

Where

$\mathcal{N}(\mu_m, \Sigma_m)$ = Gaussian distribution with mean μ and covariance matrix Σ;
$\mathrm{p_m}$ is the mixing proportion

To estimate a GMM at a new site, nine parameters (p_m, μ_m, Σ_m) need to be determined. To determine these parameters, two approaches are suggested.

First, each body class exhibits a uniquely shaped GVW distribution such that the overall GVW at a site is a mixture of the GVW distributions for each body class at that site. This means that the overall GVW distribution at a site is the volume weighted combination of each body classes' individual GVW distribution. To demonstrate, each of the five body class groups shown in Figure 4-22 consists of a three component mixture model that have been combined via a GMM of 5 by 3 (i.e. 15) components weighted by their corresponding volume to produce the overall GVW distribution shown in the bottom right of the figure. Moreover, certain body classes such as tanks and platforms exhibit unique GVW patterns. For example, tanks travel either loaded or empty for safety purposes resulting in GVW distributions with two clear peaks. Under these principles, a regression model was developed to relate body class volumes at each site to the GMM parameters (p_m, μ_m, Σ_m) for m = 1:3). Nine regression models were estimated, one for each dependent variable (i.e. each GMM parameter) using the volumes of vans, tanks, platforms, container, and 'other' trailer body class volumes as independent variables. The linear regression model takes the following form:

$$p_m = \beta_{m,0} + \sum_{b=1}^{B} \beta_{m,b} X_b$$

$$\mu_m = \beta_{m,0} + \sum_{b=1}^{B} \beta_{m,b} X_b$$

$$\Sigma_m = \beta_{m,0} + \sum_{b=1}^{B} \beta_{m,b} X_b$$

Where

m = the mth GMM component (m = 1…3)

$\beta_{m,0}$ = constant

$\beta_{m,b}$ = regression coefficient for mth GMM component for body class b

X_b = body class volumes for body class b

To produce a GVW distribution, the estimated component parameters would be combined in a final mixture as follows:

$$f_j(x) = \sum_{m=1}^{3} p_m \cdot \mathcal{N}(x; \mu_m, \Sigma_m)$$

Where

$\mathcal{N}(\mu_m, \Sigma_m)$ = the mth Gaussian distribution for site i with mean μ and covariance matrix Σ

p_m = mixing proportion

To apply the model, body class volumes at an ILD site would be estimated via the ILD signature classification model and used to predict each of the nine mixture model components. This model assumes a linear relationship between body class volume and GMM parameters and also assumes a static relationship between body class volume and GMM parameters across space. The latter is a simplifying assumption to be accounted for by the second approach.

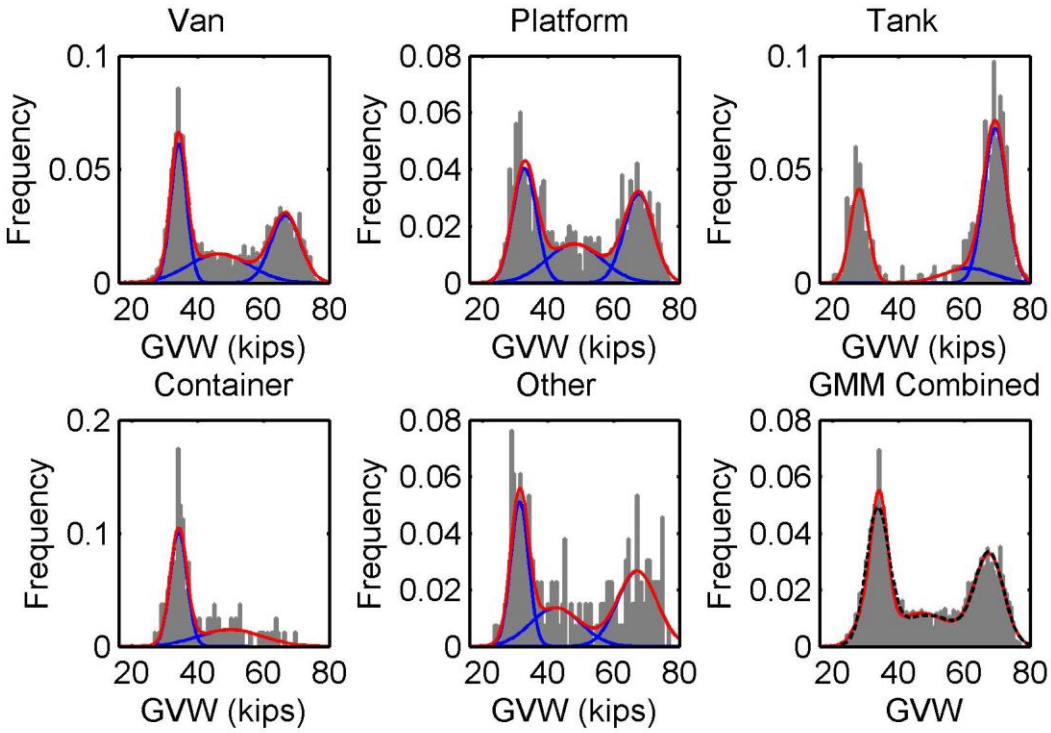


Figure 4-22 Example of GVW distribution by Body Class Group Contributing to Overall Site GVW Distribution

Second, GVW distributions are related spatially such that sensors along the same route in the same direction within the same region see similar GVW distribution patterns. Figure 4-23 shows a small subsample of GVW distributions for FHWA class 9 trucks at four WIM sites in Northern California in the southbound direction along I-5 and SR-97. Each of the sites has a significant volume of loaded trucks. This is due to shared trips along common routes and commodity flow patterns within a region. Given a reasonable assumption about spatial relationships between sites, a GMM can be estimated at an ILD site by combining the GMMs at each of the sites that are spatially related to the ILD site using the assumed spatial distances as weights in the mixture model. Thus, for site j, the GMM components μ and Σ can be estimated as follows:

$$f_j(x) = \sum_{i=1}^{N} \sum_{m=1}^{3} W_{i,j} \cdot \mathcal{N}\left(x; \mu_{1,i}, \Sigma_{1,i}\right)$$

Where

N = number of neighboring sites, i = 1…N

$\mathcal{N}(\mu_{m,i}, \Sigma_{m,i})$ = the mth Gaussian distribution for site i with mean μ and covariance matrix Σ

$W_{i,j}$ = spatial distances between sites i and j used as a mixing proportion

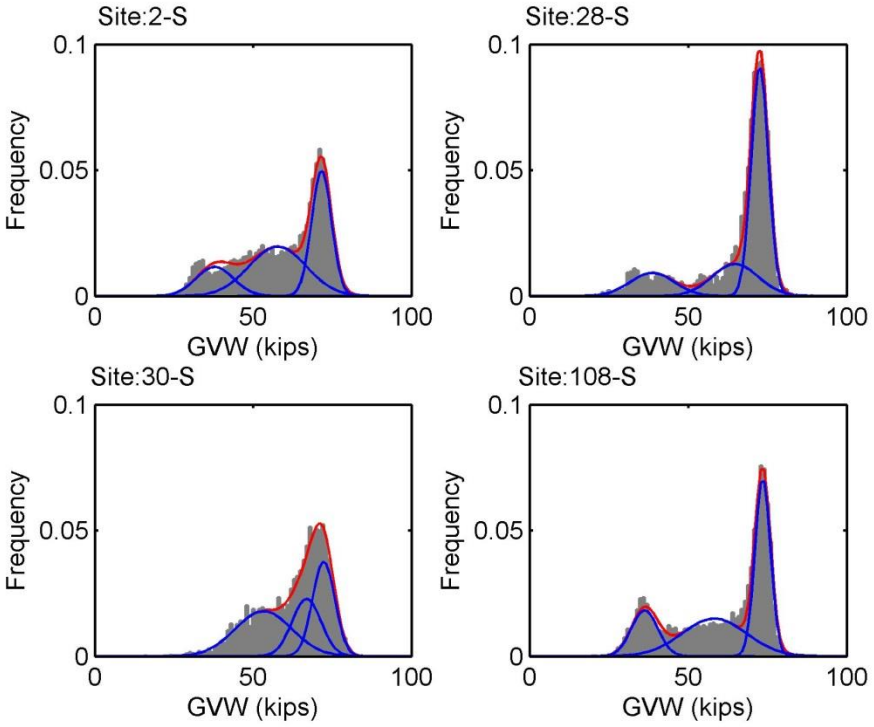


Figure 4-23 GVW Distributions along Southbound I-5 and SR-97 in Northern California

A common method to assess spatial relationships is to use coordinate locations, however this fails for directional data where sites of opposite directions share the same coordinate but tend to have very different GVW distributions. For example, at the port of Long Beach, most incoming container traffic in the westbound direction is heavily loaded while the outgoing container traffic in the eastbound direction is empty. Using route based distances would also not account for bi-directional sites, nor would it accurately capture the spatial relationship for trucks. For example, I-5 and SR-99 are parallel routes in the central valley and sites located at Willows along I-5 and Chico along SR-99 are only 90 miles apart by highway routes but would never reasonably see the same trucks and therefore might not have the same GVW weight distribution patterns.

In order to derive reasonable spatial relationships between sites, GPS data from the American Trucking Research Institute (ATRI) was used to assess the number of shared trips between sites. The ATRI truck GPS data set was considered because its samples comprise fully of truck speed data. However, both the sampling period and sampling frame are limited. The samples were obtained from four two-week periods from the month of February, May, August and November in 2010, and were limited to trucks that subscribe to the program. Trucks in ATRI tend to be long distance hauls, hence, there might be bias in the spatial patterns derived from the ATRI data. Still, ATRI GPS data can be used to evaluate the number of shared trips between WIM and ILD sites.

4.5.2 Data

A spatially diverse set of WIM sites with GVW data was needed to model the proposed two solutions described above. The data collected for body classification modelling consisted of only four sites spread across California and was therefore not dense enough to use to estimate a reasonable model. Instead,

archived WIM data from 2010 was used. Since WIM data contains GVW, axle spacing, and vehicle length measures but not body class, the WIM-only body class volume model (Section 5.2) was applied to estimate body class volumes of FHWA class 9 trucks. In summary, GMM parameters ((p_m, μ_m, Σ_m) representing the GVW distribution and body class volumes (vans, platforms, tanks, containers, other) were estimated for all 114 sites included in the archived WIM data for the year 2010.

The time aggregation level was determined by examining the time of day (off peak, AM peak, midday, and PM peak), day of week (Monday through Friday), and seasonal (fall, spring, summer, winter) changes in GVW distribution parameters. For exposition purposes, the models were developed for data disaggregated to the midday (10am to 2PM) time period on Wednesdays in the Fall season.

Finally, WIM system measurements can contain systematic measurement error due to sensor calibration issues. Although there are sophisticated methods (Jeng et al., 2015) by which to assess calibration issues, a simple approach of normalizing each vehicles GVW measurement by the weight of the steering axle was employed for this analysis. The steering axle has a relatively static weight across locations and body types so it can be used as a reference for calibration error. All GVWs displayed in the following tables and figures reference the normalize GVW.

ATRI GPS pings converted to truck trip trajectories have poor resolution due to the 15 minutes between consecutive pings. As a result, truck trip trajectories (shown as green lines in Figure 4-24) could not be directly "snapped" to the road network and linked to WIM sites to derive the spatial weight matrix. Instead, screenlines were manually drawn at each of the sites for each direction to capture the truck trip trajectories passing through a site. After placing screenlines, the number of truck trip trajectories that passed through each pair of WIM sites were counted and converted into a directional spatial weight matrix (a sub-sample shown in Table 4-27) with cells(i,j) representing the number of shared trips from WIM site i to WIM site j. Figure 4-25 shows an example of the resulting spatial relationships arising from the screenline capture approach for the northbound Lodi WIM site located along I-5 in the Sacramento area. The site with the most shared trips is WIM site number 105 with 969 shared trips.

Figure 4-24 ATRI Truck Trip Trajectories

Table 4-27 Sample of the Directional Spatial Weight Matrix from GPS Truck Trip Trajectories

Row Labels	1_N	1_S	10_N	10_S	100_S	101_N	102_N	102_S
1_N		0	70	0	0	0	135	0
1_S	0		0	64	0	0	0	128
10_N	70	0		0	0	0	1611	0
10_S	0	64	0		0	0	0	1506
100_S	0	0	0	0		0	0	0
101_N	0	0	0	0	0		0	0
102_N	135	0	1611	0	0	0		0
102_S	0	128	0	1506	0	0	0	

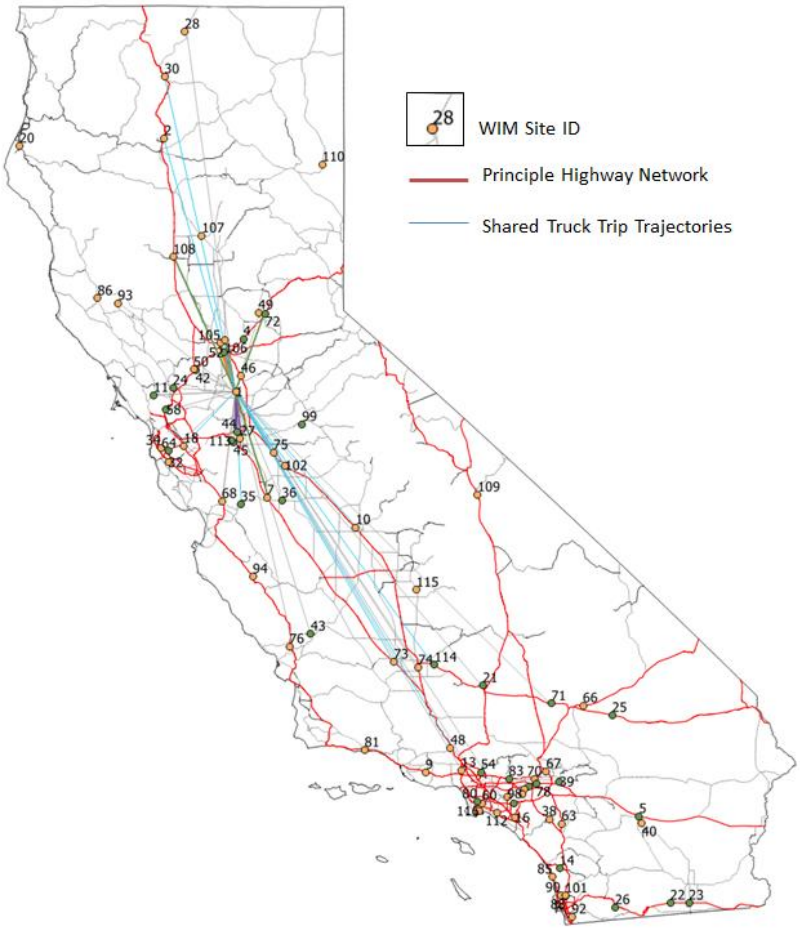


Figure 4-25 Shared Truck Trip Trajectories with the Lodi Northbound WIM Site as an origin

4.5.3 Results

The Kolmogorov-Smirnov test was used to assess the fit of the GVW distributions resulting from the regression and spatial weight based methods. The null hypothesis was that the estimated distribution is from the same continuous distribution as the observed GVW data. The alternative hypothesis was that the estimated distribution is from a different continuous distribution than the observed GVW data. Testing was performed by holding out one site from model development (i.e. estimation of the regression coefficients) and then applying estimated model to the data from the held out site. This process was repeated for all 112 of the sites in the dataset. Two sites were identified as outliers based on their observed GVW mean and variance, leaving 110 viable samples for modeling. In total, 93 stations had shared truck trajectories from which a GMM based on spatial weights could be estimated. For the regression based approach, 65.2% of the sites failed to reject the null hypothesis ($\alpha = 0.05$). This means that for 72 of the 110 sites their estimated GVW distributions matched their observed GVW distribution. For the spatial weight based approach, 67.7% of the sites failed to reject the null hypothesis ($\alpha = 0.05$) meaning that for 63 of the 93 sites the estimated GVW distribution matched the observed distribution.

Lastly, the regression and spatial weight based models were combined to produce a final GMM model for each site by weighting each of the models equally. For the combined model approach, 65.2% of the

sites failed to reject the null hypothesis. Detailed results show that for some cases, where both the regression and spatial weight models failed to match the observed data, the combined provided a suitable match.

Results for a sample of directional GVW distributions are presented in Figure 4-26 for WIM site 10 along SR-99 in Fresno located in central California and site 22 along I-8 near the southern California-Mexico border. The observed GVW distributions at each of these sites are quite different. However, with the exception of the regression model at westbound 22, all models produced statistically significant matches to the observed data.

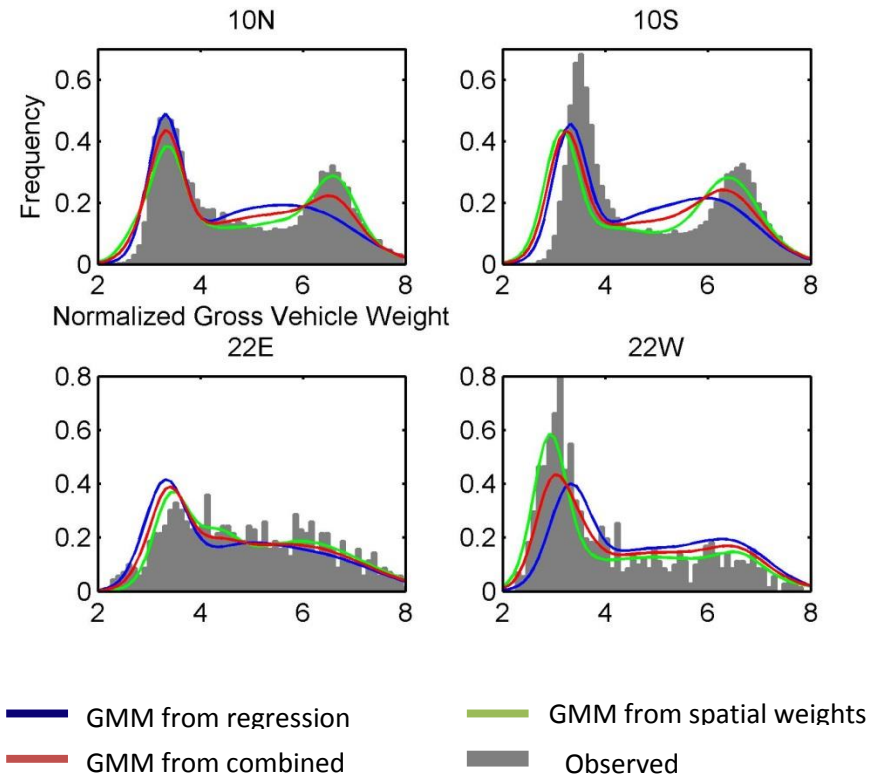


Figure 4-26 Spatial Interpolation of GVW Distributions Results

4.5.4 Conclusions

The spatial interpolation of GVW distributions is an example of the type of analysis that can be performed with knowledge of weight data. The results indicate that by combining body class volume estimates and incorporating spatial relationships between WIM sites, GVW distributions can be estimated with relative accuracy. This application serves as a baseline for what can be accomplished in regards to weight interpolation.

Several key areas for improvement should be undertaken in future studies. First, a more advanced normalization approach may be applied to the GVW data at each site prior to determining GMM parameters. Second, an obvious improvement will arise by predicting body class volumes based on ILD signature data rather than WIM data alone. More accurate GVW distributions should be possible with

improved classification data. Also, with the WIM-Signature based body classification method, GVW distributions can be stratified by body class, so that the spatial weight combination method can pull from the stratified GVW distributions rather than the overall distribution. Third, spatially based regression approaches could be evaluated rather than using a global regression model.

4.6 Summary and Conclusions

The method presented in this Chapter to integrate WIM and ILD signature technologies and the resulting models developed to predict detailed truck body class will uniquely fill the existing gaps in freight and air quality monitoring data sources by providing complete, temporally continuous, and spatially diverse truck characters data along major truck corridors. A Multiple Classifier System (MCS) approach which combined the predictions of five independent classifiers including a multilayer feed forward neural network, a probabilistic neural network, a support vector machine, a decision tree, and a naïve Bayes classifier combined using Naïve Bayes Combination (NBC) was used to improve classification generalization. The MCS with NBC proved to be more accurate than any single classifier used in the ensemble and in some cases exceeded the accuracy of even the best classifier in the ensemble. Comparisons to previous classification models using ILD signature are difficult to make due to the low number of commercial vehicle samples used in previous efforts, however, the number of body classes depicted in this Project far exceed any found in existing research. In regards to this point, previous vehicle classification problems suffered from a lack of a comprehensive set of commercial vehicle data. In this project, models were developed from around 33,000 commercial vehicle records representing over 60 body classes. The data itself is a valuable asset for truck characteristics analysis as it contains a very diverse array of body types, time periods, and locations.

5 System Hardware and Data Architecture

5.1 Overview

One of the main purposes of this study was to develop advanced truck classification models that are based on the integration of Weigh-in-motion (WIM) and inductive signature data, and on standalone inductive signature data. The system hardware and data architecture components developed in this study are described in this chapter. Additionally, since the data for model development for the integration of WIM and inductive signature data was obtained solely from the older 1060-series WIM controllers, there was a need to demonstrate that the signature data obtained from the 1060-series controllers are similar to the newer iSinc WIM controllers to ensure that the models are cross-platform compatible for future developments.

5.2 Design of WIM and Inductive Signature Hardware

Two methods of signature integration for the 1060-series WIM controller were initially considered. The first and preferred option was to develop an adapter interface to replace the existing bivalent 1060 Loop Sensor Module (LSM) module with IST-222 inductive loop signature detector cards. The second less desirable option was to splice into the inductive loop sensor leads such that the inductive loop leads would directly connect to both the 1060 LSMs as well as IST-222 detector cards. However, the concern was that the splice would adversely affect the inductive loop measurement accuracy for both detector cards.

Investigation of the potential inductive signature integration with the iSinc controller was undertaken through a trial inductive signature data collection to investigate the quality and compatibility of the inductive loop signature data obtained from iSinc controllers. This procedure helped to determine if the inductive signature data obtained from iSinc controllers would be suitable for vehicle classification applications. However, since the inductive signature logging is a proprietary feature of the iSinc controller, further development of the inductive signature feature for operations depended on International Road Dynamics (IRD), the developers and vendors of iSinc, which would require significant developmental time and costs. Because of the perceived challenges to get iSinc controllers operational with inductive loop signature data within the proposed timeline of this study, the hardware integration effort was focused on the 1060-series controller. Furthermore, at the commencement of this study, 1060-series controllers were deployed at about 80 percent of current WIM sites within the State of California. Hence, despite their age, a hardware integration solution with the 1060 series controllers would cover a much larger number of candidate sites that were available for deployment consideration.

5.2.1 Integration Design for 1060-series WIM Controllers

From the pin-out specifications of the 1060 LSM obtained from IRD as well as the IST-222, it was determined that the IST-222 was a viable candidate for adapting to the 1060 LSM. Although the voltage supplied to the 1060 LSM is 12V, as opposed to 24V which is typically supplied to 222 detector cards from traffic cabinets, the IST detector cards were designed to handle input voltages ranging between 10-24V. Hence, the lower supply voltage from the 1060 WIM controller was not expected to be an issue. However, the 1060 LSM is a four channel detector, while the IST-222 possesses only two detector

channels. To address this difference, an adapter was created to allow two IST-222 detector cards to replace an existing 1060 LSM to enable signature data processing from all four detector channels. The cross-assignment between the 1060 WIM controller LSM and two IST-222 ILD cards is provided in Appendix E. A prototype IST-222-to-LSM adapter was subsequently designed following this assignment scheme to adapt two IST-222 detector cards as a replacement for a single 1060 WIM LSM. The adapter was developed as shown in Figure 5-1, which comprises a 64-pin connector mounted on a 4x6 in PCB which shares an identical form factor of the 1060 WIM LSM and connected to two 15-pin VGA-style female connectors, with each connector designed to connect to an IST-222 detector card via a 44-pin edge connector interface.

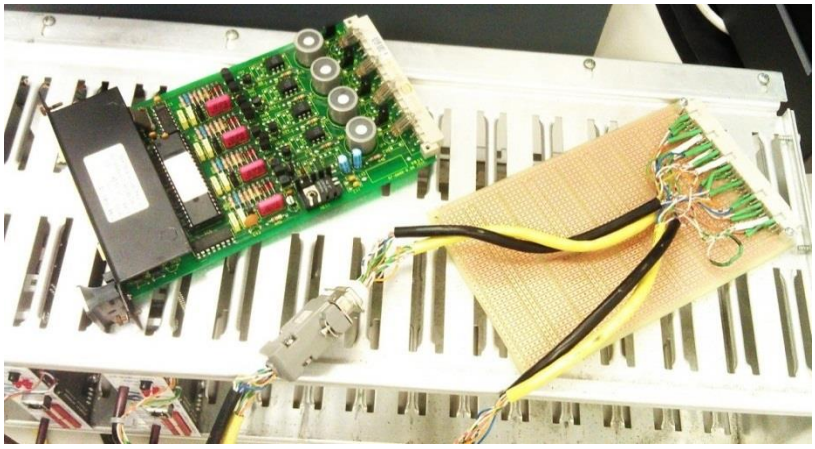


Figure 5-1 1060 WIM Controller Loop Sensor Module (Left) and IST-222 Loop Sensor Module Adapter (right)

A modified 222 input file was fabricated to interface with the IST-222 LSM adapter as shown in Figure 5-2. The back panel of the input file was removed and replaced with individual 44-pin edge connectors for each IST-222 detector card. Each edge connector was wired to a multi-pin connector for the purpose of interfacing with the LSM adapter. Because detector cards plugged into the modified 222 input file draw power directly from the 1060 WIM controller, the modified input file does not connect to an external power supply. Inductive loop signature data is logged into a field processing unit via the USB port located on the front panel of each IST-222 detector card. Schematic layouts comparing the hardware setup for a standalone 1060 WIM controller and the proposed integration with IST-222 detector cards is presented in Figure 5-3.

Figure 5-2 A modified 222 Input File fabricated for the IST-222 LSM Adapter

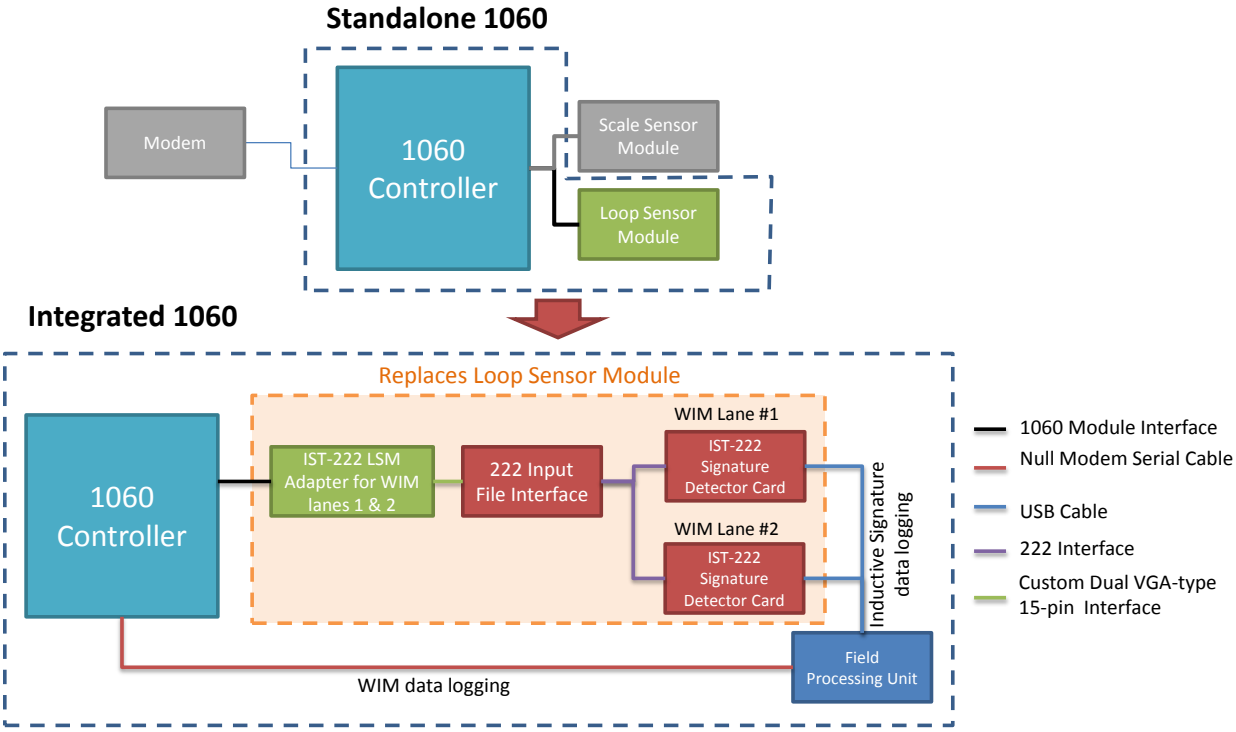


Figure 5-3 Comparison of Hardware Setup for Standalone 1060 WIM Controller (top) and 1060 WIM Controller Integrated with IST-222 Detector Cards for Inductive Signature Data Logging

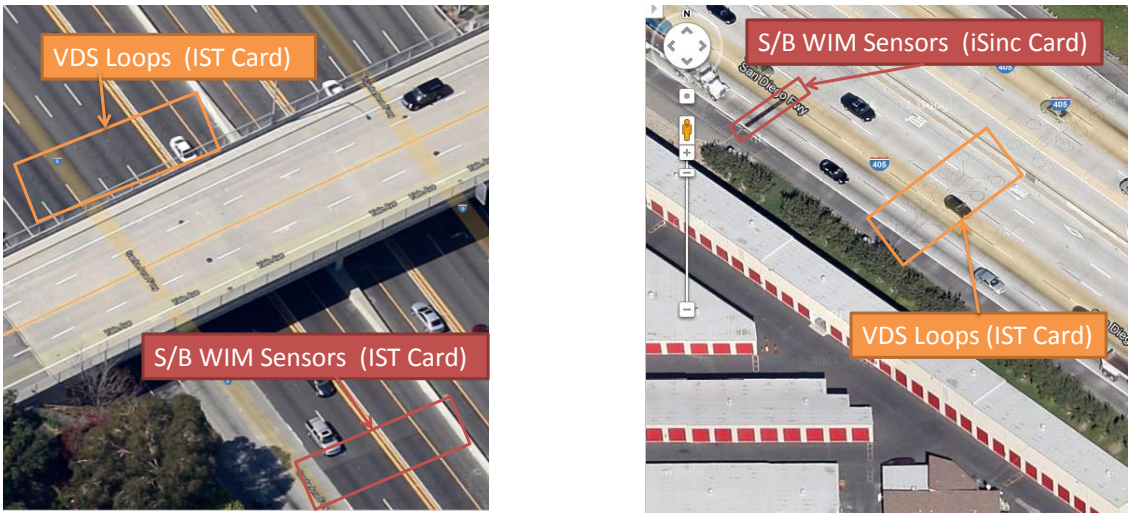
5.2.2 Comparison of Inductive Signatures between 1060 and iSinc WIM Controllers

As previously stated, there are two types of WIM controllers deployed in California. However the data used to develop the body class models in this study were collected at only one type of controller, the 1060-WIM series controller. Although the majority of WIM controllers in California at the beginning of his study were 1060 controllers, they were in the process of being replaced with the iSinc controllers. As described in the previous sections, the 1060 controllers will need to be equipped with IST-222 ILD cards to enable inductive signature data processing, while iSinc controllers have the built-in capability for sampling and processing inductive signature data, although this feature is currently available only in diagnostic mode. Because the inductive signature data logging hardware differs and performs under different sampling rates for these two controllers, there may be some differences in the quality of the inductive signatures obtained. This evaluation task was performed to determine if the models, which were developed from data collected from the 1060 series WIM controllers, are compatible with the iSinc controllers, thus eliminating the need to develop different classification models for each controller type.

5.2.2.1 Data

The ideal comparison would require both types of controllers, i.e. a 1060 and an iSinc to be located in close proximity to each other. This would allow samples of inductive loop signatures of the vehicles to be captured by both controllers. Hence each vehicle that traverses the study site would generate two signatures – one from each controller type – for direct comparison. Unfortunately, such a configuration is not available in the State of California. Hence, an alternative experimental setup was designed using two independent sites – one for each controller type – with an adjacent Inductive Loop Detector (ILD) site equipped with IST-222 ILD cards. Unlike the inductive loop sensors at the WIM sites which have a 6 foot *square* loop configuration, the inductive loop sensors at the adjacent ILD sites have a 6 foot *round* loop configuration. Consequently, inductive signature features obtained between the WIM and ILD sites will inherently possess some differences due to the geometric differences in the loop configuration. However, it can be concluded that inductive signatures obtained from the two WIM controller types are compatible, if the difference in inductive signature pairs obtained between each WIM controller and their adjacent ILD location are similar, indicating that the differences are attributed only to the loop geometry configuration between WIM and ILD locations, and not due to the controller hardware itself. To compare inductive signatures between the two hardware systems, a statistical test which removed the effects of differing loop configurations (round vs. square) was performed.

Inductive loop signatures were collected from the 1060 WIM controller equipped with IST detector cards and the iSinc WIM controller at Yale (I-5 SB) and Westminster (I-405 SB) sites pictured in Figure 5-4. At the Yale location, the ILD and adjacent WIM sites were both equipped with IST inductive loop detector cards. This location is referred to as 'IST-IST'. At the Westminster location, the ILD site was equipped with IST cards while the WIM site was equipped with iSinc LSMs. Hence, the location is referred to as 'IST-iSinc'. Since the separations between the ILD and WIM sites at both locations were less than 100 feet, each of these locations were ideal for obtaining samples of inductive signatures from the same truck across the WIM and adjacent ILD site.



(a) **(b)**

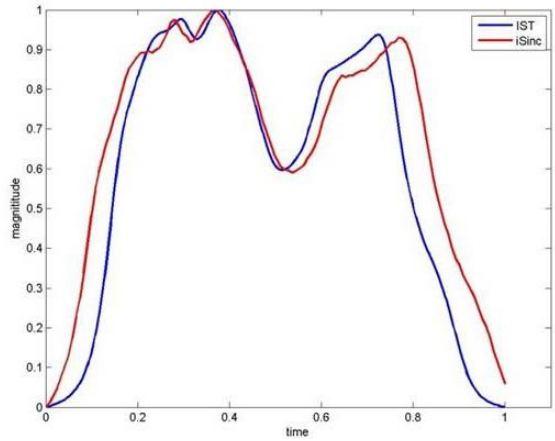
Figure 5-4 Experimental Setup: I-5 Yale location referred to as IST - IST (a), and I-405 Westminster location referred to as IST - iSinc (b)

5.2.2.2 Methodology

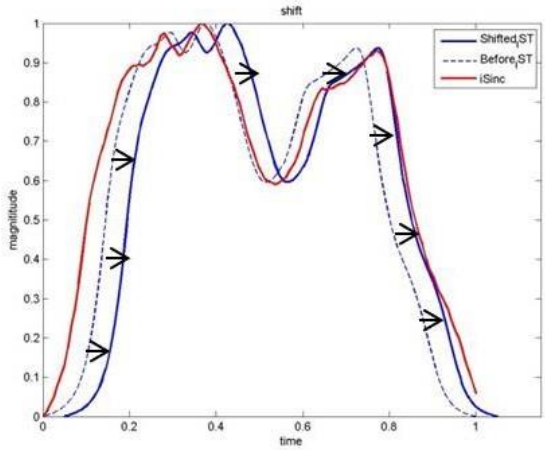
A total of 400 and 97 vehicle records were used in the comparative analysis at the Yale and Westminster locations, respectively. The analysis was performed in two steps: (1) Signature Transformation and (2) Statistical Comparison.

Signature Transformation

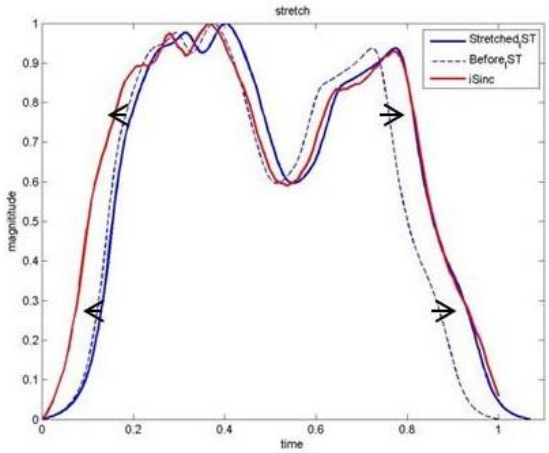
In order to remove effects from the geometric loop configurations, scaling of the signatures by the controller/detector cards, speed differences, and lateral positioning over the loop between the two sites, a signature transformation step preceded the statistical comparison. To normalize each signature, the sampled magnitudes and durations were divided by their maximum values. The second step, referred to as 'shift and stretch', involved using the WIM signature as a reference and horizontally shifting and stretching the ILD signature to achieve the best fit—more specifically, the minimum sum of vertical differences—between the two signatures.



(a)



(b)



(c)

Figure 5-5 Signature Transformation: Normalization (a), Shift (b), and Stretch Step (c)

Statistical Comparison

A two tailed t-test of the log transformation of the median and 85th percentile errors from the IST-IST and IST-iSinc was used to statistically compare the transformed inductive loop signatures. A log transformation was applied to normalize the median and 85th percentile error to satisfy the normality requirements of the t-test. The null and alternate hypothesis for the t-test is shown below:

$$H_0 : \mu_{e_IST} = \mu_{e_iSinc} \qquad H_1 : \mu_{e_IST} \neq \mu_{e_iSinc}$$

Critical values for the t-test were 0.5532 (p= 0.5820) for the median error and -1.3505 (p=0.1791) for the 85th percentile statistical comparisons at the 5% level of significance. Thus, the statistical tests confirm that the inductive signatures from the 1060 WIM controller equipped with IST-222 inductive loop detector cards and iSinc LSMs were similar. Specifically, at the 5% significance level, inductive signatures from the 1060 and iSinc were not statistically different. Therefore, this result concludes that the classification model developed from IST-222 inductive loop detector cards at the 1060 WIM sites is expected to be applicable to WIM sites equipped with iSinc controllers.

5.3 Hardware Configuration for Deployment

5.3.1 Inductive Loop Detector Sites

The components required for deployment of the truck body classification system at existing ILD sites are as follows:

- Inductive Signature Loop Detector Cards
- Field Processing Unit
- USB Cables (1 required per detector card)
- High-speed (LTE) Wireless Modem

Each inductive signature loop detector card can monitor up to two inductive loop sensors. Hence, the number of inductive signature cards required for deployment at ILD sites depending on the physical inductive loop configuration installed at each site. At dual loop sites, one inductive signature loop detector card is required for each lane that is to be monitored for truck traffic. At single loop sites, each inductive signature loop detector card may be able to monitor up to two lanes, if the inductive loops at both lanes are wired up to the same detector card slot in the detector input file. USB cables connect each inductive signature loop detector card to the field processing unit, which processes the signature data stream from the inductive signature loop detector cards. The field processing unit subsequently transmits the inductive signature data via a high-speed wireless modem to the communications server at UCI-ITS for further processing.

The field processing unit uses a fan-less design that contains no moving parts. The system dissipates heat primarily by conduction through the heat conducting fins on the external case. Data is stored in a solid state drive in place of a conventional hard drive. This solid state system design helps to eliminate common causes of system failures that are associated with overheating due to failed fan bearings, or from frost effects which may damage conventional hard drives.

The field processing units are configured with an Intel Core2-Duo dual core processor and 2GB of memory. It is also equipped with multiple USB ports to interface with the inductive signature detector cards and wireless Modem.

ILD sites are typically configured to perform one or both of the following functions: Traffic monitoring and traffic census. Sites configured to perform traffic monitoring are equipped with a traffic controller, which processes conventional ILD data and transmits traffic information continuously to a central Traffic Management Center (TMC) to monitoring of traffic conditions on the road network. Such sites are typically called Traffic Monitoring System (TMS) sites. Census sites equipped with inductive loop sensors are only configured to perform traffic census, and are typically used periodically to collect traffic volume counts. Axle-based classification counts are also collected at sits instrumented with piezo sensors. Hence, ILDs and traffic controllers are not permanently installed at these sites.

(a) Hardware Deployment at Traffic Monitoring System Cabinet

(b) Hardware Deployment at Census Cabinet

Figure 5-6 Deployments at Inductive Loop Detector Sites

5.3.2 WIM Sites

The components required for deployment of the truck body classification system at existing WIM sites equipped with 1060 series controllers are as follows:

- Inductive Signature Loop Detector Card Interface
- Inductive Signature Loop Detector Cards (1 required per monitored lane)
- Field Processing Unit
- USB Cables (1 required per card)

- Serial Null Modem WIM Data Interface Cable
- High-speed (LTE) Wireless Modem

In addition to the components required for ILD deployment, WIM deployments further require an inductive signature loop detector card interface and a serial null Modem WIM data interface cable. The inductive signature loop detector card interface consists of one or more 1060 loop sensor module (LSM) adapters which are installed in place of existing LSM cards in the 1060 series controller. These LSM adapters are connected to a modified 222 input card file for housing the inductive signature loop detector cards.

One inductive signature loop detector card is required for each lane to be monitored for truck traffic, which typically corresponds to the two rightmost lanes in each direction of travel. USB cables connect each inductive signature loop detector card to the field processing unit, which processes the signature data stream from the inductive signature loop detector cards. The field unit is also connected directly to the WIM controller via a Null Modem interface cable. This cable receives raw WIM data that has been processed by the 1060 controller. The field processing unit subsequently transmits both WIM and inductive signature data via a high-speed wireless modem to the communications server at UCI ITS for further processing.

The field processing units deployed at WIM sites are also based on a fan-less design which dissipates heat primarily by conduction through the heat conducting fins on the external case. Data is stored in a solid state drive in place of a conventional hard drive. This solid state system design helps to eliminate common causes of system failures that are associated with overheating due to failed fan bearings, or from frost effects which may damage conventional hard drives.

For WIM deployments, the field processing units are configured with an Intel i-3 dual core processor and 4GB of memory. The units are also equipped with RS-232 serial ports to interface directly with the 1060 series WIM controller, and USB ports to interface with the inductive signature detector cards and wireless Modem.

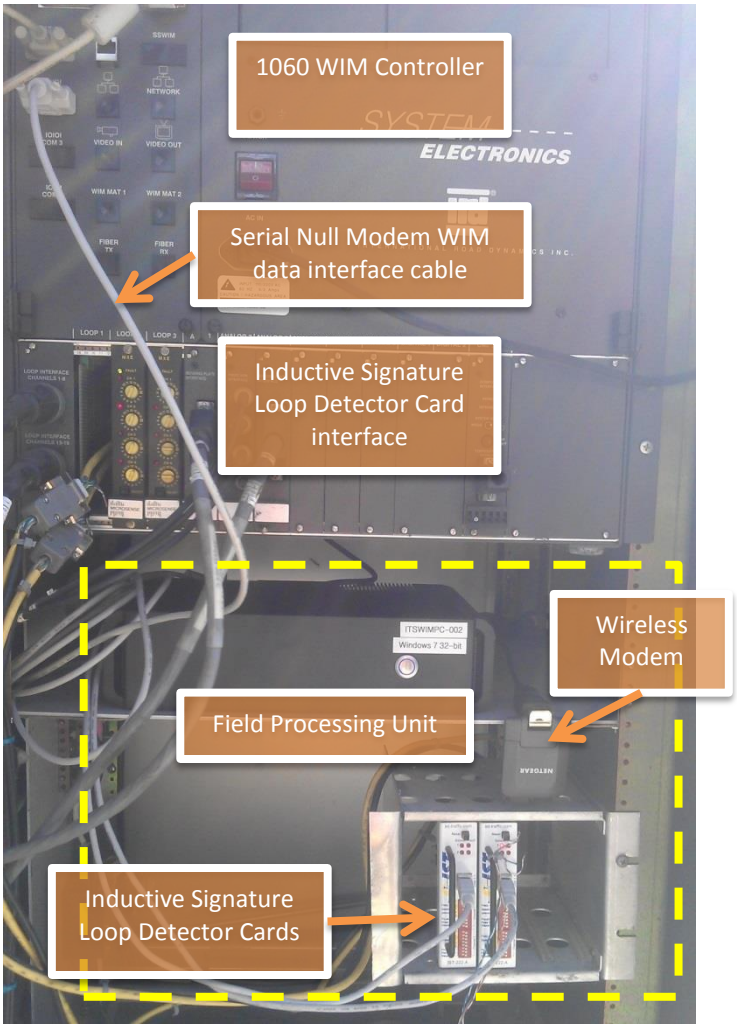


Figure 5-7 Deployment Hardware from NB 1-405 at Saigon

5.4 Data Communications Architecture

5.4.1 Overview

The signature analysis system depends on retrieving signature and WIM data from units installed at the roadside throughout the state. The primary issue to resolve is how to send raw data from the field units to a central server, where they can be processed, analyzed, and stored for future use.

An earlier prototype implementation of a system for transmitting inductive signatures for vehicle re-identification, created an infrastructure for transmitting this data using the Common Object Request Broker Architecture (CORBA) over a private TCP/IP subnetwork. In this earlier implementation, software installed on the field unit performed the following steps:

1. Using proprietary libraries from Inductive Signal Technologies (IST), connect to the proprietary IST server software running on the unit to receive data streaming from the IST signature cards
2. Process the streaming signatures packets into complete signatures
3. Process the complete signatures into lower-dimensional "feature vectors"
4. Transmit the feature vectors back to the central server using ORBIX's implementation of the CORBA event service to send structured, strongly typed data over the TCP/IP subnetwork

This similar implementation provided a useful starting point for solving the signature transmission problem in this project. In deploying the current system, however, a number of issues presented themselves as follows:

1. Inductive Signature Technologies (IST), the company that developed the signature reading hardware and software, went out of business. This meant that all existing code would no longer be supported—most critically, the proprietary libraries for receiving the signature data from the hardware.
2. The cost of the ORBIX CORBA implementation was difficult to justify for this relatively simple use case of transmitting data over the network for processing at the server
3. Seeking an alternative communications architecture seemed prudent given the strongly typed nature and relative mismatch of CORBA for this lightweight data transmission use-case (not to mention CORBA's decline as a common communications layer)
4. We also wanted a generalized solution that could be used to transmit WIM data as well as signature data.

Given this constellation of issues, we considered a number of solutions and finally settled on a transitional approach that would leverage as much of the existing infrastructure as possible, would provide a path for migration away from burdensome, legacy and commercial technologies, and would also support the addition of WIM data to the communications streams. We discuss the various issues in the following sections.

5.4.2 Supporting Legacy IST Hardware

We began by addressing the issues associated with transmitting signature data. The availability of a substantial inventory of still-functioning IST hardware and simultaneous lack of a readily available alternative on the market meant that it was necessary to try to maintain the IST infrastructure for this project. Though we initially did gain permission to review the source code of IST's server and libraries, we found the legacy code would require substantial modification to recompile and upgrade. Given the limited and unsupported lifetime of the accompanying hardware, we decided to pursue the migration strategy.

To support migration with minimal effort while resolving the primary issues, we decided to make limited modifications to the existing IST-related code but still maintain the CORBA infrastructure for the time-being to minimize the number of simultaneous changes made to the code base. However, in order to transition away from paying license fees for CORBA, we moved to an open-source CORBA implementation called OmniORB, which provided a sufficiently compatible implementation of the necessary CORBA features to make a transition possible

To rebuild the field unit software using OmniORB, we transitioned to using the Cygwin unix emulation layer atop the Windows operating system being used on the field processing units. This transition made the code more easily transferable to alternative platforms in the future by not relying on Windows-specific build tools. Instead, the system is compiled using the cross-platform gcc compiler. The Unix-like cygwin layer also provides capabilities for more rapid and automated remote deployment of system changes. Finally, the source code was placed under version control using the git distributed version control tool.

Once the field unit code was recompiled to work with OmniORB, we performed testing to establish its consistency with the prior toolset. At this point, we evaluated the changes needed in the data being transmitted over the CORBA event service. As noted, the existing software had been set up to transmit signature feature vectors to limit the data bandwidth necessary. Since that implementation, however, wireless communications speeds have improved sufficiently to permit the transmission of complete, raw signature data. To send this data, we modified the interface definition file (IDL) to include raw signature in the structure sent by CORBA and modified the field unit IST processing code to include the raw data in the transmitted objects.

5.4.3 Transitioning the Middleware

With the goal of removing the dependency on CORBA, we sought a lightweight data passing middleware that would provide flexibility on the server-side for using the signature data stream in a variety of ways. These included:

1. The primary use case of storing the raw signature data for later processing
2. The ability to send (near) real-time raw signature data to in-line processing algorithms (e.g., real-time classification)
3. The ability to pass (near) real-time signature data feeds to visualization and debugging tools

Furthermore, we wanted this middleware to be general enough to support data sent from WIM units as well.

We considered a variety of middleware solutions including message brokers such as RabbitMQ, various Enterprise Service Bus architectures, and others. However, we ultimately settled on a very lightweight solution provided by the publish/subscribe (PUBSUB) capabilities of the REDIS key-value store (or data structure server). In the PUBSUB model, a publisher process creates a channel to which one or more subscriber processes can listen to receive data. This system mirrors the CORBA Event Service model already employed by the legacy system, but replaces the fairly heavy burden linking in a library supporting the CORBA infrastructure with a very lightweight and actively supported library capable of communicating with the open source REDIS server.

Working with REDIS offers several advantages for this use case. First, data is passed on the channels as simple strings. CORBA uses strong data typing, whereby interfaces are defined and specified in the IDL files shared between systems. Any change in the data passed (e.g., changing from signature feature vectors to raw signatures) requires that all components receive an updated IDL file and are recompiled to transmit the new data types. This offers distinct advantages for tightly engineered systems. In our signature analysis as WIM data use cases, however, this tight coupling creates as many problems as it solves by requiring early software engineering decisions in the development of a rapidly evolving

system. By moving to a system that passes data as simple strings, we offload the data validation and consistency issue to the subscribing processes.

Though we want to loosen the data specification, we still need to be able to pass structured data in string form. The JSON standard provides a well-known format for representing flexible but structured data in this regard. Furthermore, the JSON format is the lingua franca for web applications, meaning its use for data transition simplifies the development of web applications using that data.

Thus, our design decision was to create a REDIS-based PUBSUB system whereby data from each field unit (IST signatures or WIM station) is assigned a specific channel to which streaming signature data is written, and various processes subscribe to these channels to receive, process, analyze, display, or store the data as appropriate for their functions. In this context, subscribers to the data need to make use of a language-specific library for commenting with the REDIS server using its well-established, open-source protocols for which there are numbers free and open source implementations.

5.4.3.1 Linking IST Units

To bring the legacy CORBA Event Service data used by the IST signature units into the REDIS infrastructure, we created a "bridge" process using the python scripting language. Python was chosen for this process because:

- we preferred to use a scripting language for its simplicity,
- the OmniORB IDL compiler provided python solid support for creating the CORBA communications stubs (but didn't for virtually all other scripting languages),
- there are well supported libraries for converting python structures into JSON strings, and
- there are well supported libraries for interfacing python to REDIS

The python bridge process uses the OmniORB libraries to subscribe to the CORBA Event Channels associated with each field unit. When data comes in on these channels, the bridge process converts the data from the CORBA structures into JSON strings use the standard libraries available. These strings are then published onto the appropriate (matching) REDIS channel

At this point, all downstream processes (subscribers) are shielded from the upstream CORBA middleware and no longer require a CORBA implementation to use the data. Instead, they need to incorporate a much lighter-weight and simpler REDIS implementation as well as language-specific JSON de-serialization library—both of which are readily available for virtually every language across multiple platforms.

5.4.3.2 Linking WIM Stations

Because we were not adapting WIM stations from an existing infrastructure, no bridge process was necessary and linking the WIM station data into the middleware system was more straightforward. In our implementation, software installed on the field unit receives WIM station data in real time, performs some basic processing, converts those data to JSON strings and writes them to the appropriate REDIS channel on the server using standard REDIS libraries.

5.4.4 Using the Data Downstream

With the data now available via the REDIS PUBSUB system, we turned our attention to making use of the data.

5.4.4.1 Signature Data Applications

We built two main downstream applications for the signature data. The first was another "bridge" process that reads the PUBSUB channels and copies the signatures into a database for long-term storage and application development. To build this process we also preferred a scripting language. This time, however, we chose to build an application using server-side javascript and the node.js platform. While we could have continued working with python, we found that javascript's integration with JSON made data processing incredibly simple and minimized the number of lines of code necessary to process the incoming data from REDIS and re-write it to the downstream PostgreSQL database. This process works by subscribing to the specified REDIS channels using the node.js REDIS module, reading the JSON strings published on those channels, converting them (natively) from JSON to javascript objects, and then writing them to PostgreSQL using the well-support node modules for database interaction. Once written to the database, the signatures can be used by downstream applications needing to perform structured queries on the signature datasets.

The second application we built is a web application providing (near real-time) visualization of the signatures coming in from the field units. In this case, we again chose to use node.js for both the reasons stated above, and for its capabilities as a web application platform. In addition to these core technologies, we used a variety of web application technologies to make the interface as dynamic as possible. These include:

- Express: a web development framework that handles routing, data management, and links the remaining technologies to deliver the web application
- Jade: an HTML templating language and processing libraries used to maintain portable HTML5 web pages
- Socket.io: a javascript library for real-time communication between the web server and web clients. Tightly integrated with Express, we used socket.io to deliver streaming signature data to the client web pages in near real-time
- d3.js: a client-side javascript library for transforming data into visualizations using, in our case, scalable vector graphics (SVG) to draw dynamic charts and tables and css transformations to illustrate data streams.

The web application provides two main pages. The first shows all field processing units known to the system, including their names and locations, as well as a timestamp showing when the most recent data was received from the unit as shown in Figure 5-8. Using this single page, a user can quickly assess the health of the system and its various components. The second page is accessed by clicking on any of the field units listed in the first page. This second page shows a graphical representation of the last signatures received from this field unit for each lane (shown in Figure 5-9). This page provides useful diagnostic capabilities, including the ability to visually assess whether the signatures being received are valid.

Site List Sites Config

ITS Signature Data Viewer

Site	DetStaId	Channel	Description	Last data streamed	Last data in db (delayed 30s or more)
ITS_ARB_WIM_001	75	IST1	KEYES		(unknown)
ITS_ARB_WIM_002	46	IST2	GALT		2015-04-01 12:16:29
ITS_ARB_WIM_003	27	TST3	TCY		2015-04-01 12:24:00
ITS_ARB_WIM_004	6006	TST4	I5_SR119		(unknown)
ITS_ARB_WIM_005	113	TST5	CARBONA		2015-04-01 12:28:42
ITS_ARB_WIM_006	6002	TST6	SR99_BARSTOW		(unknown)
ITS_ARB_VDS_001	6003	TST101	SR198_ROEBEN		(unknown)
ITS_ARB_VDS_002	10001	TST102	SR4_TCY		2015-04-01 12:38:00
ITS_ARB_VDS_003	10002	TST103	SR205_MH		2015-04-01 12:41:11
ITS_ARB_VDS_004	10003	TST104	SR152_SR33		(unknown)
ITS_ARB_VDS_005	10004	TST105	SR120_I5		2015-04-01 12:47:23
ITS_ARB_VDS_006	10005	TST106	SR88_WCX		(unknown)
ITS_ARB_VDS_007	10006	TST107	I5_HAMMER		(unknown)
ITS_ARB_VDS_008	6001	TST108	SR168-DAKOTA		(unknown)
ITS_ARB_VDS_009	6004	TST109	SR99_SCHUSTER		(unknown)
ITS_ARB_VDS_010		TST110	None		(unknown)
ITS_ARB_VDS_011	6005	TST111	SR46_SR99		(unknown)

Figure 5-8 Overview of deployed field processing units shown on the ITS Signature Data Viewer

As the web application has been used to monitor the system, additional capabilities were added, such as adding additional timestamp information to show for each field when the most recent signature data arrived in the PostgreSQL database (from the database bridge described above). By comparing to the real-time timestamp, this allows users to assess the stability of the data processing infrastructure.

Site List Sites Config

REID Monitor for TST103 (SR205_MH) 2015-06-19 17:26:56 was the last update; waiting for more...

DetStaId	10002	10002	10002	10002
ISTSN	1031	1077	1064	1022
LaneDir	3	3	4	4
Lane	2	3	2	3
timestamp_full	2015-06-19 17:26:56.279	2015-06-19 17:26:51.263	2015-06-19 17:26:53.967	2015-06-19 17:26:56.450
Samples	302	1034	297	932
rawSigSamps	189	651	210	595
VehicleCount	715	563	463	249
duration	0.2516666650772095	0.8616666793823242	0.2475000023841858	0.7766666412353516
reserved	10846	5497	18388	7563
rawsig				

Figure 5-9 Graphical representation of signatures streaming live from SR-205 at Mountain House Parkway

5.4.4.2 WIM Data Applications

For the WIM data, we built a single "bridge" process similar to the one described above for the signature data. Again, this process simply reads the PUBSUB channels and copies the WIM data into a database for long-term storage and downstream application development. This application is virtually identical to the signature bridge process except that it reads the WIM data and converts it to SQL statements for the WIM tables in the PostgreSQL database. Again, once written to the database, the WIM data can be used by downstream applications needing to perform structured queries on the WIM datasets.

5.5 Data Processing

5.5.1 Field Unit Data Processing

5.5.1.1 WIM Data Preprocessing

WIM data is initially transmitted from the 1060 WIM controller to the field processing unit via a null modem cable. A listener was coded in Java to listen for incoming WIM data transmission events through the connected RS-232serial port. The raw WIM record consists of a data packet, followed by a checksum which uses the Cyclic Redundancy Check 16 (CRC16) algorithm. Hence, each data packet is validated against its checksum for error and flagged if the validation fails. Data packets which pass the validation check are subsequently parsed, timestamped by the field processing unit at the time it is received at the serial port, and reconstituted as a JSON string record before it is transmitted wirelessly to the server. It should be noted that the timestamp from the field processing unit does not reflect the actual detection time of the vehicle. Instead, it records the time when the WIM record arrives at the field processing time from the WIM controller via the null modem interface. There the timestamp would reflect delays that incorporate the duration of the vehicle on the sensor system, the WIM

controller data processing and the data transmission speed between the WIM controller and the field unit. However this is still a very useful measure, since the actual time stamp is only recorded on the WIM controller. The issue arises because the 1060-series WIM controller is a primitive DOS-based system with no ability to automatically synchronize its internal clock with a central server. Hence, many 1060-series based controllers were found to have timestamps that deviated significantly from the actual time. In this regard, the difference between the WIM controller timestamp and the field unit timestamp of WIM records gives a close approximation to the actual deviation of the WIM controller clock. This information is critical for the process of matching inductive signature and WIM records from the same vehicle, which is described later in this chapter.

5.5.1.2 Inductive Signature Preprocessing

Inductive signature data is obtained through the advanced loop detector cards connected via USB cables. An inductance change detection threshold value of 200 is used to determine the presence of vehicles. Hence, each inductive signature record comprises of a continuous stream of data samples with values above the detection threshold, and is timestamped by the field unit at the instance of detection.

5.5.2 Server Data Processing

Inductive signature and WIM records transmitted to the communications and modeling server via the communications architecture described in section 5.4 are initially stored in temporary tables in the PostgreSQL database platform hosted on the database server, and processed hourly. The hourly processes on the server side include WIM and inductive signature data pairing for data obtained from WIM sites, followed by classification and finally data archive. The flowchart of the hourly processes is presented in Figure 5-10.

As a form of redundancy, the system always searches for the earliest available records that are stored in the temporary tables, and begin processing the data in hourly batches until the most recent hour with complete data.

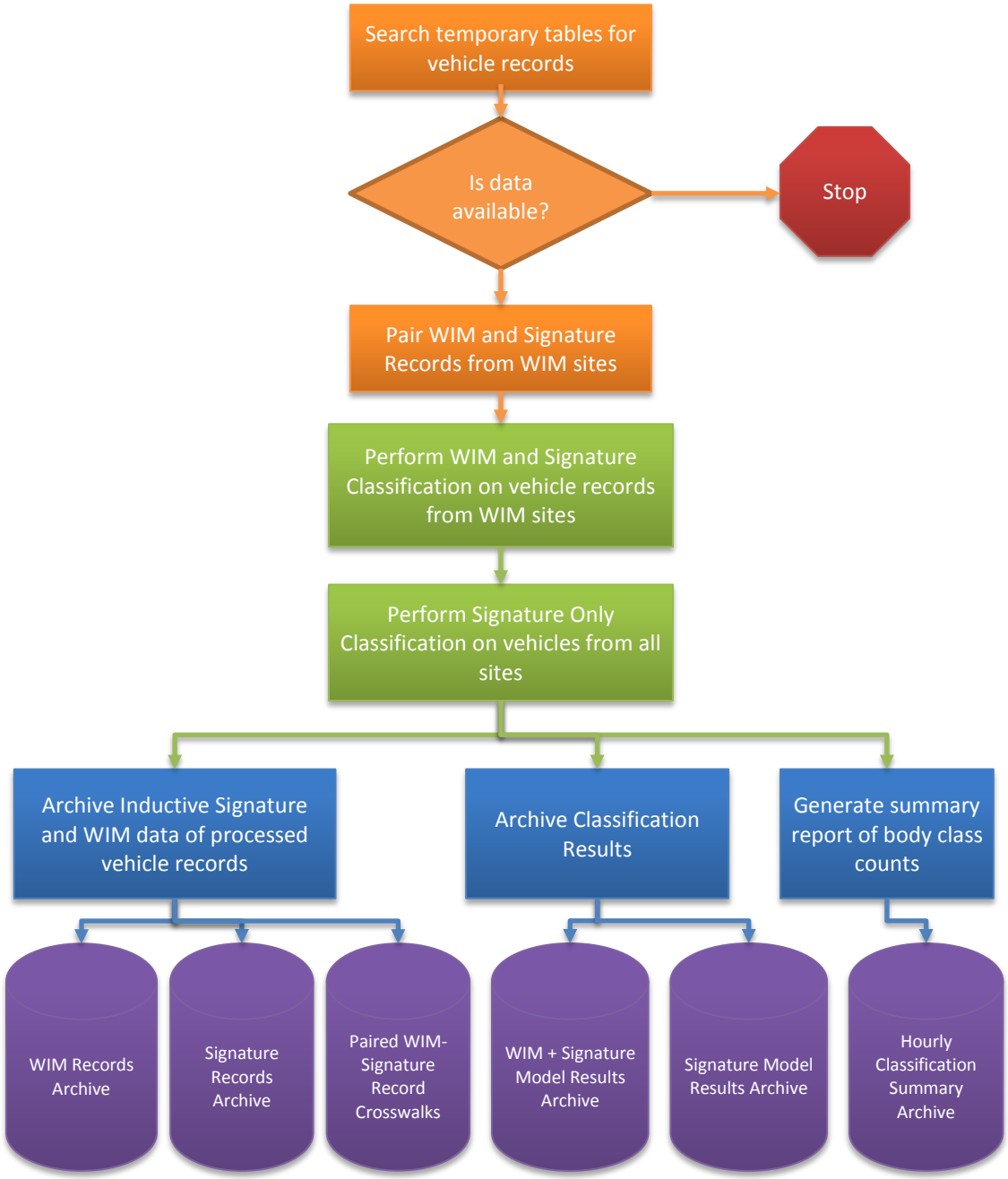


Figure 5-10 Flowchart of hourly data processes

5.5.2.1 WIM and Inductive Signature Record Pairing

As mentioned in Section 3.2, two types of data are obtained from WIM sites: WIM and inductive signature records for every passing vehicle. These two types of data are processed by separate hardware systems: WIM data is obtained from the 1060 WIM controller and signature data is obtained directly from the field processing unit via the advanced ILDs. In this process, the field PC clock sets the detection timestamps for ILD signature, while the 1060 WIM controller clock sets the detection timestamps for WIM. Ideally these two clocks should be synchronized so that the WIM and signature data can be paired based solely on their timestamps. However, the clocks have been observed to drift

and cannot be automatically synchronized since the WIM 1060 uses a DOS-based operating system that does not support synchronization with external clocks. As a consequence, the timestamps reported for the same vehicle may be very different from these systems.

To address this problem, an alignment algorithm was developed to pair WIM and signature data corresponding to the same vehicle. It should be noted that the time differences between WIM and signature, which are referred as time shifts, may drift over time. Therefore, the algorithm was designed to reset the time shifts every one hour to address drifting concerns. The algorithm starts with collecting truck data from WIM and signature based on their durations. Vehicle records with durations longer than 0.35 second were selected for the alignment process. Second, the predefined time shifts, which are ranged from -1000 milliseconds to 3000 milliseconds gapped by 100 milliseconds, were sequentially added to the timestamp of the signature data. At every iteration, WIM data, which has the closest timestamp to the updated signature's timestamp, pair to the signature data. The alignment algorithm repeats the iterations until all the predefined time shifts have been applied to the timestamp of signature data. Additionally, 1 hour aggregated duration differences between WIM and signature data are compared at every iteration, and the time shift which has the smallest aggregated duration difference is chosen as the optimal time shift. Lastly, all other signature data with durations under 0.35 seconds were paired with all WIM data which have the closest timestamp to the optimal time shift.

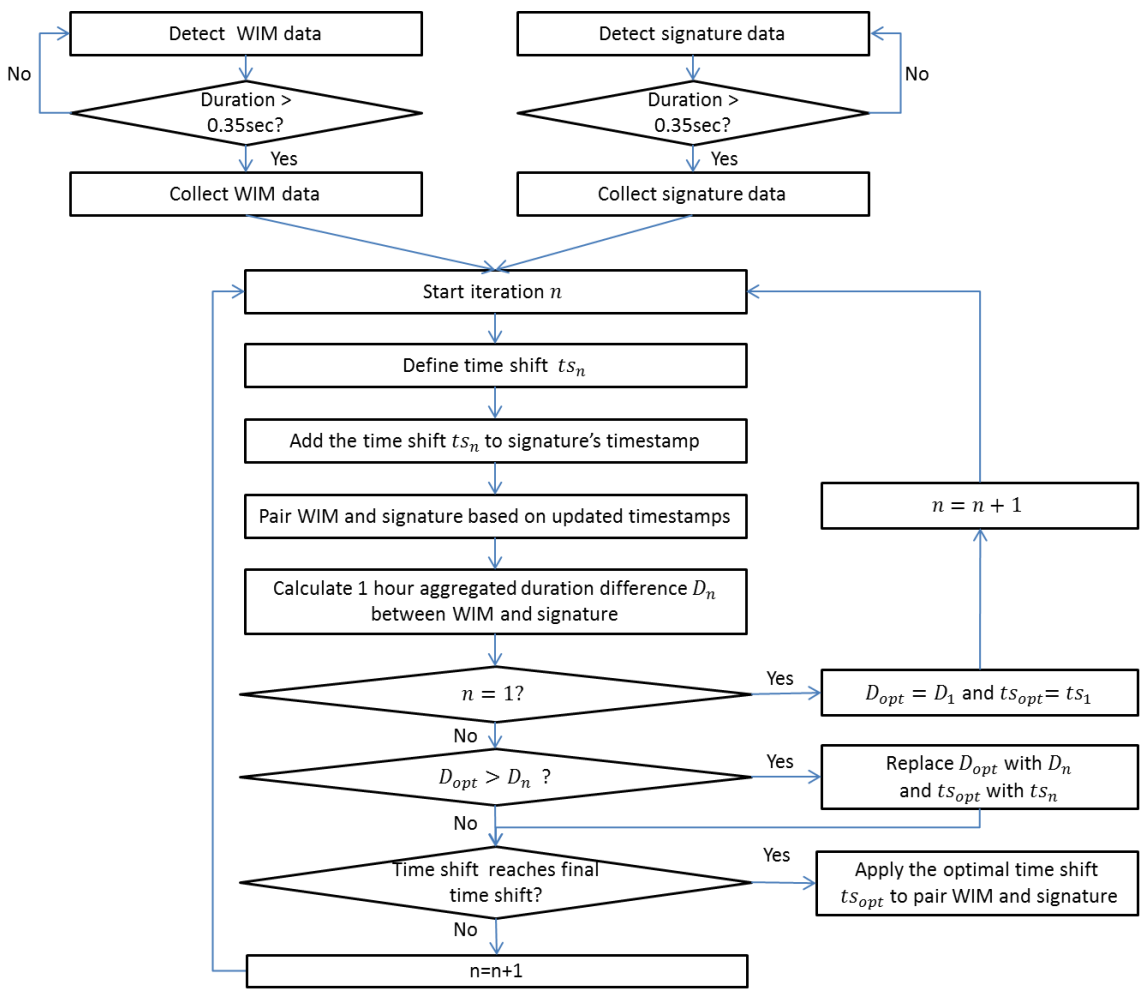


Figure 5-11 Flowchart of the WIM and signature record alignment algorithm

5.5.2.2 Classification and Data Archive

After WIM and signature data have been paired, the vehicle records are classified using the body classification models described in section 4.2. Vehicle records from WIM sites are classified using the 'Inductive Signature Body Classification Model' and the 'Inductive Signature and WIM Body Classification Model' to yield two different vehicle classification predictions. Vehicle records from ILD sites are classified only using the 'Inductive Signature Body Classification Model'. After the classification is completed for all vehicle records, there archival processes are performed. First, the WIM and signature records are transferred to the archive tables and deleted from their corresponding temporary tables the pairing information between matched WIM and signature records are also archived. Next, the results from the classification models are archived. The identifier of each vehicle record is stored with the classification results to allow the classifications to be matched with the raw data records. Lastly, the hourly volume counts by body class are calculated by site and stored archived.

6 System Deployment

6.1 Overview

This chapter describes the efforts that were undertaken to perform the test and actual deployments in the field, as well as the development of the Truck Activity Monitoring System interface to allow users to obtain on-demand truck classification summary reports from the deployed system.

6.2 Test Deployments

The hardware configuration and software for processing WIM and signature data in the field processing unit was first developed in the lab, followed by two test deployments that were performed in the field at existing WIM and ILD sites to test and refine the field system.

6.2.1 Lab Development and Testing

The software for processing WIM and signature data was developed at UCI-ITS. Test inductive loops were fabricated to simulate field inductive loop sensors, while a laptop with RS-232 serial ports was setup to simulate a 1060 WIM controller transmitting WIM data to the field processing unit via a null modem cable.

6.2.2 WIM Test Deployment

The WIM test deployment was performed at the NB Saigon WIM site located along the northbound I-405 freeway in Orange County. This reasons this test site was selected were because of its proximity to ITS-UCI and also because it equipped with a 1060 WIM controller. There were objectives for this test deployment.

Firstly, we wanted to test the custom hardware interface between the WIM controller and the field processing unit. The custom interface described in Section 5.2 was fabricated in-house at UCI-ITS and required field testing to ensure reliability of the design.

Secondly, we wanted to test the stability of the field processing unit configuration. We initially used the Microsoft Windows XP operating system (OS) due to its known compatibility with the software drivers for the IST detector cards. However, we were concerned about the end of support of this operating system by Microsoft, and decided to upgrade to the Microsoft Windows 7 32-bit OS, and were able to resolve the driver compatibility issues with the OS.

Lastly, we wanted to test and refine the software for processing the WIM and inductive signature data, and develop redundancy measures to ensure system stability to minimize the need for on-site maintenance. Wireless modem connectivity issues were occasionally encountered which caused losses in data transmission. To address this, we designed a persistence test that connects to the server hourly to detect if WIM and signature data had been successfully transmitted. The field processing unit was designed to automatically restart several times if it is unable to connect to the database or if it is unable to find any data transmitted from the field processing unit. If the restarts fail to rectify the problem, the system goes into a cold shutdown. A wake-up alarm was configured in each system to initiate a warm

boot at a specific time of day after such shutdown cases, or for system restart after power failure events.

6.2.3 ILD Test Deployment

The ILD test deployment was performed at the N. Sand Canyon ILD site along the I-405 freeway in the City of Irvine. The purpose of this deployment was to test the system's reliability for actual deployment at ILD sites. This deployment was also used to test a new inductive signature detector card developed by CLR Analytics, which is expected to provide inductive signature cards for future deployments in place of IST cards. A prototype algorithm was developed to process the signature data obtained from this detector card for future applications.

Redundancy measures that are similar to the WIM deployment were also implemented on the systems used for ILD deployment.

6.3 Deployment in the California San Joaquin Valley

A total of sixteen locations were deployed with the TAMS in this study. Figure 6-1 presents an overview of the deployed locations in the region of San Joaquin Valley Air Basin. The deployments were chosen from existing WIM and ILD sites located at interstate freeways and state route highways with significant truck volumes.

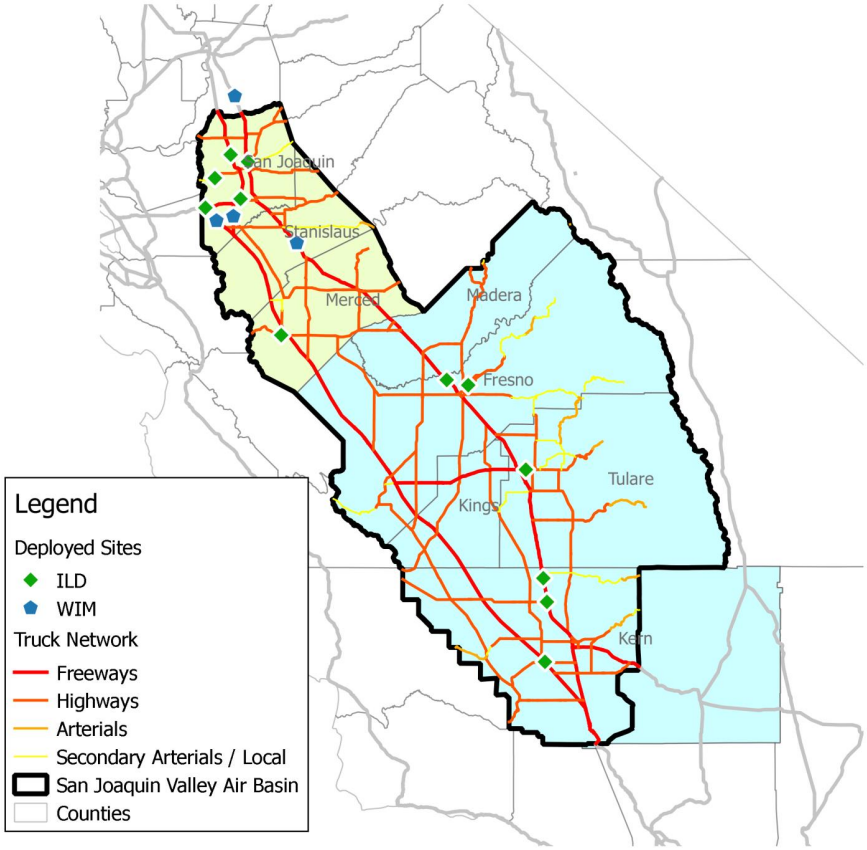


Figure 6-1 Overview of site deployments

6.3.1 WIM Site Deployments

Four WIM sites were selected for deployment for the WIM-Signature classification system as shown in Figure 6-2. These sites were the only ones available configured with 6-foot inductive loop sensors and equipped with 1060 WIM controllers that were compatible with the hardware interface developed by UCI-ITS. Although there were other sites in the California San Joaquin Valley equipped with 1060 WIM controllers, one was configured with 12x6 foot rectangular inductive loop sensors (Lodi), while other sites were either unavailable due to construction activities (Fresno), or were scheduled for upgrade to iSinc controllers (Porterville).

Three of the selected sites are located within the San Joaquin Valley Air Basin (Tracy, Carbona and Keyes). The Tracy and Carbona WIM sites are located on the I-5 and I-580 freeways, respectively. These sites effectively capture the split truck flows to and from the I-5 freeway from the south, with the Carbona WIM site capture flows between the San Francisco Bay Area and the southern half of the California San Joaquin Valley. The Keyes WIM site captures major truck flows on the SR-99 freeway south of Stockton. The fourth WIM site (Galt) is located just outside the San Joaquin Valley Air Basin on SR-99 and was designed to capture truck flows between the San Joaquin Valley Air Basin and the Sacramento metropolitan area.

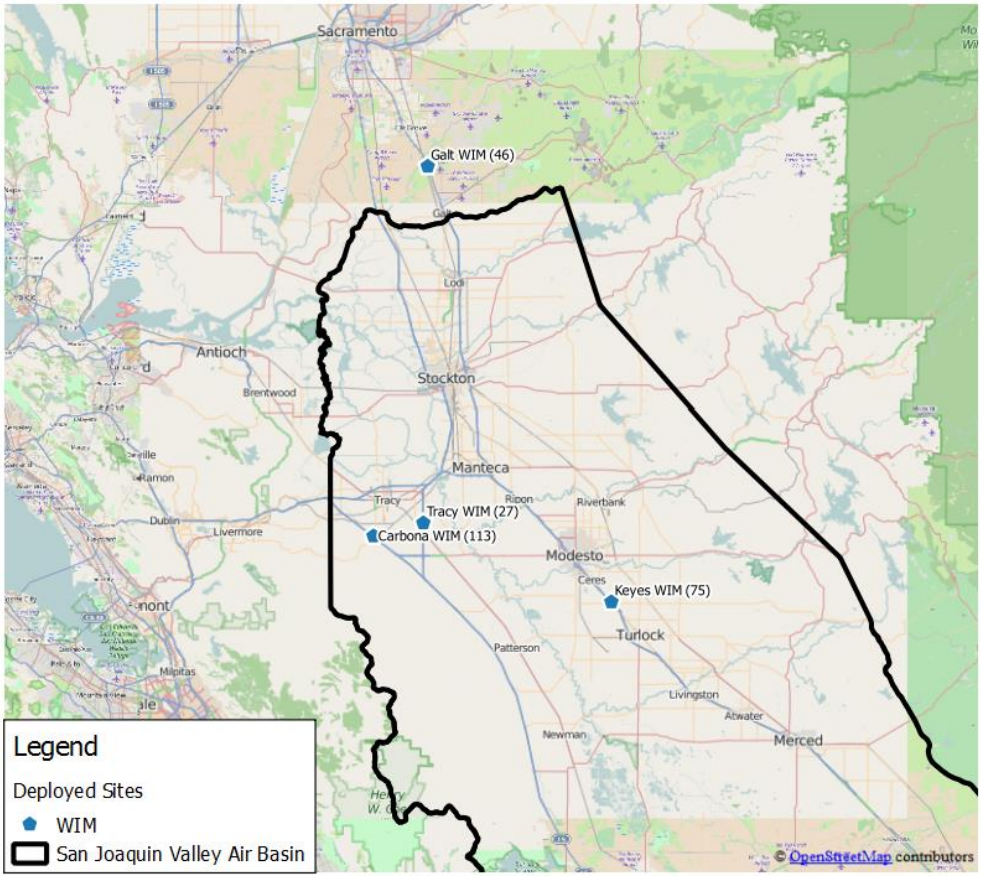


Figure 6-2 Locations of deployed WIM sites

We coordinated with Caltrans Traffic Operations HQ and IRD field technicians to deploy each site. This involved requesting access to the WIM cabinet and configuring the 1060 controller to output WIM data

via the WIM controller's available RS-232 serial port. The IRD technician performed a final verification to ensure functionality of the WIM controller after each deployment was completed.

6.3.2 ILD Site Deployments

Twelve ILD sites were selected for deployment for the signature only classification system. The California San Joaquin Valley Air Basin is divided into two Caltrans Districts: District 6 to the south and District 10 to the north. District 6 encompasses Madera, Fresno, Tulare, Kings and Kern counties, while District 10 incorporates San Joaquin, Stanislaus and Merced counties. Since, deployment at existing ILD sites required coordination at the Caltrans district level, the sites were deployed in two phases, grouped by the Caltrans district they belonged to.

The sites available for deployment in Caltrans District 6 were limited due to recent events of copper theft affecting the traffic detector systems infrastructure. The locations of deployed ILD sites in Caltrans District 6 are presented in detail in Figure 6-3. In the order from north to south SR99_BARSTOW is located on the SR-99 freeway, and was designed to capture truck flows north of Fresno. A location on the SR-180 freeway east of the SR-99 freeway was initially deployed. However, the field unit experience unstable data connectivity issues, and was subsequently re-located to the SR-168 freeway instead. This new location, named SR168_DAKOTA, is located on the SR-168 freeway to capture truck flows serving the northeast region of Fresno. SR198_ROEBEN is located on the SR-198 freeway east of the SR-99 freeway, in the City of Visalia. Further south, SR99_SHUSTER is located on the SR-99 freeway in Delano. This site was designed to capture truck flows between Bakersfield and Fresno. SR46_SR99 is located on the SR-46 highway just west of the SR-99. Since the east end of SR-46 terminates at SR-99, this location captures truck flows to and from the SR-99 along the east-west corridor of SR-46. Lastly, I5_SR119 was designed to capture truck flows on the south end of the San Joaquin Valley Air Basin. It is located on the I-5 freeway just north of the SR-119 interchange.

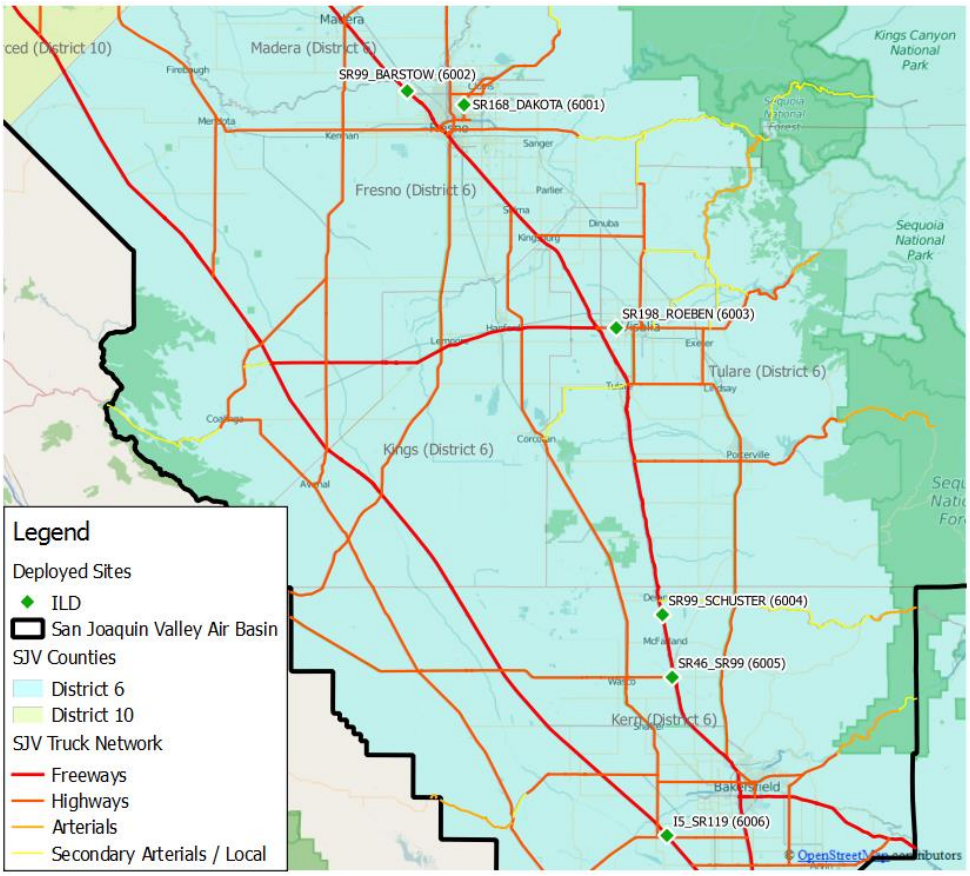


Figure 6-3 Locations of deployed ILD sites in Caltrans District 6

The locations of deployed ILD sites in Caltrans District 10 are presented in detail in Figure 6-4. In the order from north to south, I5_HAMMER is located on the I-5 Freeway in Stockton. This location was designed to capture truck flows entering and leaving the San Joaquin Valley Air Basin to the north via the I-5 truck corridor. SR88_WCX is located on the SR-88 highway, just east of SR-99. This location was intended to capture truck flows serving the northeast of the San Joaquin Valley Air Basin. SR4_TCY is located on the SR-4 highway. This location was designed to capture truck flows towards the Contra Costa county area. SR120_I5 is located on the SR-120 freeway just east of the I-5 freeway. It was designed to capture truck flows on the east-west corridor between the I-5 and SR-99 freeways. SR205_MH is located on the SR-205, and effective captures truck flows between the San Francisco Bay Area and the San Joaquin Valley Air Basin. It is also strategically located between the Port of Oakland and the Union Pacific (UP) Lathrop rail terminal to capture container truck traffic between the two facilities. Lastly, SR152_SR33 is located on the SR-152 highway west of the I-5 freeway. The location was designed to capture truck flows to and from the Monterey County region.

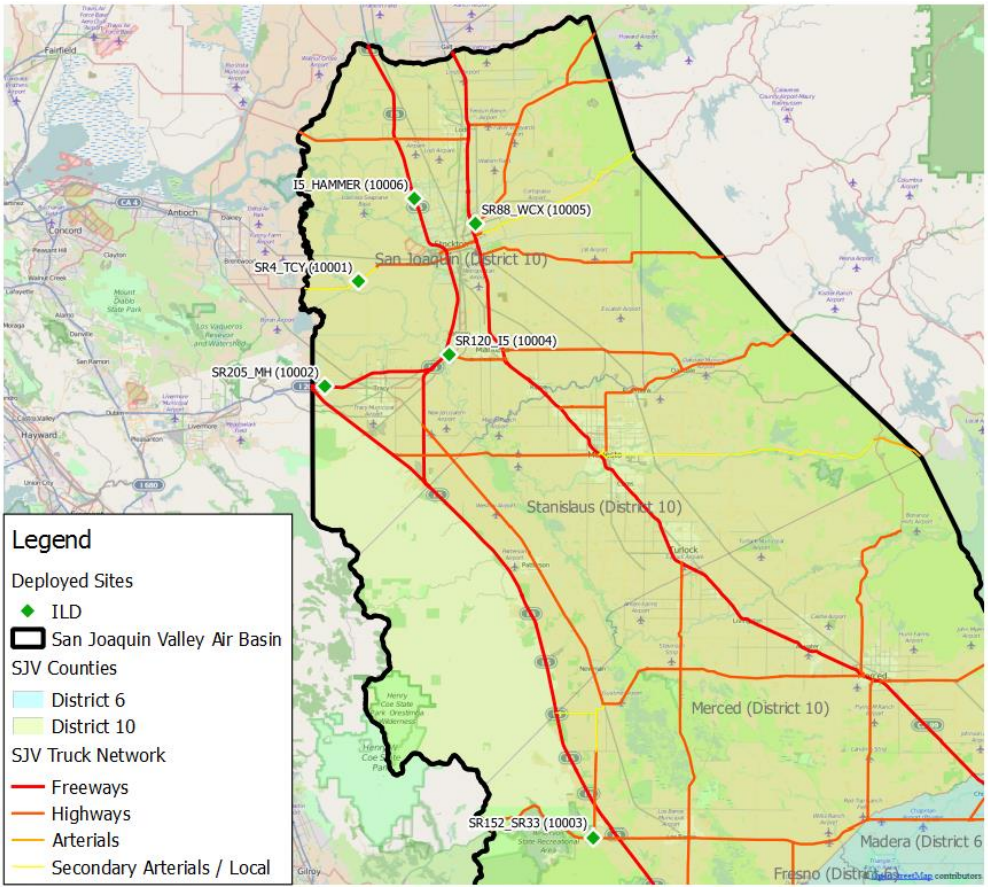


Figure 6-4 Locations of deployed ILD sites in Caltrans District 10

6.4 Truck Activity Monitoring System

The prototype Truck Activity Monitoring System (TAMS, http://freight.its.uci.edu/tams) was developed to provide access to on-demand summary truck classification reports via an interactive web-based user interface.

6.4.1 System Design

TAMS was developed using a combination of Java, JavaScript and Java Server Pages technology. The backend of this system is based on a Java platform that performs data queries with the database server. The results from the data queries are processed through the Java Server Pages middleware which transmits the results to the web browser. JavaScript is then used to process the results and present them on the web interface.

6.4.2 Web Interface Design

The TAMS welcome page presents an overview of the system and the map of deployed site locations for the study as shown in Figure 6-5. A menu bar located on the upper left of the page under the title provides navigation to features currently built into TAMS.

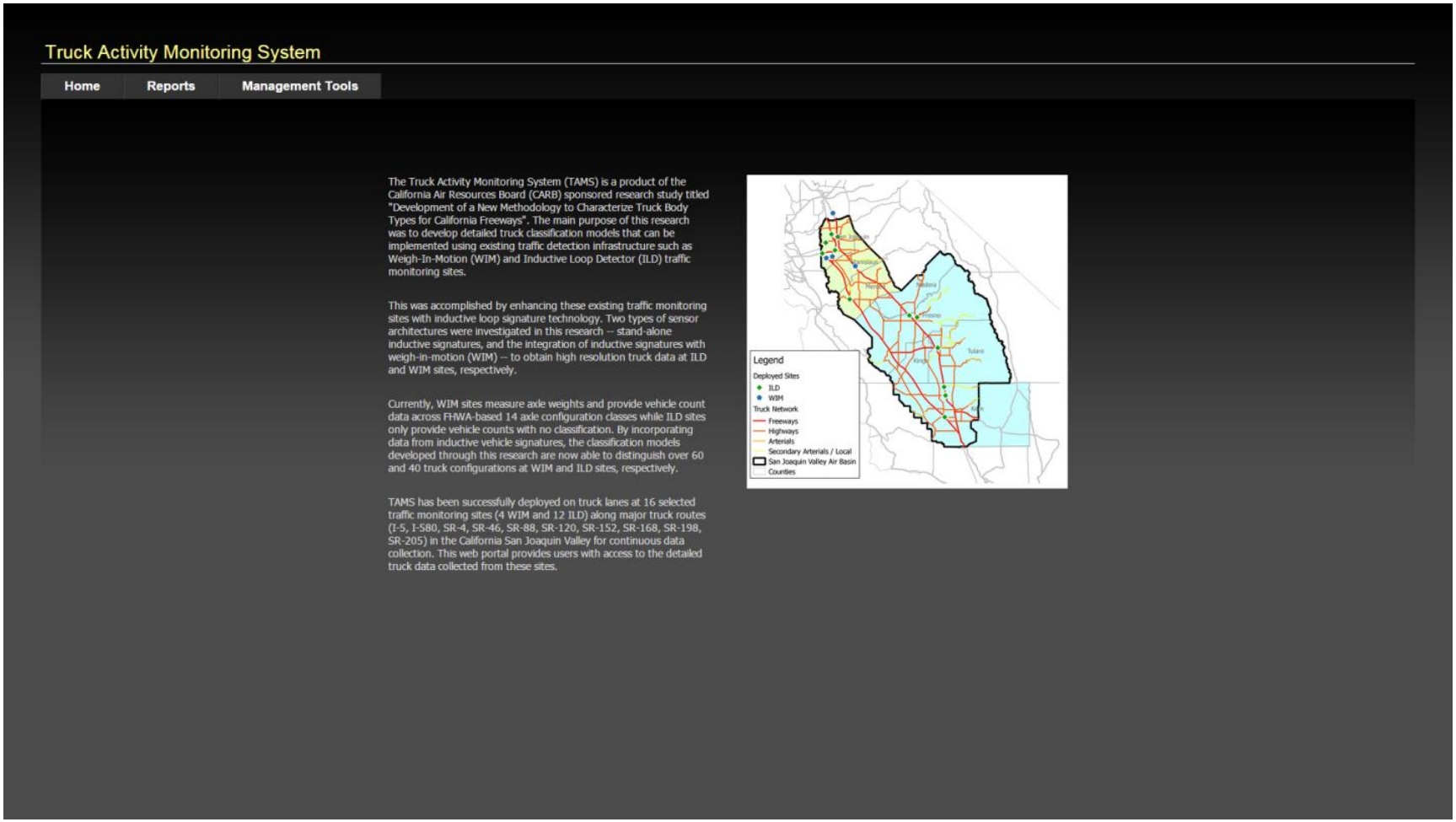


Figure 6-5 TAMS welcome page

To access the main user-interface for obtaining data reports, navigate to Daily Tables under Reports in the menu bar. This will bring up the TAMS Daily Tables interface. The desired site can be selected either via the pull-down box located on the top left corner, or through the interactive map interface. After the date is selected from the calendar interface, an initial summary report will be presented, showing the breakdown of daily volumes by lane and aggregated by vehicle class categories. For WIM sites, the aggregated vehicle classes are presented according to the FHWA scheme as shown in Figure 6-6. In the case of ILD sites, five aggregated vehicle classes are presented (shown in Figure 6-7) as follows:

- PC - passenger cars,
- SU - single unit trucks with no trailers,
- Single - trucks pulling a small single trailer,
- Semi - tractors pulling a semi-trailer, and
- Multi - tractors pulling multi-trailers

A detailed breakdown of hourly volume counts by detailed truck classes can be obtained by clicking on the individual daily volume entries. Examples of detailed hourly classification volumes for WIM and ILD sites are presented in Figure 6-8 and Figure 6-9, respectively.

Galt WIM, SR-99 @ Dillard Rd, WIM site
Summary Data for Tuesday, Jun 23 2015

Vehicle Class	NB ML 1	NB ML 2	SB ML 1	SB ML 2
FHWA 2	8436	3841	7136	3952
FHWA 3	3365	1605	4308	1754
FHWA 4	30	59	24	57
FHWA 5	115	668	220	550
FHWA 6	3	86	7	87
FHWA 7		23		31
FHWA 8	1	111	1	109
FHWA 9	1	1630	44	1776
FHWA 10		11		7
FHWA 11		234		251
FHWA 13		2		1
FHWA 14		38		55
FHWA 15	178	13	136	29

Figure 6-6 Daily volumes by lane and aggregated FHWA classes at WIM site

SR205_MH, SR-205 @ Mountain House Pkwy, ILD site
Summary Data for Tuesday, Jun 16 2015

Vehicle Category	EB ML 2	EB ML 3	WB ML 2	WB ML 3
PC	21130	12453	19801	10074
SU	1461	633	1068	607
Single	208	215	129	144
Semi	719	1617	457	1255
Multi	47	105	19	51

Figure 6-7 Daily volumes by lane and aggregated vehicle classes at ILD site

Galt WIM, SR-99 @ Dillard Rd, WIM site
Summary Data for Tuesday, Jun 23 2015

Click on individual summary volume counts to obtain detailed hourly breakdown by body class

Vehicle Class	NB ML 1	NB ML 2	SB ML 1	SB ML 2
FHWA 2	8436	3841	7136	3952
FHWA 3	3365	1605	4308	1754
FHWA 4	30	59	24	57
FHWA 5	115	668	220	550
FHWA 6	3	86	7	87
FHWA 7		23		31
FHWA 8	1	111	1	109
FHWA 9	1	1630	44	1776
FHWA 10		11		7
FHWA 11		234		251
FHWA 13		2		1
FHWA 14		38		55
FHWA 15	178	13	136	29

Galt WIM, SR-99 @ Dillard Rd: NB ML 2, FHWA Class 9: Breakdown by Hour

Body Class	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Total
20ft Intermodal Container								3							4	1			1					1	10
40ft Intermodal Container	1	2	2	4	8	4	3	4	7	6	4	6	7	10	2	2		1	1	2	2		3	2	83
40ft Intermodal Container Reefer	1	1			1		2	2	2		3	3		2	2	1	3		2				1		26
Agricultural Van	5	1	3	6	3	4	5	6	7	9	5	6	6	5	5		4	4	2	3	1	3		1	94
Automobile Transport		1	1			2				2			1	1		2			1						11
Drop Frame Van	3	4	4	4	6	7	10	4	10	6	1	10	7	12	8	6	5	6	5	6	5	2	2	6	139
Enclosed Van Reefer	20	20	9	25	22	19	22	25	24	27	28	31	36	26	12	13	9	17	16	26	12	13	17	15	484
Enclosed van	5	4	1	3	8	6	3	9	7	9	12	23	14	8	8	2	4	6	7	15	7	8	7	5	181
Livestock	1			1	2	3	6	2	3	5	6	2	1	1	5	5	1		3		1				48
Low Boy Platform	2	3	1	1	3	3	8	15	13	15	10	12	6	12	14	12	11	8	11	4	10	4	5	5	188
Open Top Van/Dump	4	3	4	1	4	8	7	7	11	8	9	4	11	3	9	4	4	2	4	5	3	5	5	3	128
Platform	3	3	2	3	9	11	11	10	9	19	9	11	20	9	17	8	13	7	9	3	9	8	3	8	214
Tank			1	2		3	2	3	1	3	3	2		1		1			1					1	24

Figure 6-8 Detailed hourly classification volumes at WIM site

SR99_SCHUSTER, SR-99 @ Schuster Rd, ILD site
Summary Data for Tuesday, Jun 23 2015

Click on individual summary volume counts to obtain detailed hourly breakdown by body class

Vehicle Category	NB ML 2	NB ML 3	SB ML 2	SB ML 3
PC	7680	6027	11262	6928
SU	4132	449	654	500
Single	73	189	36	94
Semi	748	2058	549	1254
Multi	20	268	51	157

SR99_SCHUSTER, SR-99 @ Schuster Rd: SB ML 3, Tier 2 Class Semi: Breakdown by Hour

Body Class	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Total
20ft Container		1					1		11	7	5	1	1	3	3	2	8	6		3	3	1		1	57
40ft Container								1	2	1	3	1	1	1	4	3	3	3	3	1					27
40ft Container Reefer						1				1		1			1	1				1	2			1	9
53ft Container	1															1									2
Agriculture									1	1				1	1						1				5
Auto						1	1	2	2	3		1	5	3		1		5	2		4		2	1	33
Beverage									1							1									2
Bulk Waste													1												1
Drop Frame Van							2	1	1	5	5		5	1	4	3	1	1	4	1	3		1		38
Dump						1	1	1			1	2	1	1	3		2		1	1					15
Enclosed Van (FHWA 8)	1						1		3	1	2	1	1	1	1	1	2	1	1	1	1				19
Enclosed Van (FHWA 9)	5	5	1		1	9	9	3	10	21	13	14	14	27	21	8	25	33	17	15	24	2	8	2	287
Enclosed Van Reefer (FHWA 8)						1		1		3	3			1											9
Enclosed Van Reefer (FHWA 9)	2	1	1	1	2	10	5	2	11	23	14	22	20	19	36	11	36	36	32	28	21	2	25	4	364
Livestock							2				1														3
Logging						1		2					1	1	1			1			1				8
Low Boy Platform				1		2	3		4	6	3	2	5	5	5	4	1	8	4	2	3	1	1	1	61
Open Top Van					2	3	5	1	11	8	7	4	3	8	7	2	3	4	1	3				1	73
Platform	1	2	1	4	5	6	9	3	14	13	15	10	9	17	20	11	13	14	12	5	3		3	1	191
Tank				2		2	4	3	4	3	2	3	4	1	7	2	4	1	3	1	1	1	2		50

Figure 6-9 Detailed hourly classification volumes at ILD site

6.5 Summary of Findings

Due to technical difficulties associated with the deployment effort, about a month of data was collected for the TAMS before the expiration of the wireless data subscription service. Nonetheless, several insights and observations were obtained from the data collected from the deployed sites.

6.5.1 Truck Travel Pattern Case Study

Due to the limited data available the results presented in this section are mostly from daily observations and are not expected to be statistically significant.

In this analysis, we focused on sites SR205_MH, I5_HAMMER and SR120_I5, which are located on the SR-205, I-5 and SR-120 freeways, respectively. The map of these three locations is presented in Figure 6-10. SR205_MH is located between the Port of Oakland and the City of Lathrop, which has a significant rail facility.

The tabular heat maps in Figure 6-11, Figure 6-12 and Figure 6-13 show the hourly volume of tractors pulling semi-trailers by trailer configuration on June 11, 2015 at their corresponding locations. Colored cells in each table represent the hourly volumes corresponding to the truck body configurations by hour-of-day. The cell color scheme provides an overview of the hourly volume patterns over the course of the day, where red represents the highest hourly truck volumes observed for the day, while green represents the lowest hourly volumes. Cells with no observed volumes are shaded in grey. This facilitates a quick assessment of the predominant truck volumes at the location and the peak volumes of each truck configuration.

Although enclosed van trailers were the most common body configuration, and operate throughout the day, there were also approximately 250 40-foot container trailers (refrigerated and non-refrigerated) traveling along this corridor in both directions. In addition, while tractors pulling enclosed vans were observed to operate throughout the day, the 20- and 40-foot containers were mainly observed during the daylight hours. Whereas 53-foot containers are used only domestically, 20- and 40-foot containers are typically used for at port facilities. Hence, these observed container traffic on the SR-205 corroborate with the operating hours of the Port of Oakland in the San Francisco Bay Area.

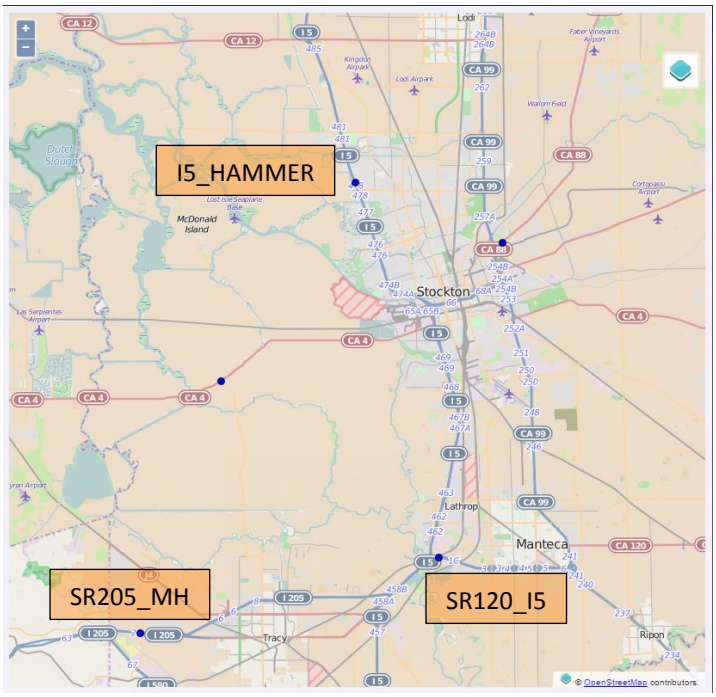


Figure 6-10 Location of detector sites SR205_MH, I5_HAMMER and SR120_I5 for analysis of intermodal freight movement

SR205_MH, SR-205 @ Mountain House Pkwy: EB, Tier 2 Class Semi: Breakdown by Hour of Day

Body Class	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Total
20ft Container			1		2	2	3	7	3	3	7	6	8	7	6	16	12	7	11	7	2		1	2	113
40ft Container	1					3	12	21	7	8	19	11	28	9	8	24	17	23	29	22	3	1	2	1	249
40ft Container Reefer									2	1	1	3	1	2		4	1	1	1	1					18
53ft Container		2			2	1		4	6	2	5	3	9	8	7	5	3	4		3	3		2	1	70
Agriculture						1		5	2		1	1		2	1	1	1	2							17
Auto			1				1	2	2	1		1	3	2	3	2	2	2	1	3	1				27
Beverage						1			1		1	1	1		2			1			1				9
Bulk Waste										1	1	1	1				1						1		6
Container Chassis										1															1
Drop Frame Van				4			1	1	2		2	1	2	2	3	2	3	3	5	2	3	1	1	2	40
Dump						2	2	10	3	3	2	1	3	4	6	2	1		1	1	2	1			44
Enclosed Van (FHWA 8)	1					1	4	2		2	4	4	2	4	6	1	5	5	1	3	2		2		49
Enclosed Van (FHWA 9)	6	4	3	5	7	12	26	50	51	27	46	43	52	67	48	34	42	39	28	34	25	15	24	23	711
Enclosed Van Reefer (FHWA 8)				1				2	3	1		1		2	1	2	1	2	1				1		18
Enclosed Van Reefer (FHWA 9)	3	4	3	7	1	6	17	26	23	14	34	20	27	27	15	30	22	22	7	25	14	13	8	8	376
Livestock												1													1
Logging		1		1	2	1	2		1			2	3	8	4		2	12	4	1	4	4	4	2	58
Low Boy Platform			1	3	1		2	7	6	5	7	6	6	9	8	6	7	4	6	5	2	1	1	1	94
Open Top Van					2	3	9	7	9	7	11	3	6	11	3	9	4	1	10	5	3	1	1	6	111
Platform		2	2	2	4	3	11	21	37	21	29	12	29	34	21	19	16	17	13	5	2	2	1	3	306
Tank	3	3		1	1	2	7	11	15	6	13	6	7	10	8	4	2	7	12	7	9	2	4	4	144

(a) Eastbound SR-205

SR205_MH, SR-205 @ Mountain House Pkwy: WB, Tier 2 Class Semi: Breakdown by Hour of Day

Body Class	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Total
20ft Container	1		1	1		1	2	3	5	3	10	5	5	7	4	3	2	1	3	4	1	1	3		66
40ft Container	2		1	1		2	4	3	8	6	19	15	15	28	16	16	16	16	12	17	10	3	7	3	220
40ft Container Reefer								1	3	1	2	1		4	4	3		2		2	1	1	2	2	29
53ft Container			1	1				4	1	1	3	4	1	3	3	8	3	3	2	2	1	2		2	45
Agriculture						2		1		3	1	2	1			1		1		1			2	1	16
Auto							1	6			2	1					1								11
Beverage						1	1	3	1		1			2							1				10
Bulk Waste					1																				1
Container Chassis																						1			1
Drop Frame Van		1			2		2	4	1	1	1		2	2	1	2	1			3				1	24
Dump	1			1	2	2		1	4	3	3	4	1	6	5	2	3	2	1	3	2	1		1	48
Enclosed Van (FHWA 8)	2			1	1	1	2	1	6	2	7	3	2	3	3	3	2	7	2	1	2	1	4		56
Enclosed Van (FHWA 9)	18	5	3	4	4	1	18	29	30	17	45	21	33	34	39	30	30	24	18	24	26	15	20	23	511
Enclosed Van Reefer (FHWA 8)			1				1				1	1			1	3			3	1					12
Enclosed Van Reefer (FHWA 9)	8	3	3	6	3	2	7	7	11	13	16	5	24	18	16	18	19	27	10	18	13	6	13	9	275
Livestock									1											1					2
Logging				2			3	1	3				1	1	1	1							1		14
Low Boy Platform	2				1	1	3	17	2	6	5	7	11	6	5	6	3	1	1	1		3	2		83
Open Top Van	3		4	3	2	1	1		11	6	6	1	5	14	4	4	3	1	4	6	7	1	2	3	92
Platform	7		3	5	4	5	12	15	20	9	17	13	9	9	18	15	10	6	7	13	3		5	4	209
Tank		4		4	2	1	2	4	9	5	10	8	9	11	8	9	5	7	3	3	4	5	3	2	118

(b) Westbound SR-205

Figure 6-11 TAMS heat map of hourly directional semi-tractor truck volumes along the SR-205 freeway at Mountain House Parkway

In contrast, the volumes of 40-foot containers were less significant on the I-5 freeway north of Lathrop (shown in Figure 6-12) as well as on the SR-120 east of Lathrop (shown in Figure 6-13). The truck volumes in the westbound direction along the SR-120 freeway were not presented as there were suspected data issues due to weak inductive signatures with low magnitudes. Hence, there is strong evidence that Lathrop serves as a significant trip end for port-related container traffic from the San Francisco Bay area.

I5_HAMMER, I-5 @ Valley Forge: NB, Tier 2 Class Semi: Breakdown by Hour of Day

Body Class	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Total
20ft Container		2	1	1	3	1	6	3	2	3	1	4	2	3	3	3	1	4	1	2	1	2			49
40ft Container				1		4	1	4	4	4	3	1	2	2	3	2	1		4	2	2	1			41
40ft Container Reefer			1	3		1	1	1	2	1	2		1	3	2		1	2	1	1		1			24
53ft Container	3	5	6	6	13	20	38	27	14	25	22	20	23	28	12	9	11	12	9	8	9	12	2	7	341
Agriculture	1	1			2	2	3	1	5	3	1		4	1	3		1		4		1		1		34
Auto	1	1		3	4	3	9	6	5	2	4	5	6	5	2	8	3	5	4	2	1	5			84
Beverage						3	1		1			3		1	2	1	2			1	1	1			17
Bulk Waste							1	1	1	4	1	1			1		1			1			2		14
Container Chassis		1							1				1												3
Drop Frame Van	3				3	3	2	9	10	12	8	12	5	4	5	6	6	4	7	4	3	4	2	1	113
Dump		2	2	6	6	3	3	6	7	3	3	4	3	2	4	8	1	1		2	2	3	1		72
Enclosed Van (FHWA 8)			2	3	5	5	6	6	2	2	9	6	5	2	6	6	5	2	5		2		2	1	82
Enclosed Van (FHWA 9)	52	56	47	70	71	110	131	128	150	129	114	120	110	118	105	105	92	88	83	65	65	80	55	44	2188
Enclosed Van Reefer (FHWA 8)							3		1	1	1	1	1	3			1	1							13
Enclosed Van Reefer (FHWA 9)	35	42	54	42	48	61	56	60	72	68	73	65	73	61	61	62	50	36	57	43	44	59	44	41	1307
Livestock								2	2	1		4	3	4	2					2					20
Logging	2					4	3	5	1	3	1	1	1	1		3		3	4	1	1	2			36
Low Boy Platform	3	6	3	5	3	21	18	17	16	22	16	22	19	10	12	22	17	9	9	11	8	4	1	5	279
Open Top Van	6	3	1	4	7	9	12	14	11	14	13	13	10	13	6	8	10	8	2	2	2	2	2	4	176
Platform	7	4	10	14	25	35	44	41	54	49	35	40	44	34	39	40	18	22	28	24	27	13	18	8	673
Tank	3	5	6	7	17	27	25	24	21	30	17	18	24	23	16	17	19	12	9	7	6	6	5	4	348

(a) Northbound

I5_HAMMER, I-5 @ Valley Forge: SB, Tier 2 Class Semi: Breakdown by Hour of Day

Body Class	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Total
20ft Container	1	3		2	1	2	4	3	4	2	1	4	5	2		4	1	4	2	1	1	6	2	1	56
40ft Container	1		1			1		4	2	2	2	1	3	2	3	3	3	2	1				2		33
40ft Container Reefer				1		3	2	1	1		1	2	1		1	1	3	3		3	2	1			26
53ft Container	2	6	3	6	1	3	7	7	11	18	24	25	19	32	38	38	34	25	15	11	6	7	10	3	351
Agriculture					1		2		1		3	2	3			1	2		2	1					18
Auto	1	1				3	2	2	2	4	4	2	1	9	7	10	8	3	5	1	4	2		2	73
Beverage				1					1	1		3		3	1	2	2	2	3					1	20
Bulk Waste			1		1		1	1	1		5	1			2			1	3			1			18
Container Chassis														1	1		1								3
Drop Frame Van			1	1	1	2	2	4	6	4	9	12	6	8	6	6	5	9	5	4	1	4	2	3	101
Dump	1				3	2	2	6	4	3	7	5	10	13	5	5	4	7		3	3		2	2	87
Enclosed Van (FHWA 8)				2	2	2	4	2	2	9	2	4	4	5	4	3	5	3	2	2	1	2			60
Enclosed Van (FHWA 9)	37	38	30	36	61	44	64	108	94	122	147	160	157	161	174	150	142	111	89	89	63	47	63	56	2243
Enclosed Van Reefer (FHWA 8)	1				1			1		1	1													1	6
Enclosed Van Reefer (FHWA 9)	27	27	22	17	20	38	39	58	70	74	86	75	90	86	99	83	76	54	61	63	38	34	27	26	1290
Livestock				1					2	1					2		1	2		1					10
Logging								3	1	1	1		2	4		3		3	3						21
Low Boy Platform	2	2	1	1	2	4	8	11	6	10	11	24	12	19	16	20	18	10	5	9	3	2	1	2	199
Open Top Van	1		1	4		7	2	8	9	9	3	18	14	15	10	9	7	3	2	9	3	1		1	136
Platform	9	6	6	2	12	21	28	30	22	34	41	54	50	39	38	38	39	31	28	42	24	15	17	12	638
Tank		3	1	4	8	10	13	14	25	25	33	19	29	31	24	27	16	21	18	9	6	3	1		340

(b) Southbound

Figure 6-12 TAMS heat map of hourly directional semi-tractor truck volumes along the I-5 freeway at Valley Forge

SR120_I5, SR-120 @ I-5: EB, Tier 2 Class Semi: Breakdown by Hour of Day

Body Class	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Total
20ft Container							5	8	14	18	13	20	12	22	10	13	12	8	9	2	1	2	3	1	173
40ft Container		1		1		5	4	7	6	4	3	4	3	7	4	5	6		7		1				68
40ft Container Reefer						1	6	2	2	5	5	3	6	5	7	3	2	2				1			50
53ft Container							1											1							2
Agriculture							2			1	2	1		3	2	1	1	2				1			16
Auto		1			1		4	5	3	2	1	5	5	2	8	7	2	3	6	2	1			1	59
Beverage							2	3	3	3	4	3	4	2	2	2		1							29
Bulk Waste								1					2		1										4
Container Chassis														1	1		1								3
Drop Frame Van						3	12	10	13	7	19	9	11	9	8	11	6	8	7			1	2		136
Dump							4	2	2	1	2	2	4		1	4				1		1	1		25
Enclosed Van (FHWA 8)						2	6	5	3	4	3	2	5	5	4	4	2		2	1					48
Enclosed Van (FHWA 9)				2	5	14	30	36	29	65	53	64	64	59	58	50	32	34	30	9	15	9	2	5	665
Enclosed Van Reefer (FHWA 8)					2		2			2	2	2	1	2	1	2	3	1		1					21
Enclosed Van Reefer (FHWA 9)		1			1	5	15	20	16	34	21	32	23	34	20	19	17	15	12	4	4	1	2	1	297
Livestock								1	1			1	2	2	1			1							9
Logging						2	7	3	4	6	1	5	4	5	1	4	2	3	4	3					54
Low Boy Platform		1		2		6	13	18	19	19	28	25	20	14	24	22	20	9	18	6	2	3	1	1	271
Open Top Van					1		15	8	12	10	19	14	9	10	8	13	4	7	6	2	1	1			140
Platform	1	1	2		9	17	41	58	68	80	73	82	86	85	78	42	49	31	50	16	6	7	2	10	894
Tank					3	1	4	1	1	5	2	8	6	6	5	2	3	2	1			1		1	52

Figure 6-13 TAMS heat map of hourly eastbound semi-tractor truck volumes along the SR-120 freeway east of the I-5 freeway

6.5.2 Problems Encountered

The reliability of the field hardware systems deployed in this study was excellent, with no significant failures associated with the main hardware components. Most sites remained online throughout the deployment period. Most systems that went offline were able to get re-connected automatically after a day as designed.

6.5.2.1 Detector related issues

Some sites were observed to drop lanes occasionally, which appeared to be a random occurrence. It was determined that the cause may be within the proprietary design of the IST detector cards. However, it appears that this problem may be resolved by monitoring the status of the detector cards, and re-initializing the card if the status of the card is detected to be offline.

It was observed that deployments at WIM sites generally encountered more issues than ILD sites. The field technicians occasionally reported receiving bad data such as misclassifications from the detector cards to the WIM controller. There were also instances when the custom detector interface may have been accidentally disconnected during routine maintenance of the WIM cabinets. This was mostly due to the initial unfamiliarity of the WIM field technicians with the newly added hardware, and is not expected to be a significant issue moving forward. In such cases, the affected LSM adapters were replaced with the original 1060 LSMs to ensure full functionality of the WIM controller.

It was also observed that the signature data from some lanes had abnormally low peak magnitudes at a few sites. As a result, it parts of the inductive signatures were truncated because they did not meet the minimum detection threshold, and signatures from multi-unit trucks were occasionally split into one or more separate signatures consequently recognized erroneously as multiple vehicles. It was subsequently observed that the affected lanes tend to produce higher passenger vehicle volumes and lower overall truck volumes, as shown in Figure 6-14.

I5_SR119, I-5 @ SR-119, ILD site
Summary Data for Tuesday, Jun 23 2015

Vehicle Category	NB ML 1	NB ML 2	SB ML 1	SB ML 2
PC	7487	7389	1727	3207
SU	188	118	1389	1753
Single	18	35	88	68
Semi	206	521	712	2715
Multi	24	51	11	214

Affected lanes

Figure 6-14 Classification count errors due to low signature magnitudes

At other locations signatures were observed to have a high background noise, which affect the quality of the signature data. The consequence of this is a higher misclassification that results in abnormally high volumes for trucks with a large signature variation, such as auto, platform and logging trailers. These errors may be caused by poor connections between the inductive loop sensors and detector cards. Results from the affected site – SR99_SCHUSTER – before and after the connection issues were resolved are presented in Figure 6-15.

SR99_SCHUSTER, SR-99 @ Schuster Rd, ILD site
Summary Data for Wednesday, May 20 2015

Vehicle Category	NB ML 2	NB ML 3	SB ML 2	SB ML 3
PC	9272	4011	9754	4127
SU	1535	5169	1034	598
Single	57	522	89	143
Semi	669	11368	771	3083
Multi	30	195	61	331

SR99_SCHUSTER, SR-99 @ Schuster Rd: NB ML 3, Tier 2 Class Semi: Breakdown by Hour

Body Class	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Total
20ft Container	25	19	12	12	15	21	15	33	30	30	19	26	15	20	20	20	16	13	13	9	14	9	9	10	425
40ft Container		1	1			2	1	1	1		3	1	1	1				1	3		2	1	1	3	24
40ft Container Reefer	5	9	7	4	1	6	6	2	7	3	5	4	4	8	7	8	6	3	1	2	5	2	2	4	111
53ft Container				1												1		1		1					4
Agriculture	4	3	5	3	2	4	5	8	8	5	2	7	4	6	7	5	8	3	1	3	7	1	2	3	106
Auto	241	137	180	120	88	139	126	171	222	134	105	77	69	118	111	102	94	41	37	39	82	61	37	76	2607
Beverage	6	5	5	2	6	6	4	8	11	5	2	8	6	8	9	4	4	5	3	3	3	1	2	2	118
Bulk Waste	1						1							2	1		1	1	1						8
Container Chassis					1	1			1			1	1	1									1		7
Drop Frame Van	56	20	43	16	19	30	28	44	58	42	30	38	25	47	52	43	43	15	24	19	39	32	8	27	798
Dump	3					1	4		1	2	2	1	3	3	1	3	4	1	1	1		3		3	37
Enclosed Van (FHWA 8)	1		1	1	1		2	1	3	4	1	1	1	2	1	4	7		1	1		2	2	1	38
Enclosed Van (FHWA 9)	83	50	45	52	43	48	53	63	63	72	61	42	71	76	58	60	61	39	30	35	45	31	27	31	1239
Enclosed Van Reefer (FHWA 8)	3						2		1			2	1	2				2						1	14
Enclosed Van Reefer (FHWA 9)	7	5	6	6	7	1	15	12	9	11	19	18	18	12	12	22	23	12	11	16	5	9	4	4	264
Livestock	3	1								1	1				1	1	1								9
Logging	13	8	8	18	6	11	13	24	19	28	15	28	22	13	15	19	18	9	10	10	14	7	8	11	347
Low Boy Platform	225	159	225	119	139	178	218	186	308	138	89	119	66	126	113	89	121	39	56	78	140	168	114	232	3445
Open Top Van				3	1	3	1	1	2	3	5	2	2	6	3	2		7	2	1			4	3	51
Platform	64	48	36	46	43	85	78	83	103	74	73	71	77	81	93	80	64	70	58	58	60	48	42	45	1580
Tank	1	5	6	3	4	5	4	12	5	4	6	5	9	10	10	10	6	5	6	8	2	2	4	4	136

(a) Before connection issues were addressed

SR99_SCHUSTER, SR-99 @ Schuster Rd, ILD site
Summary Data for Monday, Jun 22 2015

Vehicle Category	NB ML 2	NB ML 3	SB ML 2	SB ML 3
PC	7999	5341	12207	5966
SU	4724	591	587	819
Single	79	289	40	144
Semi	638	2412	407	1443
Multi	16	303	29	222

SR99_SCHUSTER, SR-99 @ Schuster Rd: NB ML 3, Tier 2 Class Semi: Breakdown by Hour

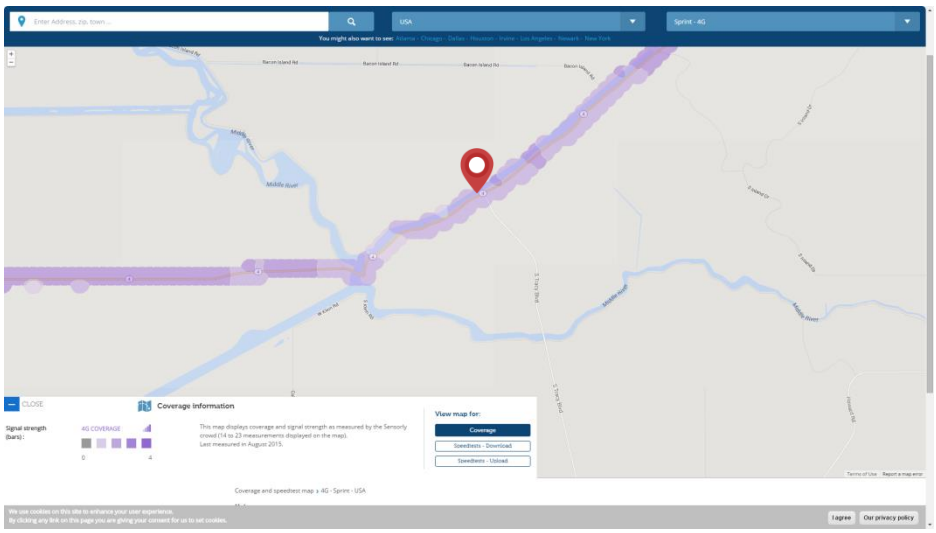
Body Class	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Total
20ft Container					1	2	6	4	2	1	1	1	1	3	2	3	3	2	2		1				35
40ft Container			1			5	6	7	4	5	2	2		3	5	5	6	2	4	1	1	1			60
40ft Container Reefer							2	3	3		2	1	2	1	1	1	1	1	2		1				21
53ft Container					1	1		1	1	1	1			2	2	1	1		2		1		1	1	17
Agriculture			1				1	3		1	1	3	1		1				1						13
Auto					2		3	4	1	1	1	1	3	5		1	3	1	2	1		1	1		31
Beverage								2				1					1								4
Bulk Waste						1	1	1				1	2		2	1		1							10
Container Chassis											1					1						1			3
Drop Frame Van					2		2	1	2	3	1	3	1	1	1	1		3	2	1					24
Dump					2	2	1	1	3		3	3		5		1	1	3	2	1		1	1	1	31
Enclosed Van (FHWA 8)							2	1	2	1		3	2	2			1	2	1		1				18
Enclosed Van (FHWA 9)		2	1	3	19	21	45	40	35	32	24	32	24	33	35	45	56	53	40	22	30	24	12	8	636
Enclosed Van Reefer (FHWA 8)												1	1		1										3
Enclosed Van Reefer (FHWA 9)		1	1	2	13	25	33	39	48	25	31	43	35	46	38	50	65	75	54	28	38	38	17	10	755
Livestock						1	1								4		2	1	1			1		1	12
Logging		1			1		1	2				1				2			1	1		1			11
Low Boy Platform	1				2	3	6	3	17	9	12	8	5	9	7	7	10	10	5	1	2	3	2		122
Open Top Van			2			4	5	5	5	6	8	6	5	3	5	6	4	5	4	1	3	2	1	2	82
Platform			1	2	19	18	25	24	37	24	21	10	24	24	28	17	23	30	23	7	10	15	2	1	385
Tank				1	6	5	8	6	9	6	5	7	8	6	13	14	9	16	7		6	4	2	1	139

(b) After connection issues were addressed

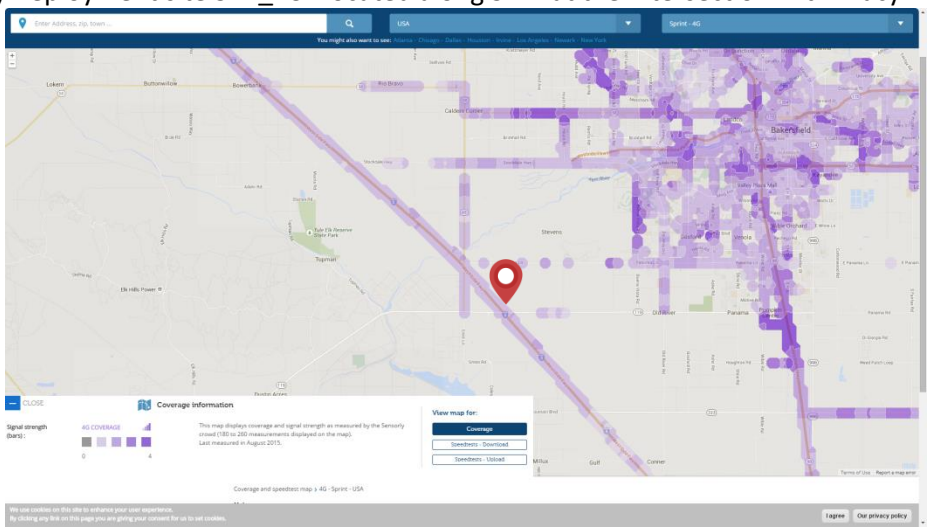
Figure 6-15 Classification count errors due to high signal noise

6.5.2.2 Wireless data communications

Wireless data communications was a concern prior to deployment due to the probability of poor reception coverage at some sites. This issue was observed at two sites: SR4_TCY along SR-4 at the Tracy Blvd intersection and I5_SR119 deployed along the I-5 freeway north of the SR-119 interchange. Verification with cell coverage maps provided by Sensorly[5] (shown in Figure 6-16) confirmed that both these locations had fairly weak cell signals provided by Sprint PCS, which was the wireless service provider for the system.



(a) Deployment site SR4_TCY located along SR-4 at the intersection with Tracy Blvd



(b) Deployment site I5_SR119 located along I-5 freeway north of SR-119 interchange

Figure 6-16 Coverage maps for deployment sites affected by data communications issues (obtained from Sensorly)

[5] www.sensorly.com

6.5.2.3 Other unresolved issues

Apart from sites with known data communications problems, it was observed that SR205_MH located along the SR-205 freeway occasionally went offline due to causes yet to be determined. However, it appears that the system returns online due to the remediating features that were built-in, which forced the system to shut down and perform a cold restart during such situations.

7 Conclusions

This project has demonstrated and applied inductive signature technology to detect truck body class at inductive loop detector (ILD) and weigh-in-motion (WIM) sites in California. Extensive data collection efforts across the state resulted in an exceptional inventory of truck body types that were used to develop two truck body classification models: a standalone inductive signature only model and an integrated WIM and inductive signature model, which were designed for implementation at ILD and WIM sites, respectively. The standalone inductive signature only model is capable of distinguishing over 40 truck configurations, while the model combining WIM and inductive signature data is capable of predicting up to 63 body classes, both with overall accuracies above 70%. In addition to the body class models developed around inductive signature technology, the extensive dataset was leveraged to develop a methodology to predict body class volumes from WIM data for sites not equipped with inductive signature technology which is useful for backcasting tasks related to the validation and calibration of the California Statewide Freight Forecasting Model (CSFFM). Furthermore, the procedure designed to estimate gross vehicle weight distributions at ILD sites showed promising results. The deployment of inductive signature technology and corresponding body classification models to 16 sites in the California San Joaquin Valley gives practitioners and researchers a valuable tool to assess detailed truck activity, freight movements and impacts. The results obtained at selected sites have been shown to corroborate strongly with existing freight facilities in the region.

With new information on truck activity by body types, results from this study are expected to improve heavy duty vehicle classification in the EMFAC model and the California Vehicle Activity Database (CalVAD), and provide critical data that is required for the analysis of freight movement that will benefit the California Statewide Freight Forecasting Model (CSFFM) and other freight- or truck-related studies.

7.1 Implementation Effort and Cost

Site implementation hardware at existing ILD sites primarily comprises a field processing unit, signature capable ILD cards, and wireless modems. The number of ILD cards required varies by site, depending on the loop configuration (single or double inductive loop sensors per lane) and the number of lanes to be monitored. They typically range been two and four for the monitoring of truck traveled lanes. For compatible WIM sites with 1060 series controllers, some additional hardware is required to interface the ILD cards with the controller as well as to receive WIM data records from the WIM controller. This includes a customized ILD adapter and a serial null modem. Hence, depending on the number of ILD cards required, the hardware implementation costs for an existing loop detector site may range from approximately $2,500 and $4,000. The implementation costs at WIM sites are expected to be about $500 higher than ILD sites. In addition, a wireless data communications monthly subscription cost of about $40 is required for data transmission from the field units to the server. This monthly recurring costs may be eliminated if existing high speed communications infrastructure can be utilized.

7.2 Limitations

Some limitations were observed during the development and deployment of the classification systems in this study:

- The traffic sensors used in this study assume that most drivers observe good lane discipline. Vehicles straddling between lanes or changing lanes while over the sensors will generate erroneous data. Hence, it is important to identify locations where there are no significant weaving movements.
- Sensor data quality is typically affected under severe congestion. Although the quality of inductive signature data is not affected by a vehicle's speed, it is sensitive to significant accelerations and decelerations relative to a vehicle's speed because they introduce a skew to the signature waveform. This may affect the accuracy of the classification models developed in this study. Hence, the system performance may be compromised during congested traffic conditions. It should be noted that this weakness is not confined within inductive signature technology; congested conditions are known to affect other conventional traffic measures such as axle spacing measurements at WIM sites and volume count accuracy at conventional ILD sites.
- Inductive signature magnitudes were observed to have a high signal-to-noise ratio for some lanes at certain sites. This may be attributed to the physical quality of the loop installations, as poor splices or deteriorating insulation may cause such problems. There are two concerns with a high signal-to-noise ratio. It may lead to false detections of noise signals as vehicles, and result in misclassification errors due to the noise in signature data being misidentified as vehicle signature features. This problem may be addressed by verifying the condition of inductive loop sensors and leads, and their connections.
- The amplitudes of inductive signatures were found to be low at some sites. As a consequence, trucks pulling trailers with a high ground clearance may be detected as multiple vehicle signatures. These fragmented signatures are typically misclassified as passenger vehicles or single unit trucks. Hence, affected sites will typically undercount multiple unit trucks, and provide an overestimate of passenger car and light truck volumes.

7.3 Follow-up Studies and Future Recommendations

This study has demonstrated the potential to identify truck body configurations at an unprecedented level of detail. As a follow-up effort, Caltrans has sponsored a $1M study to further enhance the classification models and expand the number of deployed sites to over 90 across the State of California. These future deployments will be located along major truck corridors within metropolitan areas, at regional cordon lines, and near state boundaries. This expanded study will also further enhance the TAMS web interface through which users can examine and download individual body class predictions in addition to hourly summaries by location.

With the emergence of alternative drivetrain technologies in trucks, it will be a worthwhile effort to investigate the potential of inductive signature technology in identifying trucks by drivetrain as well. This will be particularly useful in providing truck activity emissions estimates through monitoring truck activity by drive train technologies and provide information on the best policies to maximize the benefit of technology investments.

8 References

ASTM Standard E1318-02, ASTM Standard Specification for Highway Weigh-in- Motion (WIM) Systems with User Requirements and Test Methods. ASTM International, West Conshohocken, PA, 2009, DOI: 10.1520/E1318-09, 2009, www.astm.org

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A., Classification and regression trees. CRC press, 1984.

Brown, G., Wyatt, J., Harris, R., and Yao, X., Diversity Creation Methods: A Survey and Categorization, Journal of Information Fusion Special issue on Diversity in Multiple Classifier Systems, Volume 6, Issue 1, March 2005, pp 5-20.

Bureau of Transportation Statistics, Freight in America, January 2006, http://www.bts.gov/publications/freight_in_america/, accessed online June 2012

California Air Resources Board, Mobile Source Emissions Inventory, Appendix G, Emissions Analysis Methodology and Results: Medium-Heavy Duty Trucks, Heavy-Heavy Duty Diesel Trucks, and Regulated Bus Categories, http://www.arb.ca.gov/regact/2010/truckbus10/truckbusappg.pdf, accessed June 2012.

Caltrans Truck Services, Data Weigh-in-Motion Webpage, http://www.dot.ca.gov/hq/traffops/trucks/datawim/, accessed online June 2012.

Cetin, M., Monsere, C.M., Nichols, A.P., Ustun, I., Key Factors Affecting the Accuracy of Reidentification of Trucks over Long Distances Based on Axle Measurement Data. Transportation Research Record: Journal of the Transportation Research Board 2243, 2011, pp. 1–8.

Chang, L. Y., & Wang, H. W. Analysis of traffic injury severity: An application of non-parametric classification tree techniques. Accident Analysis & Prevention, 38(5), 2006, pp. 1019-1027.

Dietterich, T. G., Ensemble Methods in Machine Learning, Multiple Classifier Systems Lecture Notes in Computer Science, Volume 1857, 2000, pp. 1-15.

Federal Highway Administration Washington, D.C., Truck Characteristics Analysis, 1999, https://www.fhwa.dot.gov/ohim/tvtw/trchanal.pdf.

Federal Highway Administration, U.S. Department of Transportation, 2001., Traffic Monitoring Guide. FHWA-PL-01-021, http://www.fhwa.dot.gov/ohim/tmguide/index.htm.

Gentili, M., Mirchandani, P.B., Locating Active Sensors on Traffic Networks. Ann Oper Res 136, 2005, pp. 229–257.

Gentili, M., Mirchandani, P.B., Locating sensors on traffic networks: Models, challenges and research opportunities. Transportation Research Part C: Emerging Technologies 24, 2012, pp. 227–255.

Hastie, T., Tibshirani, R., Friedman, J., Hastie, T., Friedman, J., & Tibshirani, R. The elements of statistical learning, New York: Springer, Vol. 2, No. 1, 2009.

Hernandez, S., Tok, Y.C.A., Ritchie, S.G. Density Estimation using Inductive Loop Signature based Vehicle Re-identification and Classification, UCI ITS Working Paper, UCI-ITS-WP-13-4, 2013.

Highway Statistics 2010, Office of Highway Policy Information, Annual Vehicle Distance Traveled in Miles and Related Data, Table VM-1, http://www.fhwa.dot.gov/policyinformation/statistics/2010/vm1.cfm, accessed June 2012.

Inductive Signature Technologies, Advanced Inductive Detector Specification Sheet, Accessed online from http://www.ist-traffic.com/datasheets/ist222ss.pdf , August 30, 2012.

Institute of Transportation Studies (ITS), University of California, Irvine, Assessment and Development of Commodity Flow, Logistics, and Other relevant Goods Movement Data Sources to Facilitate Statewide Freight Modeling, Final Report for the California Department of Transportation, February 2010.

Jeng, S.T., Tok, Y.C.A., Ritchie, S.G.,Freeway corridor performance measurement based on vehicle reidentification,Intelligent Transportation Systems, IEEE Transactions on,Vol. 11 (3), pp 639-646, 2010.

Jeng, S.T., Chu, L., and Cetin, M., Weigh In Motion Station Monitoring and Calibration using Inductive Loop Signature Technology, presented at the 94th Annual Meeting of the Transportation Research Board, Washington, D.C., January 11-15, 2015.

Kuncheva, L., Combining Patter Classifiers, John Wiley and Sons, New York, 2004.

Kwigizile, V., Mussa, R., & Selekwa, M., Connectionist Approach to Improving Highway Vehicle Classification Schemes: The Florida Case. Transportation Research Record: Journal of the Transportation Research Board, (1917), 2005, 182-189.

Lam, W.H.K., Lo, H.P., Accuracy of O-D Estimates from Traffic Counts. Traffic Engineering & Control, Vol. 31. 1990.

Lee, C. E., & Souny-Slitine, N. Final research findings on traffic-load forecasting using weigh-in-motion data. Work, 987, 7., 1998

Li, X., Ouyang, Y., Reliable sensor deployment for network traffic surveillance. Transportation Research Part B: Methodological 45, 2011, pp.218–231.

Liu, H., Tok, Y.C.A., and Ritchie, S.G., Development of a Real-Time On-Road Emissions Estimation 46 and Monitoring System, IEEE ITSC 2011 Conference Paper, October 5-7, 2011, Washington, D.C.

Liu, H.X., Danczyk, A., Optimal Sensor Locations for Freeway Bottleneck Identification. Computer-Aided Civil and Infrastructure Engineering 24, 2009, pp. 535–550.

Lu, Q., Harvey, J., Le, T., Lea, J., Quinley, R., Redo, D., and Avis, J., Truck Traffic Analysis using Weigh-in-Motion (WIM) Data in California, University of California, Berkeley, Institute of Transportation Studies, Pavement Research Center, June 2002.

NCHRP, 2008, National Cooperative Highway Research Program (NCHRP) Synthesis 386, http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_syn_386.pdf

NCFRP, 2014, National Cooperative Freight Research Program (NCFRP) Repot 39

Nichols, A., and D. M. Bullock, Quality Control Procedures for Weigh-in-Motion Data. Publication FHWA/IN/JTRP-2004/12. Joint Transportation Research Program, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana, 2004.

Papagiannakis, A. T., Quinley, R., & Brandt, S. R. NCHRP Synthesis project 386: High Speed Weigh-in-Motion System Calibration Practice. Transportation Research Board of National Academies, Washington DC., 2008.

Pursula, M. and Pikkarainen, P., Neural Network Approach to Vehicle Classification with Double Inductive Loops, Proceedings of the 17th ARRB Conference, Part 4, 1994, pp. 29-44.

Prozzi, J.A. and Hong, F., Effect of Weigh-in-Motion System Measurement Errors on Load-Pavement Impact Estimation, Journal of Transportation Engineering, Vol. 133, No. 1, 2007, pp 1-10.

Quinley, R., WIM Data Analyst's Manual, Publication FHWA-IF-10-018, FHWA, U.S. Department of Transportation, 2010.

Sherali, H.D., Desai, J., Rakha, H., A discrete optimization approach for locating Automatic Vehicle Identification readers for the provision of roadway travel times. Transportation Research Part B: Methodological 40, 2006, pp.857–871.

The California Vehicle Activity Database (CalVAD) Viewer, California Traffic Management Laboratories (CTMLabs), http://calvad.ctmlabs.net/ Accessed September 10, 2013.

Titterington, D.M., Murray, G.D., Murray, L.S., Spiegelhalter, D.J., Skene, A.M., Habbema, J. D. F., 25 and Gelpke, G. J., Comparison of discriminant techniques applied to a complex data set of head 26 injured patients, Journal of the Royal Statistical Society, Series A, Vol. 144, Issue 2, 1981, pp. 27 145–175.

Tok, A., Hernandez, S., Ritchie, S.G., Accurate individual vehicle speeds from single inductive loop signatures, presented at the 88th Annual Meeting of the Transportation Research Board, Washington, D.C., 2009.

Vehicle Inventory and Use Survey (Discontinued), U.S. Census Bureau, 2002 Economic Census.

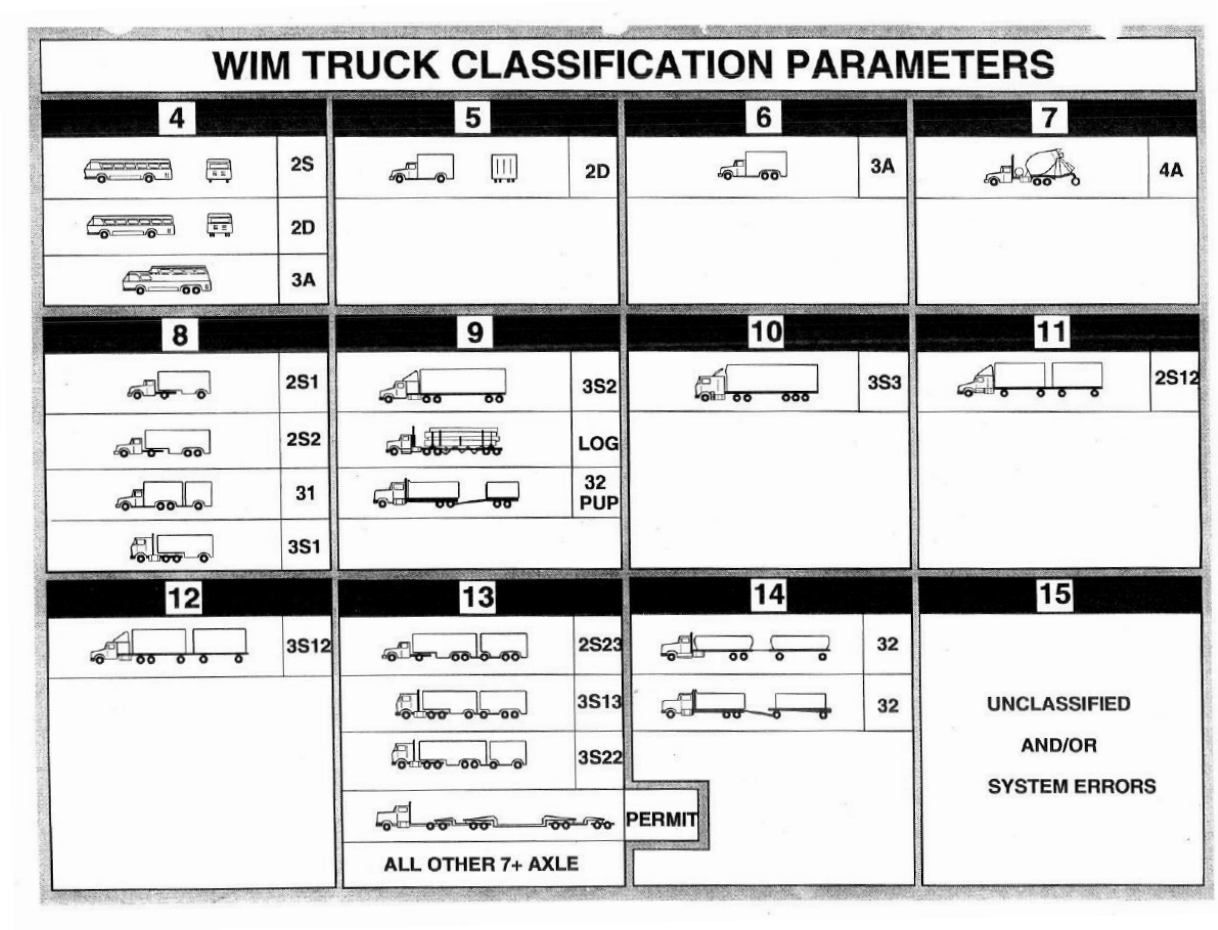
Yan, X., & Radwan, E. Analyses of rear-end crashes based on classification tree models. Traffic injury prevention, 7(3), 2006, pp. 276-282.

Yang, H., Zhou, J., Optimal traffic counting locations for origin–destination matrix estimation. Transportation Research Part B: Methodological 32, 1998, pp.109–126.

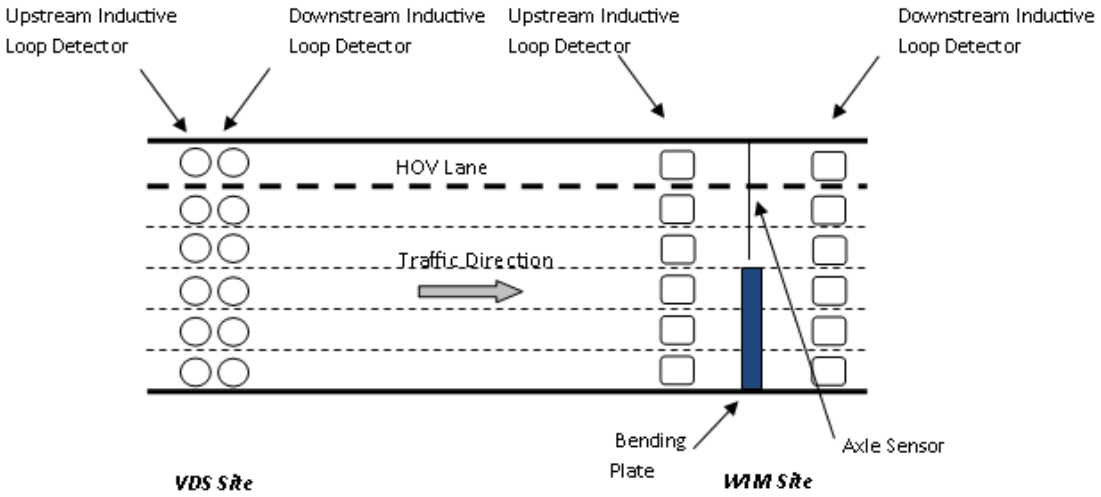
Appendix A – FHWA-CA Classification Scheme

FHWA-CA Classification Scheme for Commercial Vehicle Classes 4 through 14

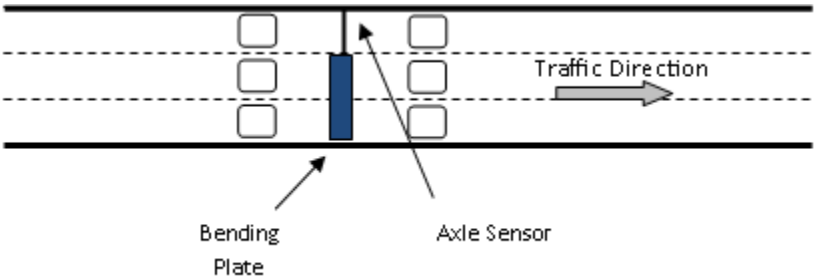
(http://www.dot.ca.gov/hq/tpp/offices/ogm/trucks/WIM_Truck_Classification_Diagram.pdf)



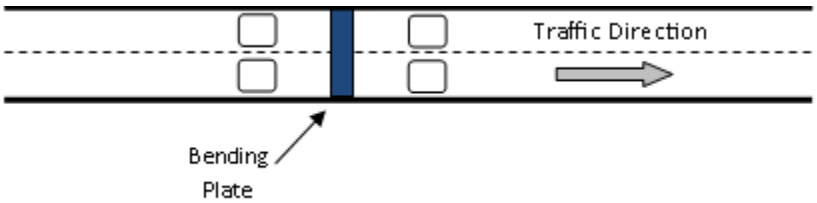
Appendix B – Data Collection Site Configurations



Irvine Site Configuration



Fresno Site Configuration



Willows and Redding Site Configuration

Appendix C – Database Structure for Data Groundtruth System

Vehicle Records with user input body classification

vehid	station	lane	time	truck_axle	trailer_axle	total_axle	truck_body	trailer_body	unsure	commodity
821355176810120	2	2	1355176810120	1	99	2	5	99	0	-99
821355176831500	2	2	1355176831500	3	7	5	35	26	0	-99
821355176868890	2	2	1355176868890	2	99	2	6	99	0	-99
821355176884520	2	2	1355176884520	1	99	2	5	99	0	-99
821355176918740	2	2	1355176918740	1	99	2	5	99	0	-99
821355176928090	2	2	1355176928090	3	4	5	52	110	0	-99
821355176933710	2	2	1355176933710	1	2	4	5	32	0	-1
821355176942580	2	2	1355176942580	2	11	5	51	1	0	1
821355176944480	2	2	1355176944480	3	4	5	52	11	0	-99
821355176948650	2	2	1355176948650	3	7	5	12	1	0	1
821355176964310	2	2	1355176964310	3	4	5	52	110	0	-99
821355176973490	2	2	1355176973490	2	99	2	44	99	0	-1
821355176991300	2	2	1355176991300	1	1	3	7	51	0	-1
821355176996300	2	2	1355176996300	3	11	6	52	21	0	-99
821355177042900	2	2	1355177042900	2	11	5	51	11	0	-99
821355177046760	2	2	1355177046760	2	2	4	6	32	0	-1

Signature Records

vehid	wimsigid
821355176810120	831355176811267
821355176831500	831355176832649
821355176868890	831355176870046
821355176884520	831355176885740
821355176918740	831355176919901
821355176928090	831355176929260
821355176933710	831355176934880
821355176942580	831355176943801
821355176944480	831355176945701
821355176948650	831355176949873
821355176964310	831355176965531
821355176973490	831355176974659

WIM Records

vehid	wimid
821355176831500	1904399
821355176928090	1904400
821355176942580	1904403
821355176944480	1904404
821355176948650	1904405
821355176964310	1904406
821355176973490	1904407
821355176991300	1904408
821355176996300	1904409
821355177042900	1904410
821355177046760	1904411

Photo Records

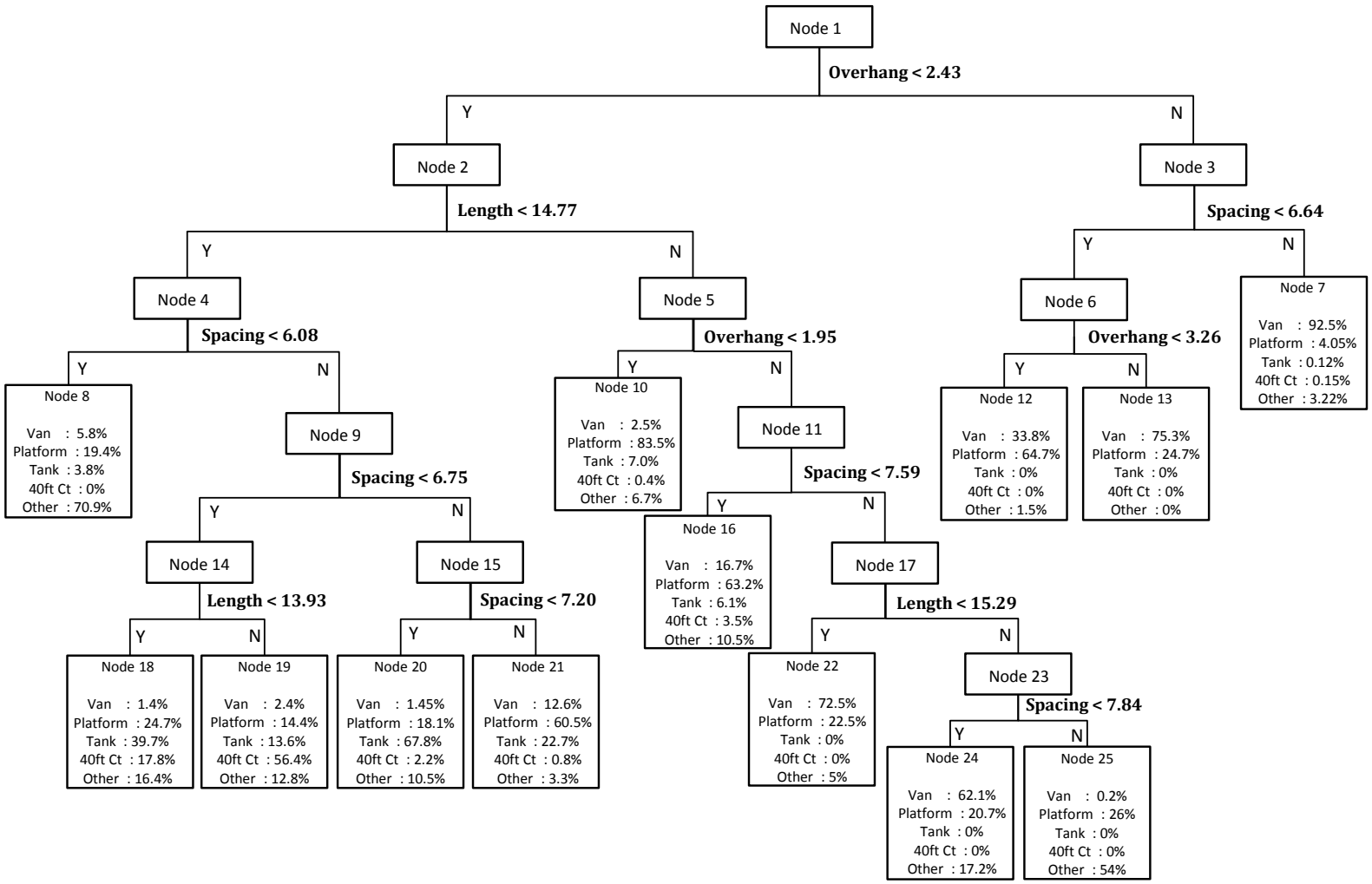
vehid	photoid	photoorder
821355176810120	821355176810120	1
821355176831500	821355176831500	1
821355176831500	821355176831800	2
821355176831500	821355176832140	3
821355176868890	821355176868890	1
821355176884520	821355176884520	1
821355176918740	821355176918740	1
821355176928090	821355176928090	1
821355176928090	821355176928400	2
821355176928090	821355176928730	3
821355176933710	821355176933710	1
821355176933710	821355176934020	2

Combine Signature, WIM, and Photo data by Vehicle ID field

Signature, WIM, and Photo Records linked by Vehicle Record

vehid	wimsigid	wimid	photoid
821355176831500	831355176832649	1904399	821355176831500
821355176831500	831355176832649	1904399	821355176831800
821355176831500	831355176832649	1904399	821355176832140
821355176928090	831355176929260	1904400	821355176928090
821355176928090	831355176929260	1904400	821355176928400
821355176928090	831355176929260	1904400	821355176928730
821355176942580	831355176943801	1904403	821355176942580
821355176942580	831355176943801	1904403	821355176942880
821355176942580	831355176943801	1904403	821355176943220
821355176944480	831355176945701	1904404	821355176944480
821355176944480	831355176945701	1904404	821355176944790
821355176944480	831355176945701	1904404	821355176945120
821355176944480	831355176945701	1904404	821355176945450

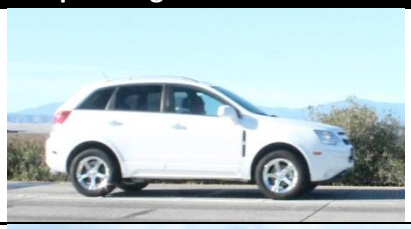
Appendix D – Modified Decision Tree Model Architecture for Predicting Truck Volume by Body Configuration using WIM data

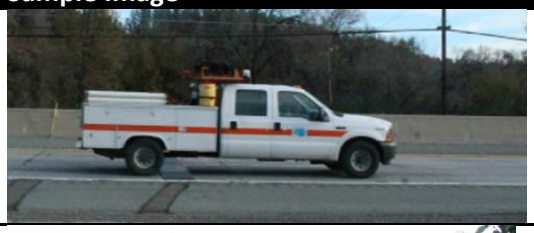


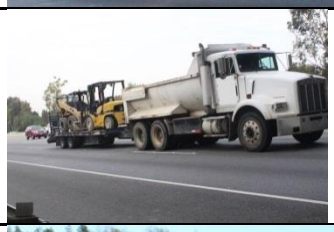
Appendix E – Pin-out Cross Assignment between 1060 WIM and IST-222 Detector Cards

1060 Pin	IST Card No./Pin	1060 Pin	IST Card No./Pin
1		33	
2		34	#1/H
3	#1/F	35	
4		36	
5	#1/D	37	
6		38	#1/E
7	#1/7, #1/20, #2/7, #2/20	39	
8		40	
9	#1/X	41	
10		42	#1/W
11	#1/A, #2/A	43	
12		44	#1/J
13	#1/K	45	
14		46	#1/L, #2/L
15		47	
16		48	#2/H
17	#2/F	49	
18		50	
19	#2/D	51	
20		52	#2/E
21		53	
22		54	
23	#2/X	55	
24		56	#2/W
25		57	
26		58	#2/J
27	#2/K	59	
28		60	
29	#1/C, #2/C	61	
30	#1/B	62	#2/B
31		63	
32		64	

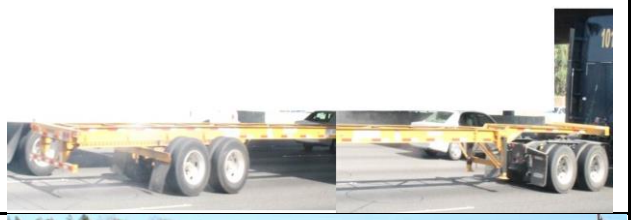
Appendix F – Vehicle Classification Scheme for Inductive Loop Signature Only Classification Model

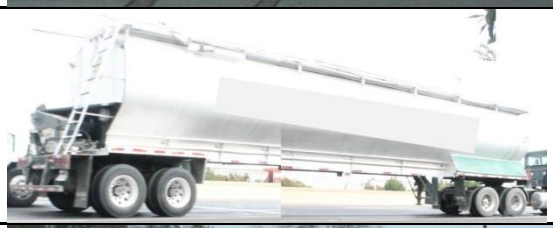
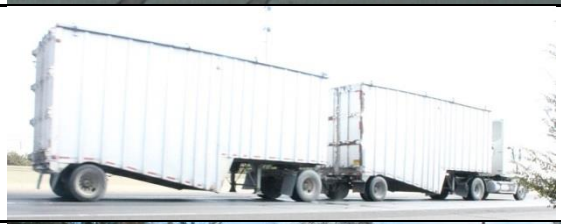
Vehicle Category	Body Configuration	Sample Image
Passenger Car	Passenger Vehicle	
Single Unit Truck	20ft Bus	
Single Unit Truck	30ft Bus	
Single Unit Truck	Beverage	
Single Unit Truck	Bobtail	
Single Unit Truck	Van/Platform	
Single Unit Truck	Concrete	

Vehicle Category	Body Configuration	Sample Image
Single Unit Truck	Utility/Service	
Single Unit Truck	Dump/Tank	
Single Unit Truck	Dump Triple Rear	
Single Unit Truck	Multi Stop Van/RV	
Single Unit Truck	Street Sweeper	
Vehicle with Single Trailer	Passenger Vehicle w/Trailer	
Vehicle with Single Trailer	Pickup/Utility w/Trailer	
Vehicle with Single Trailer	Platform w/Trailer	

Vehicle Category	Body Configuration	Sample Image
Vehicle with Single Trailer	Concrete w/Lift Axle	
Vehicle with Single Trailer	Dump	
Vehicle with Single Trailer	Dump w/Lift Axle	
Vehicle with Single Trailer	Dump w/Trailer	
Vehicle with Single Trailer	Van w/Trailer	
Vehicle with Single Trailer	Tank w/Trailer	
Vehicle with Single Trailer	RV w/Towed Vehicle	
Vehicle with Single Trailer	Service w/Trailer	

Vehicle Category	Body Configuration	Sample Image
Vehicle with Single Trailer	RV w/Small Trailer	
Vehicle with Single Trailer	Livestock w/Trailer	
Tractor with Semi Trailer	20ft Container	
Tractor with Semi Trailer	40ft Container	
Tractor with Semi Trailer	40ft Container Reefer	
Tractor with Semi Trailer	53ft Container	
Tractor with Semi Trailer	Agriculture	
Tractor with Semi Trailer	Auto	

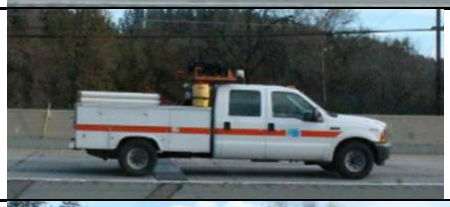
Vehicle Category	Body Configuration	Sample Image
Tractor with Semi Trailer	Platform	
Tractor with Semi Trailer	Beverage	
Tractor with Semi Trailer	Bulk Waste	
Tractor with Semi Trailer	Container Chassis	
Tractor with Semi Trailer	Drop Frame Van	
Tractor with Semi Trailer	Enclosed Van Reefer (FHWA 9)	
Tractor with Semi Trailer	Enclosed Van Reefer (FHWA 8)	
Tractor with Semi Trailer	Enclosed Van (FHWA 9)	
Tractor with Semi Trailer	Enclosed Van (FHWA 8)	

Vehicle Category	Body Configuration	Sample Image
Tractor with Semi Trailer	Dump	
Tractor with Semi Trailer	Livestock	
Tractor with Semi Trailer	Low Boy Platform	
Tractor with Semi Trailer	Open Top Van	
Tractor with Semi Trailer	Tank	
Tractor with Semi Trailer	Logging	
Tractor with Multiple Trailers	Agricultural Van	
Tractor with Multiple Trailers	Platform/Tank	
Tractor with Multiple Trailers	Bottom/Belly Dump	

Vehicle Category	Body Configuration	Sample Image
Tractor with Multiple Trailers	Van/Platform (Low Chassis)	
Tractor with Multiple Trailers	Enclosed Van	
Tractor with Multiple Trailers	End Dump	
Tractor with Multiple Trailers	Hopper	

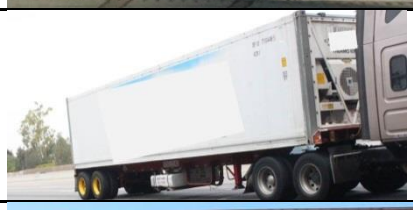
Appendix G – Vehicle Classification Scheme for Combined Weigh-In-Motion and Inductive Loop Signature Classification Model

FHWA Class	Body Configuration	Image
2	Passenger Vehicle	
3	Passenger Vehicle	
4	30ft Bus single rear axle	
4	30ft Bus rear tandem	
4	RV	
4	Van/Platform	
5	12 Pass Van	

FHWA Class	Body Configuration	Image
5	20ft Bus	
5	30ft Bus	
5	Bobtail	
5	Other	
5	Passenger Car	
5	Utility/Platform/Pickup	
5	Van/Platform	
6	Bobtail	

FHWA Class	Body Configuration	Image
6	Bus	
6	Concrete	
6	Dump	
6	Dumpster	
6	FHWA 6 with Trailer	
6	Garbage	
6	Platform/Van/Tank/ Other	
7	Concrete_Tandem_Lift Axle	

FHWA Class	Body Configuration	Image
7	Dump_Tandem_Lift Axle	
7	Dump_Triple_No Trailer	
7	Garbage_Triple_No Trailer	
8	Beverage	
8	Livestock	
8	Low Boy Platform/Drop Frame Van	
8	Platform	
8	Van	

FHWA Class	Body Configuration	Image
9	20ft Intermodal Container	
9	40ft Intermodal Container	
9	40ft Intermodal Container Reefer	
9	53ft Intermodal Container	
9	Agricultural Van	
9	Automobile Transport	
9	Drop Frame Van	
9	Enclosed van	

FHWA Class	Body Configuration	Image
9	Enclosed Van Reefer	
9	Livestock	
9	Logging	
9	Low Boy Platform	
9	Open Top Van/Dump	
9	Platform	
9	Tank	
10	Low Boy Platform/Drop Frame Van	

FHWA Class	Body Configuration	Image
10	Platform	
10	Van	
11	Agricultural Van	
11	Bottom Dump	
11	Hopper	
11	Platform	
11	Tank	
11	Van	

FHWA Class	Body Configuration	Image
12	Bottom Dump	
12	Platform	
12	Van	
13	Multi-Trailer 7 or More Axles	
14	Dump	
14	Livestock	
14	Platform	
14	Tank	

FHWA Class	Body Configuration	Image
15	Unclassified	